ADVANCED INFORMATION TECHNOLOGY (AICITSS)

COURSE MATERIAL
MODULE – I

Board of Studies
The Institute of Chartered Accountants of India, New Delhi

The objective of this background material is to provide uniform reference material to the students undergoing Advanced Information Technology under AICITSS.

All attempts have been made to make the discussion simple and comprehensive. Students may note that the material has been prepared with an objective to help them in acquiring requisite knowledge and skills in the subject and gain hands on experience.

This is also expected to serve as a source of reference book in their future education and training. In case students have any suggestions to make for further improvement of the material contained herein, they may write to Board of Studies, ICAI Bhawan, A-29, Sector 62, Noida. Queries can alsobe sent to : helpdeskadvitt@icai.in.

All care has been taken to provide the material in a manner useful to the students. However the material has not been specifically discussed by the Council of the Institute or any of its Committees and the views expressed herein may not be taken to necessarily represent the views of the Council or any of its Committees.

Revised Edition : February, 2021

ISBN : 978-81-8441-716-6

Price : ₹ 800/- (For All Modules)

Published by The Publication Department on behalf of The Institute of Chartered Accountants of India, ICAI Bhawan, Post Box No. 7100, Indraprastha Marg, New Delhi- 110 002, India.

PREFACE

The impact of Information Technology (IT) on several aspects of accounting profession and practice has been pronounced over the last two decades. The revolutionary developments of various IT tools and techniques have a far reaching impact on theorganizations.

Forensic Accounting and Fraud Detection provide support to Forensic Accountants to perform two broad functions: litigation support and investigation. The main purpose of this topic is to give an idea of Cyber Crime, Legal Aspects of Forensic Accounting to CA students.

Microsoft Excel is undoubtedly most user friendly tool and being extensively used all over the world. The objective of this course is to enable CA Students learn advanced Excel techniques and using them effectively in their day to day work for cleaning & analyzing big data, making quick reports & Charts and also to make dynamic dashboards for top management. Methodology Hands–on exercises in Excel. The Course Intended for CA Students undergoing Articleship in middle & big Multi-National Companies and playing with huge data where excel is used extensively.

The Institute of Chartered Accountants of India has been making earnest efforts to develop a contemporary body of knowledge and skill set for its students by updating its curriculum from time to time.

Inorder to impart synchronized and uniform theoretical and practical knowledge to all the aspiring CA students in the IT area, the Institute has established its own IT Labs equipped with computers of latest configuration, software and other infrastructural facilities at almost all its branches and regional offices.

In addition, Advanced Database Concepts, Computer Assisted Audit Techniques (CAAT) and Advanced ERP Concepts using Tally are the main components which would be covered during the training programme.

This uniform course material has been prepared by Board of Studies of the ICAI in accordance with the course contents covered in the specially designed curriculum to disseminate quality education to its students and the same has been printed by The Board of Studies for distribution to CA students across India and abroad.

We hope that this course material would help the students in building their IT skills which is amust for all, in the current scenario.

CONTENTS

UNIT-1
BASIC CONCEPT OF FORENSIC ACCOUNTING AND FRAUD DETECTION

CHAPTER

1 Basic Concept of Forensic Accounting and Fraud Detection

LEARNING OBJECTIVES

- Understanding the fundamentals of Forensic Accounting
- Frauds and its impact on business
- Identify different types of fraud
- Different methods of Fraud detection
- Fraud triangle and Fraud tree
- Role of Forensic Accountants

Section 1-Basic Concept of Forensic Accounting

1.1 Introduction

Scientific bookkeeping is a fundamental device for revealing or request of budgetary wrongdoing and the bearing of equity, giving unequivocal data about the realities discovered identified with money related wrongdoing. It is another territory, however lately, banks, insurance agencies and even police have expanded taking assistance of scientific bookkeepers. The expansion in cushy wrongdoing and the troubles looked by law implementation offices in revealing misrepresentation have likewise added to the development of the calling.

The legitimacy of the outcomes found by Forensic Accountant relies upon the information, aptitudes, and experience of the measurable bookkeeper. A legal bookkeeper must be fit for coordinating information and aptitudes in the assessment, examination, translation, announcing, and tribute backing of proof. In court, the legal bookkeeper can be a specialist witness, an advisor, or assume different jobs, for example, trier of certainty, uncommon ace, court-selected master, ref, judge, or go between.

1.2 What is Forensic Accounting?

Forensic accounting is a particular field of accounting that explores fraud and examines historical data which can be used in legal proceedings. Forensic accounting is a sensible blend of accounting, auditing, and investigative abilities to conduct financial fraud inquiries. It is useful for court action and analytical accounting.

What does Forensic Accounting cover?

The two major aspects within forensic accounting practices are:

1. **Litigation support services.** A forensic accounting expert measures the damages experienced by the parties implicated in legal disputes and can aid in settling conflicts, even before it reaches the courtroom. In the

event that a conflict reaches the courtroom, the forensic accounting professional could give evidence as an expert witness.

2. **Investigative/fact-finding services**. A forensic accountant must determine whether illegal matters such as employee felony, securities embezzlement (including tampering and distortion of financial accounts), identity theft and insurance racket have taken place.

The scope of forensic accounting is to:

1. Look for evidence of unusual development in the accounting and financial systems.
2. Design accounting processes for verifying important premises and data. A forensic accounting orientation also calls for skills in identifying possible fraud.
3. Perform audit type processes on a routine schedule in order to reduce transaction processing risks.
4. Cover a broad range of businesses and locations that require customary or continuous surveillance of all transaction processing systems.

Types of Forensic Accounting

There are various types of forensic accounting that can take place, and they are typically combined by the types of legal proceedings that they fall under. Below are some of the most common examples:

- Financial theft (customers, employees, or outsiders)
- Securities fraud
- Bankruptcy
- Defaulting on debt
- Economic damages (various types of lawsuits to recover damages)
- M&A related lawsuits
- Tax evasion or fraud
- Corporate valuation disputes
- Professional negligence claims
- Money laundering
- Privacy information
- Divorce proceedings

1.3 General Nature of Forensic Accountancy Work

Forensic Accountants: Scope of work

Although the focus is fundamentally on accounting concerns, the function of a forensic accountant may also cover a more generic investigation including collection of evidence.

Almost every accounting firm in the country today has forensic accounting sections. Within these sections, there may be sub-differentiations; for instance, some forensic accounting experts may specialise in insurance applications, personal injury declarations, fraudulence, construction or royalty audits.

Forensic accountants may also offer their services in retrieving profits from crime and in relation to appropriation proceedings relevant to definite or assumed proceeds of crime or dubious transfer of funds. Unique only to India, there is a specific species of forensic accountants known as Certified Forensic Accounting Professionals.

Forensic Accounting in India

Given the nature and types of fraud in India, the Reserve Bank of India (RBI) has compulsorily made forensic accounting audit mandatory for all banks within the country. The establishment of Serious Fraud Investigation Office (SFIO) in India has become the turning point for forensic accountants in the country. The indications of the growing demand for the field are:

- The growing list of online criminal offences
- Breakdown of regulators to trace and detect cyber-security frauds
- The long chain of co-operative banks going bust

How is a forensic accounting analysis different from an audit?

The difference between the public's expectation of the purposes and objectives of an audit and the CA's responsibilities under Generally Accepted Auditing Standards has been referred to as the "expectation gap". Forensic accounting can help to bridge the expectation gap.

In comparison, forensic accounting and audit differ in specific ways, as shown below:

Sr.No	Particulars	Audit	Forensic Accounting
1.	Objective	Express opinion as to "true & fair" presentation.	Whether any fraud has actually taken place in books.
2.	Techniques	'Substantive' and 'Compliance' procedure. Sample based.	Investigative, substantive or in depth checking.
3.	Period	Normally for a particular accounting period	No such limitations.
4.	Verification of stock, estimation of realizable value of current assets,	Relies on the management certificate/Management representation	Independent verification of suspected/selected items

	provisions/liability estimation, etc		
5.	Off balance-sheet items (like contracts etc.)	Used to vouch the arithmetic accuracy & compliance with procedures.	Regularity and propriety of these transactions/contracts are examined.
6.	Adversed findings, if any	Negative opinion or qualified opinion expressed, with/without quantification.	Legal determination of fraud and naming persons behind such frauds.

In practice, there are differences in mindset between forensic accounting and audit:

- "Investigative mentality" vs. "professional skepticism". A forensic accountant will often require more extensive corroboration.
- A forensic accountant may focus more on seemingly immaterial transactions.
- A forensic accountant will often look for indications of fraud that are not subject to the scope of a financial statement audit.

Section 2-Various Types of Fraud and Related Topics

2.1 What is Fraud?

Fraud is a type of criminal activity, defined as: 'abuse of position, or false representation, or prejudicing someone's rights for personal gain'. Put simply, fraud is an act of deception intended for personal gain or to cause a loss to another party.

The general criminal offence of fraud can include:

- deception whereby someone knowingly makes false representation
- or they fail to disclose information
- or they abuse a position.

Apart from the general meaning let us study some notable definitions of Fraud as per various statutes and standards. Although definitions vary, most are based around the general theme mentioned above.There are numerous definitions for misrepresentation, contingent upon the perspective considering. As per The American Heritage Dictionary, Second College Edition, extortion is characterized as 'a trickery purposely rehearsed to make sure about unjustifiable of unlawful increase'. Davia et al. (2000) rework this in a few things that must be distinguished, while articulating an instance of extortion:

- a casualty
- details of the charming demonstration thought to be deceitful
- the casualty's misfortune
- a culprit (i.e., a suspect)

- evidence that the culprit acted with goal
- evidence that the culprit benefitted by the act(s)

In criminal law, extortion is the wrongdoing or offense of intentionally misleading another to harm them – for the most part, to acquire property or administrations unjustifiably. Extortion can be cultivated through the guide of manufactured items. In the criminal law of customary law locales it might be designated "burglary by misdirection," "theft by stunt," "robbery by extortion and duplicity" or something comparable.

The Companies Act, 2013 defines Fraud in relation to affairs of a company or anybody corporate, to include,

(a) any act

(b) omission,

(c) concealment of any fact

(d) abuse of position committed by any person or any other person with the connivance in any manner with intent to deceive to gain undue advantage from, or to injure the interests of,

- the company or
- its shareholders
- or its creditors or
- any other person

Whether or not there is any wrongful gain or wrongful loss;

- “wrongful gain” means the gain by unlawful means of property to which the person gaining is not legally entitled
- “wrongful loss” means the loss by unlawful means of property to which the person losing is legally entitled.

Fraud is also defined in Para 11(a) of SA 240 issued by the Institute of Chartered Accountants of India – “Auditors Responsibilities relating to Fraud in Audit of Financial Statements” as ‘An intentional act by one or more individuals among

- Management
- those charged with governance
- employees or
- or third parties, Involving use of deception to obtain an unjust or an illegal advantage

2.2 Classifying Fraud

The outline of misrepresentation to 'word related extortion and misuse' is a decent begin to consider the ideal extent of extortion. However still, a further order is advantageous. There are various methods of grouping word related misrepresentation. The order most utilized is where two sorts of extortion are recognized:

(1) Fiscal summary parity extortion and

(2) Resource burglary misrepresentation.

The primary contrast between the previous and the last is that there is no robbery of advantages engaged with the previous. (Davia et al., 2000) Bologna and Lindquist (1995) order extortion on numerous ways, among them misrepresentation for versus against the organization, inside versus outer misrepresentation, the executives versus non-the board misrepresentation and exchange versus proclamation extortion.

Some of them cover the previously mentioned characterization into budget summary parity misrepresentation and resource robbery extortion.

For instance, resource robbery extortion will be misrepresentation against the organization and exchange misrepresentation, without being delegated inner, outer, the executives or non-the board misrepresentation. Different mixes can be made as such.

2.3 What Constitutes Fraud?

Any trickery completed for getting an illegitimate advantage or bit of leeway over another is misrepresentation, in the broadest sense. At the point when such duplicity is completed generally for a business intrigue a monetary benefit or for any close to home

Advantage in some structure or the other to in the long run increase a monetary or business benefit, at that point it is business misrepresentation that we are discussing. At the end of the day misrepresentation is any sort of mischief or harm or unjust act which is endeavoured or completed to the disadvantage of another purposefully.

Extortion includes intentional misleading or deception of realities or potentially huge data to get excessive or unlawful money related favourable position.

There are three clear components to depict misrepresentation:

- Act or oversight of act.
- Damage to another or to the impediment of another.
- Intention.

By and large, be that as it may if the over three components are fulfilled, the extortion is considered to have been executed. It doesn't make a difference assuming in the end, the demonstration doesn't really bring about harming the other or regardless of whether there is no budgetary misfortune. For instance, a buy director who approaches different sellers' citations in a delicate/offer determination exercise may convey touchy and ordered data to a supported merchant to land him the position. In any case if the supported merchant for reasons unknown doesn't land the position, it doesn't imply that the misrepresentation has not occurred. It has only neglected to harm, maybe in the short run time frame. The extortion exists and is lying latent like an infection or microorganisms in a body, sitting tight for a helpful second.

2.4 Types of Frauds

With regards to the circumstance clarified in part I, a contracted bookkeeper might be all around encouraged to contemplate and comprehend specific kinds of cheats and controls which he is probably going to run over. The circumstance and innovation applied may vary in every task, except the center nature fakes will be the equivalent. Coming up next are a portion of the regular fakes which a contracted bookkeeper needs to comprehend:

1. Trojan Horse Frauds:

These are cheats which are submitted in two sections or stages. Where frameworks set down are sensibly acceptable and very much characterized, the fraudster isn't sure whether a demonstration of misrepresentation would be alright for him. In such circumstances, the misrepresentation is done in two stages. The primary stage is that during which the flexibility and quality of a control is tried by the potential fraudster.

In the event that the demonstration is seen or addressed, the fraudster offers a conceivable clarification of faked numbness or a straightforward slip by on his part. In such occurrences, the extortion generally doesn't enter the subsequent stage. In any case, if the demonstration isn't seen or addressed, the fraudster gathers up the solidarity to submit the misrepresentation. The subsequent stage is then actuated during which the real demonstration of harm is embraced after the fraudster is fulfilled that the basic period has passed and that he can proceed with least of hazard.

An ordinary case is the place some assortment of cash is stifled, yet not stashed right away. Assume a bookkeeper gets some cash available to be purchased of garbage or scrap. In a Trojan Horse sort of extortion, he won't pocket the cash promptly, neither will he set up the money receipt. In the primary stage, he would accomplish something like this: He would put the money got in an envelope, notice 'continues of scrap deal' on it in striking letters, mark the date on the envelope and beginning it. He will keep it some place in one of the drawers in the workplace. For a couple of days, maybe weeks or even months, he would state or fail to address it. After an adequately prolonged stretch of time has passed, when he is sensibly sure that the piece deal has been overlooked, (alluded to as 'the basic time frame') the envelope with the cash would be stolen.

This demonstration of taking is the subsequent stage. The subsequent stage is attempted just if the principal stage has yielded agreeable outcomes, for example the framework or strategy has end up being powerless or delicate.

This is a run of the mill Trojan Horse extortion. The name is gotten from the Greek folklore of the Trojan Horse where the Greeks couldn't get through the Trojan guard and they made a blameless looking wooden pony to test the Trojans. The pony was proposed to test the quality of Trojans safeguard. In the event that the Trojans pulverized the pony, the city would have been invulnerable. Be that as it may, tossing alert to the breezes, the Trojans took the pony inside the city and at a helpful time around evening time, fighters who were covered up inside the pony came out and overpowered the Trojans. The Trojan pony was effective.

On a fundamental level, a fraudster utilizes an article, demonstration, or a test as a Trojan pony to assess the degree of hazard or wellbeing in every circumstance. He assesses the aftereffects of the demonstration and afterward concludes whether to proceed. These cheats are executed with impressive arranging behind them and a clever psyche engineering them. The best preventive advances that could be taken are:

a) recognize key controls in each framework and

b) to continually test key controls.

For the initial segment, help from hazard the executives experts in this field might be viable, notwithstanding the standard consistence and inner reviews. As respects the testing of key controls, utilizing tiger groups testing could be successful. These 'tiger' groups test the quality of a control by endeavouring to break it. Different conditions are falsely made, and the framework is tried essentially. Conditions under which the controls don't work or levels at which techniques wither under tension are then inspected for conceivable therapeutic and restorative measures. Such tests empower the administration to assess the strength of controls against potential dangers and presentations and conceivably take fitting measures. Additionally, this tosses a spanner in the plans of a potential fraudster surprisingly. The tiger groups approach has a twofold impact in that, from one perspective, it expands the board readiness and then again, it debilitates the certainty and lessens the wellbeing level for execution of a fake.

2. Disaster Frauds:

These are cheats which flourish with circumstances of debacle, mayhem, turmoil, and confusion. The fraudsters work under the shield of the disarray made in such circumstances. In case of a cataclysm, for example, fire, floods, seismic tremors, and different mishaps, normally the association is reeling from the outcome and stun. It gets illogical or difficult to consent to frameworks and strategies and data and proof can be handily smothered. Information on resource area and whereabouts, shortcomings and qualities of controls and admittance to other touchy data can be utilized or abused. Resources, assets, and data can be taken, sold, harmed, or devastated for ulterior purposes. Similarly, as a patient recouping from a mishap must be incredibly mindful so as to abstain from getting a perilous contamination, so likewise an association must be extremely wary while attempting to settle itself after a calamity.

Take the instance of a distribution center manager who was in a stockroom where there was a colossal fire. There were supplies of electronic things, for example, number crunchers, memory chips, and different things which were lost by fire. When, the fire could inundate the whole stockroom, a few stocks could be rescued. Anyway, the protected probably won't have unveiled any or all the stocks rescued to the insurance agency. That is the reason each protection guarantee must be painstakingly investigated and inspected from a few edges to guarantee that all enlightens fit agreeably. It is unpleasantly basic for a petitioner to expand the guarantee and later clarify it away as a mistake whenever got. The safeguarded additionally has the comfort of expressing that all records are lost or in a confusion and can't be recovered. These sort of catastrophe cheats are the generally straightforward ones, yet there can be more pernicious ones too. The case of the distribution center attendant is a case of a straightforward one, where at the very least, a portion of the stocks might be diminished from the salvageable amounts. Anyway, the malignant fakes are those where a calamity is made as a shield to conceal the extortion previously dedicated. Such a catastrophe can influence an association in a lot of more noteworthy measurements and can take a very long time to recuperate from.

There was an exemplary instance of a misfortune making lodging in a distant area where another supervisor had the option to show benefits. There was an abrupt turnaround and the lodging began making cosmic benefits despite the fact that business conditions were melancholy and room inheritance had not detectably gone up. Incredibly, the benefits were proven through full money receipts; consequently, there was no space for

question in the psyches of the administration. Anyway, what was truly happening was that the administrator was auctioning off costly resources of the inn, for example, light fixtures, artworks, cut glass show pieces, and so on. Exquisite teakwood furniture was supplanted by business pressed wood furniture. Some portion of the cash was steered back to the inn as deal continues to show a rise in the business and increase the certainty of the administration, and some portion of the cash was stashed. At the point when the lodging was completely deprived of every one of its assets, the supervisor chose to set it ablaze to get away from responsibility and conceal his extortion.

Such a misrepresentation is undeniably more genuine and influences the associations' money related position, altruism and can draw in even legal liabilities. One of the best techniques is to assess the sheer coordination's of a circumstance. The hypothesis of opposite rationale clarified in Chapter III is one of the best approaches to recognize such cheats. It advocates application trial of 'difficulty': Are any of the occasions evidently mind blowing and incomprehensible? For the situation above of the lodging, there were tremendous money receipts on long stretches of strike, bandhs, foreboding events, though on days when the director was on leave or sick there were nil deals. Recognition of cheats in a debacle circumstances in an immediate way is generally exceptionally troublesome.

3. Achilles Heel Frauds:

The fraudster applies the rule that no chain can ever be more grounded than its most vulnerable connection. These fakes are the executed by those with a sharp brain. Those fraudsters have an eye for detail. They can perceive the littlest shortcoming or introduction in a framework and endeavour it. Such fakes are fittingly alluded to us the Achilles Heel cheats, named after the unbelievable Achilles, who was powerful however for his heel. It was his heel which was the main weak piece of his body which was assaulted, and which in the end achieved his defeat. A few frameworks are that way.

They might be the most impeccably structured frameworks with controls and systems of the most noteworthy request. In any case, a little shortcoming can permit a misrepresentation of high measurement to happen. To represent this point, an organization had recently gained a bookkeeping framework. As in each bookkeeping framework, there were a few information archives, every one of which had a few fields for data. The administration had paid the product off the rack, on suggestion of an expert, and had not given a lot of consideration to its moment subtleties. Anyway, the bookkeeper did and found that the module for receipts had two fields for dates: the date of 'receipt' of a check and the genuine date of the check called the date of the 'instrument'. For instance, a check might be set up by a borrower on state November one (date of the instrument) however may arrive at the bank by post just on the fifth of November.

The product gave contribution of the two dates to encourage examination of assortments and relating sections in the records. The administration was wilfully ignorant that the bookkeeper was entering the instrument date in the framework to clear a supported borrower's record. Appropriately the indebted person was given a credit reflectively for an instalment made even a month late, simply by dating his check a month sooner than the date of its conveyance.

This qualified the indebted person for get all rewards and the brisk riser motivating forces which he imparted to the bookkeeper. Such little things frequently are unnoticeable yet could have unbelievable effect. In the above situation where the organization had a board of 1000 account holders, the misrepresentation harmed the organization's working cash-flow to a normal of very nearly 20 lacs for every month. What should be possible to

comprehend the conditions in which these sorts of cheats can happen? Each component of the bookkeeping framework must be thoroughly analysed and perceived before review. From the perspective of the administration, altered frameworks are obviously superior to prepared off-the-rack frameworks. Every association has its own qualities and shortcomings, its own way of life, mentality, and worker spirit. Once in a while does an outsider framework adjust consequently. Except if customization is done, a framework can acquire innate shortcomings which can be abused by sharp leaning individuals. Framework testing is basic toward the start as well as for quite a while during genuine execution to learn, comprehend and battle blemishes which surface just on down to earth use.

4. Corporate Espionage:

There was a basic new alumnus enlist junior level agent in the buy and acquisitions branch of a designing organization. His forerunners indicated that he was of a lower working-class foundation drawing a pay of Rs 5,000 and by and large small methods. His dad was owing debtors and had unexpected frailty. One fine day, the agent welcomed his partners to a gathering and reported that he had as of late bought another level.

At the point when the associates went to the gathering, they were paralyzed to see that the level was a tremendous four-room level which had been adorned with marble and rock fittings and costly furnishings. The clarification given was that he had raked in boatloads of cash on the stock trade. Anyway what didn't bode well was that he neither had the cash to make the underlying venture nor the nerve, activity nor the insight to embrace theory. He didn't appear to have a known dealer and couldn't even obviously talk about which scrips he had put resources into and during which periods he had made benefits. Obviously, the stock trade profit was simply a front for some other source. Since he had practically no dynamic power, it was unimaginable that he would have the option to bring in cash through seller fakes or finance fakes or cost fakes. At that point how could he create the cash? A field examination was completed and it was discovered that he was selling organization data. Data about the organization's examination, items, innovation, and techniques, which he had simple admittance to be a saleable item.

Today, data conveys the most noteworthy worth and cheats which offer data about an organization to its rivals are on the ascent and significant time and cash is being spent on this by enormous corporate houses. Such fakes are practically difficult to forestall and prevent. For a certain something, data is immaterial and can't be missed, for example, a stole resource or cash. Also, offices for shipping it through email, web and so on., are effectively accessible. This makes it considerably more unreasonable to screen abuse or burglary of data. No measure of security will totally forestall a fake. The main thing that should be possible is to limit harm. This can be somewhat accomplished by applying the 'distraction technique' . Utilization of bait stockpiling is discovered very to be extremely powerful. Keep delicate data in a few spots, of which just one has the full right data. The odds of the right data spilling out are that path diminished to 33%.

This must be enhanced with consistent watching of all need frameworks to see that there are no security penetrates and infringement, anyway little irrelevant or inadvertent they may appear to be. In one EDP office, it was discovered that there were PC breakdowns which had an example. They happened distinctly during obligation hours of an EDP supervisor. It was later discovered that under the appearance of doing fixes, information access limitations must be loose, and significant information records were duplicated, or harmed or downloaded available to be purchased to contenders. Accordingly, even mishaps must be inspected for causes and vile prospects. Resistance of security rules must be seen intensely to authorize discipline. Finally, unique

instructional courses must be held to teach representatives at all levels with regards to the perils of such corporate undercover work exercises, the organization approaches, and correctional activity.

5. Technical Frauds.

These are fakes which can happen directly before the eyes of the administration and it may not realize that it has been swindled. This is on the grounds that specialized angles are outside the ability to understand and fakes utilizing these as a spread, are troublesome, if not difficult to identify. A misrepresentation of a high size was seen for a situation where a plastic part producing organization utilized the administrations of outer sellers having molds for production of such segments. The crude material was sent to the sellers (disintegrates) to measure and profit the parts based for standards fixed ahead of time. Obviously, the merchant was giving a decent yield for the material sent and the organization had no motivation to whine.

Anyway truly, the seller was not using the crude material completely. Around 90 % of the crude material got was blended in with scrap and prepared. The parity 10 % unused crude material was utilized for individual utilization or sold outside, and the returns took. Since the volumes of creation were high, even a simple 10 % scrap blended added up to a crude material sparing proportional to Rs 25 lacs yearly. The quality control division of the organization didn't have adequately modern instruments to assess the nature of segments delivered. The standards for input yield proportions had not been amended for quite a while. The sellers additionally kept the quality control examiners glad to guarantee a smooth endorsement and acknowledgment of prepared material. Such fakes can occur in any sort of assembling action where the controlling specialists have practically zero specialized information, or where there are not adequate apparatuses to check such cheats.

Here too such cheats are hard for anybody to recognize. In the above outline of the plastic segments, it was during the physical confirmation of stocks which created the primary peculiarity. An enormous heap, clearly of scrap, was seen at the seller's site. This was unordinary since the seller was working together just for the above organization, and according to protection strategies taken out by the organization, just loads of crude material were relied upon to be on the site. The organization likewise affirmed this was not the standard creation scrap amassed over quite a while. It was something other than what's expected. Further, the physical crude material stocks gave a few deficiencies on correlation with records. On inquisitive about both these above issues, the seller was obviously upset. As respects the lack he promptly acknowledged the duty and was set up to make it up to the organization. As respects the heap of scrap, he expressed this was a heap having a place with one of the neighbouring processing plants, permitted incidentally to be kept, to help with their capacity issues. Anyway, the reviewers were not completely persuaded. They took a round in the manufacturing plant and asked about the heap of scrap with one of the laborers. The real truth was out in the open; that piece was second rate crude material to be blended in with the organization's acceptable crude material got for preparing in a specific extent. The secretly spared great crude material was arranged outside and continues pocket.

6. Bank frauds

This area has the highest potential of fraud. The raw material is money itself. Frauds can be perpetrated within a bank itself or by outsiders. Insiders may manipulate funds, loans, and apply teeming and lading between favoured accounts. Outsiders could defraud a bank by furnishing fabricated, duplicated or altered demand drafts, cheques, bills of exchange, and other negotiable instruments. Apart from these borrowers also often

cheat banks in hypothecation agreements by inflating inventories or even providing substandard or spurious stocks with little or no value.

Chartered Accountants may find themselves as auditors, investigators, or a part of the inspection team. These days even pre facility audits are asked to be carried out. These are audits in the garb of investigations to ensure that funds are going into safe and reliable hands

7. Digital Fraud

In recent years, digitization methods have been shown as the main reason for the change in information and communication techniques. Among the reasons for the development of these methods, the fact that the businesses have removed international borders for their economic and commercial activities and switched from traditional forms of business transactions to digital enterprise forms is shown. As a result of the transformations into digital business form, the concepts and practices of ecommerce, e-business and digital accounting have been inevitable. With the development of Internet technology, costs for businesses to generate knowledge and deliver them to users have considerably decreased.

Several high-profile methods include the following:

- Business E-Mail Compromise (BEC): A sophisticated scam targeting businesses working with foreign suppliers and companies that regularly perform wire transfer payments. The scam is carried out by compromising legitimate business e-mail accounts through social engineering or computer intrusion techniques to conduct unauthorized transfers of funds.
- Data Breach: A leak or spill of data which is released from a secure location to an untrusted environment. Data breaches can occur at the personal and corporate levels and involve sensitive, protected, or confidential information that is copied, transmitted, viewed, stolen, or used by an individual unauthorized to do so.
- Denial of Service: An interruption of an authorized user's access to any system or network, typically one caused with malicious intent.
- E-Mail Account Compromise (EAC): Similar to BEC, this scam targets the general public and professionals associated with, but not limited to, financial and lending institutions, real estate companies, and law firms. Perpetrators of EAC use compromised e-mails to request payments to fraudulent locations.
- Malware/Scareware: Malicious software that is intended to damage or disable computers and computer systems. Sometimes scare tactics are used by the perpetrators to solicit funds from victims.
- Phishing/Spoofing: Both terms deal with forged or faked electronic documents. Spoofing generally refers to the dissemination of e-mail which is forged to appear as though it was sent by someone other than the actual source. Phishing, also referred to as vishing, smishing, or pharming, is often used in conjunction with a spoofed e-mail. It is the act of sending an e-mail falsely claiming to be an established legitimate business in an attempt to deceive the unsuspecting recipient into divulging personal, sensitive information such as passwords, credit card numbers, and bank account information after directing the

user to visit a specified website. The website, however, is not genuine and was set up only as an attempt to steal the user's information.

- Ransomware: A form of malware targeting both human and technical weaknesses in organizations and individual networks in an effort to deny the availability of critical data and/or systems. Ransomware is frequently delivered through spear phishing emails to end users, resulting in the rapid encryption of sensitive files on a corporate network. When the victim organization determines they are no longer able to access their data, the cyber perpetrator demands the payment of a ransom, typically in virtual currency such as Bitcoin, at which time the actor will purportedly provide an avenue to the victim to regain access to their data.

Section 3-Red Flags

3.1 What is Red Flags?

How might one survey whether there is misrepresentation? The response to that question is to figure out how to recognize signs or manifestations of misrepresentation. The core of misrepresentation location lies in recognizing the manifestation or pointer of extortion. While analysing a patient, a specialist searches for indications of a disease, side effects, for example, a rash or an ejection or a bubble or changes in temperature or circulatory strain. At the point when he notification such manifestations, he dispatches further examinations through blood tests, sweeps, and X-beams to find out the genuine idea of the disease and the profundity of contamination or harm. In a similar way a reviewer searches for warnings, a term utilized in misrepresentation recognition, for side effects of bad behaviour.

3.2 Concept of Red Flags

In extortion discovery language the term 'warning' is normally utilized. As clear warnings are only side effects or markers of circumstances of cheats. Sap 4 itself gives a couple yet the rundown is interminable. There is consistent exploration being completed over the world. Coming up next are a portion of the warnings which inspectors are probably going to run over and comprehend.

1. Absence of revolution of obligations or delayed presentation in a similar zone:

For instance, in the above case of specialized cheats, where proper or advanced specialized apparatuses for assessment of value are not accessible, there was no arrangement of steady pivot of obligations, revolution of providers and merchants. Further, if there was finished reliance on outside gatherings the warning is much more underlined. Such revolution of obligations ought to be on a 'shock premise', and if conceivable, with newcomers at every possible opportunity. Where indistinguishable production lines or branches are working, pivoting obligations over the branches can create exceptional outcomes. Turn of obligations is a preventive measure.

2. Close nexus with merchants, customers, or outside gatherings:

There would be an irreconcilable situation if a representative, especially at a senior level, were to have close relations with a customer. For instance, if an advance official were to go on a Caribbean voyage with a borrower, all things considered, his fellowship may come in the method of his obligations in regards to the

advance observing commitments. The autonomy of an individual can be obvious from the manner in which he acts with outside gatherings.

3. Sudden Losses:

An organization doing very well unexpectedly makes immense misfortunes. While there could be real reasons, fumble of assets and assets are more probable. These misfortunes are probably going to have been there from the start stewing under window dressed records. When the air pocket blasts, anyway the misfortunes eject, and it seems like unexpected misfortunes have hit the business. These beginnings of the misfortunes could have been guiding of assets, expansion of cost or concealment of pay or a blend of all.

4. TGTBT condition (Green Flags)

TGTBT represents ***Too Good to Be True***. This demonstrates beautiful gleaming report might be outfitted though in genuine terms there are melancholy conditions. The delineation given before where a lodging administrator demonstrated a turnaround of the misfortunes into fabulous benefits were only continues of stealthily sold fixed resources. Viably the inn was quickly watering down its capital and the administrator was helping himself in the deal. What is discovered from this model is that unexpected benefits should be demonstrated. There is no alternate way to progress. All the sparkles are not gold. Another outline is where an EDP director appeared to be an exceptionally committed man. He would be the first to enter the workplace and generally the last to leave. He was exceptionally regarded. Yet, what an examination later uncovered was that he was engaged with controlling information and sure in-house applications for receivables and payables to move credits to certain supported gatherings. Truth be told he was controlling his own credit record to show nil past due despite the fact that he had not paid for a while. The point here is that what gives off an impression of being acceptable should be tried before tolerating reality esteem.

5. Generation of 'vagrant' reserves :

Assets which are held in a guardian limit and for which there is no responsibility are flourishing spots for cheats. Assets gathered by trusts or gifts in real money assortment boxes are commonplace models where there is no responsibility on either side. Neither does the contributor concern himself about the utilization of the assets nor does the recipient have an immediate case or even mindfulness in regard of such assets. Anyway, such assets can be considerable and welcome culprits of extortion. In this way such circumstances are where any agent must explore first.

6. Disaster circumstances:

Mishaps where books have been lost, or harmed, or calamities, for example, fire, seismic tremor, floods and so forth are different spots where fraudsters can eat. The conditions offer ideal covering conditions.

7. Missing Documentation :

This is the surest indication of misrepresentation and basically every circumstance of missing records either has been made to stifle an extortion or if such a circumstance happens to rise it is utilized to design a fake. For instance, a few records have been taken or lost during moving of an office. Purposely the rundown of records moved could have missing things to accommodate a future safeguard against ulterior thought processes. The bigger the association the more prominent the odds of losing or losing records. Chances to sell cheats by

strangely vanishing records are simple and aplenty. In this way missing documentation is a solid sign of conceivable outcomes of misrepresentation.

8. Chaotic conditions :

As a result of fiasco circumstances, conditions where records are financially past due, untidy state or unreconciled, all things considered are misleadingly made. The explanation given regularly is lack of staff or assets, however this is a greater amount of a reason. In a few such circumstances it was discovered that there was no self-assured exertion to increment or proportion the accessible assets. The deficiency was fake and an optical hallucination to permit this tumultuous circumstance which thusly magnificently covered mystery and detestable plans

9. Irrational conduct :

Conduct which isn't happening to the workers' position and which doesn't remember the propriety of an office regularly originates from profound established frailty which could be indicative of deceitful expectations. For instance, an individual who is consistently impolite and discourteous, or excessively clandestine, is probably going to carry on in that manner to smother extortion or tricky act. The goal is kept others under control in order to keep away from coincidentally uncovering blame or to cover some malafide demonstration.

3.3 Concept of Green Flags or the unrealistic' (TGTBT) disorder

While warnings are immediate manifestations of something not right, something hindering to an element's advantage, green banners are obvious markers of something great or supportive to the substance. Both are special cases to a framework yet, one shows alerts while the different acquires a misguided feeling of solace. For instance, deficiencies in stocks, missing reports or missing checks, lack in assortment, modifications in records, and so forth are all 'warning' signs demonstrating that something is out of order and that something is negative to the element. Such warnings straightforwardly suggest the chance of extortion in some structure. Nonetheless, there are different signs which suggest the opposite, a suspicion that all is well and good and comfort, and frequently, such signals don't initiate alerts. In specific situations where the green banners are outrageous in nature, they can likewise be delegated 'unrealistic' condition clarified and showed somewhere else in this book. The main thing which distinguishes green banners is that they are abnormal indications of peculiarities or irregularities, yet obviously innocuous or maybe even supportive

A few instances of green Flags are:

1. One-way mistakes, for example, circumstances where clerks or vendor who show overabundance stocks, however never any deficiencies. The proprietors are maneuvered into a misguided feeling that all is well and good and confidence, yet actually these could be a concealment for more difficult issues, for example, other equal organizations run surreptitiously or some other misrepresentation which is hidden by demonstrating something that the proprietors will like or wouldn't fret.

2. Messy bookkeeper abruptly turns entirely capable and embraces some uncommon work, for example, a long-standing, past due compromise of receivables or payables, which proves to be fruitful.

3. A bookkeeper/worker deliberately pays from his own pocket to compensate for a slip by. This is bound to be misrepresentation and diverting consideration from the focal point which could have more genuine implications.

4. Worker doesn't take progresses/money skim when he goes on outstation visits for organization purposes. He spends from his own assets and presents travel bills for repayment months or even two or after three years. Such an ability to spend colossal sums for the organization and not quickly recouping them doesn't generally originate from respectable thought and love for the organization. Such a limit connotes, that the worker more likely than not is producing pay from other stealthy sources which might be hindering to the organization.

5. Outrageous conduct of being neighbourly or devoted or agreeable or brutal should be taken a gander at with doubt. In one such circumstance, there was a showcase of incredibly cruel conduct by the Quality Control (QC) supervisor in dismissing food rolls fabricated by a subcontractor. The dismissals were extremely serious, nearly to the tune of double the real creation. The misrepresentation was very much idea out and is clarified for the situation concentrate beneath of amazingly brutal conduct.

6. An expansion of point 5 above, pirating great material in the pretense of dismissal. Overwhelming dismissal of stores, extras, and gear in processing plant: a way to pirate out great stock. For instance, acquisition of soybean where the dismissed material really a few layers of good stock underneath the terrible stock has which is emptied. The great material is stacked and secured with dismissal layers on top while coming out. The internal/outward vehicle register indicated a bigger number of trucks going out than those coming in.

3.4 Techniques to discover Red Flags

Is there characterized technique for spotting red or green banner? The appropriate response positive idea there is no single, extraordinary characterized methodology. The strategy would need to be explicit for each kind of circumstance. This capacity to spot warnings k a salt which might be intrinsic in an individual, however it can likewise be developed by careful and steady utilization of sound review methodology. It is a component of good vehicle thinking and liberality Though there can be various circumstances and one of a kind ecological impacts, the accompanying three suggested brilliant guidelines can significantly encourage spotting sed banners in many review circumstances:

- **Rule one**: Acquire information on business and industry however much in detail as could be expected.
- **Rule two**: Assess the sensibilities of the current key interior controls in the business. sensibility of controls ought to be seen from the two limits whether they are excessively frail and excessively solid since either outrageous could demonstrate adverse to the business.
- **Role three**: Apply the COF test-Clues Don't Fittest on the discoveries in rule two above with regards to the business and industry condition. This is finished by comparing all industry and business condition related realities and data along with control irregularities and abnormal review discoveries. Consider the accompanying representation in a scone producing organization
- **Rule four**: Apply a dispassionate judgment for all abnormal amicable suggestions. Sitting late consistently, needing to accomplish work past one's position and force, ready to spend from individual assets for the organization as one of stray occurrences might be okay yet such a large number of them are consistently dubious. Exercise proficient wariness and be dispassionate in taking judgment.

3.5 Case Study of Extreme Harsh Behaviour

There was an enormous global organization ABC having a few auxiliaries and gathering organizations in food preparing and showcasing business. A portion of its fundamental items were jams, squashes, and dairy items. One of its littler subordinate organizations was promoting of scones which were prepared and stuffed by an outside co-packer C. was a little league work specialist who had been helped by ABC in setting up his shed years back. Over the most recent couple of years C has developed sensationally. His little shed. which had begun with just a couple of used machines bought from ABC, presently had mushroomed into a gigantic modernized plant with a few such machines including modernized stopping hardware. CO items premium excellent bread rolls with a decent brand name for ABC. The examiner discovered nothing in the books of record and supporting records that was straightforwardly inconvenient to the organization. Anyway by adhering to the adoration three brilliant guidelines' he had the option to discover a significant extortion by the co-packer C. Applying decide one he discovered that market for bread rolls in the nation was very huge over INR 5000 crores, developing yearly at 15%. 40% of the market was in the disorderly division, which included misleading marked and ether unbranded rolls. Per decide two he saw that the OC control practice was acceptable to the point that half of C's creation was dismissed and borne by C. Per decide 3 he felt that the QC control was too inflexible to be in any way sensible and correspondingly Cs development was excessively quick and to immense to have the option to endure and retain such enormous scope dismissals

Unmistakably business and industry standards demonstrated that such co-packers like C worked an amazingly slight edges which was more baffling. He compared different perceptions and discoveries and according to rules one and two and he inferred that the mass dismissal of creation was conscious and Chad some ulterior intention in tolerating the dismissal docilely with no dissent whenever. His examination uncovered that the enormous sloppy area gave a market to sell such dismissals without any problem. His hunch was demonstrated right. QC supervisor in conspiracy with C was intentionally dismissing enormous amounts of bread rolls on unstable grounds to empower him to sell those at low rates in the disorderly area. He could bear to do this since all economies of scale were given from ABC to C and his general expense of handling was two and the enormous volume of surreptitious deals gave him a sizeable benefit despite the fact that he sold them at much lower rates than the recorded cost.

Subsequently, the previous model demonstrates that industry and business conditions when Juxtaposed with control inadequacies of rigidities could give an evaluator or warnings for additional examination. On the off chance that a reviewer adheres to these straightforward brilliant guidelines his standard review methodology, there are acceptable possibilities that he may distinguish fakes and in bad behaviours, assuming any, which may even rise above books of record. He would he there give important input to the administration or the authority selecting him

Section - 4 Role of Forensic Accountant

4.1 Future Opportunities in Forensic Field

Fakes are of such change and broad impacts that it is unimaginable to incorporate each type or even all significant sorts of cheats. Anyway, certain normal circumstances in which individuals are probably going to end up in have been visualized and quickly portrayed underneath:

1. Conventional Investigation tasks as a continuation of reviews :

These are normal SAP 4 circumstances where the review discoveries have uncovered certain abnormalities and there is a doubt of misrepresentation or blunder. The administration may request that the examiners stretch out their review to apply such broadened or altered techniques as might be important to survey, assess and decide the nature and degree of misrepresentation. This sort of task is a customary examination and needs no elaboration. Such examinations could cover money thefts, resource misfortunes, income spillages through expanded or duplicated solicitations, concealment of pay, swelling of liabilities, flattening of receivables and the rundown could continue forever.

2. Investigations by Statutory specialists:

Examinations in regard of infringement under any arrangement under the Income Tax Act, Companies Act, could be required by any of the particular specialists. Indeed, even Police, CBI, CID, and the Economic Offenses wing could require the administrations of contracted bookkeepers. Such administrations could incorporate assurance of cases from speculators of different sorts, appraisal of assets lost or abused, resistance of endorsed methodology, bank fakes and whatever other financial offense where information on bookkeeping, record-keeping and pertinent relevant laws could be valuable. In the ongoing very much distributed tricks, for example, Harshad Mehta trick, C R Bhansali, Neek Leeson, and Ketan Parikh, huge number of contracted bookkeepers have been approached to give important experiences regarding the nature and procedure of the fakes executed.

3. Bank fakes.:

This zone has the most elevated capability of extortion. The crude material is cash itself. Fakes can be executed inside a bank itself or by pariahs. Insiders may control assets, credits, and apply overflowing and replenishing between preferred records. Untouchables could swindle a bank by outfitting created, copied, or changed interest drafts, checks, bills of trade, and other debatable instruments. Aside from these borrowers likewise regularly cheat banks in hypothecation understandings by swelling inventories or in any event, furnishing unsatisfactory or false stocks with almost no worth. Sanctioned Accountants may end up as evaluators, examiners, or a piece of the investigation group. These days even pre office reviews are approached to be completed. These are reviews in the clothing of examinations to guarantee that assets are going into sheltered and solid hands

4. Business hazard assessment:

This is another region of expert open door for sanctioned bookkeepers. Each undertaking is consistently laden with dangers. What shifts is the degree and degree of the hazard. Take for instance a situation where an organization must attempt another venture for which it requires a huge fund say Rs 100 crores. In the current monetary business sectors, there are a lot of specialists offering a plenty of administrations. Regularly such methods for financing are acquired through experts not very notable to the borrowers and conceivably of questionable certifications. They offer new and untested financing plans through banks or financing foundations or IDBI, or RBI, and so on. In such circumstances once in a while forthright or advance instalment are to be made which run in significant sums. In such conditions either the monetary officials of the organization who could be sanctioned bookkeepers or review firms might be solicited to ask into the achievability from the plan just as the unwavering quality of the specialist. Since the stakes included are commonly high, such tasks offer a

moving open door for contracted bookkeepers to gain the energy about the customers. Comparative circumstances could emerge when another seller, or another customer or another endeavor is to be gone into and the organization needs to guarantee that there is no hazard. In every single such circumstance the primary concern is to guarantee that the customer limits its odds of being hoodwinked.

5. Insurance case fakes.:

Cases for loss of stocks and loss of benefits of huge qualities, especially surpassing Rs 5 crores are normally studied in detail by most insurance agencies. Frequently these cases are swelled, with or without aim. In such circumstances also contracted bookkeepers could be called upon to survey, ask and explore into cheats.

6. Compliance confirmations:

There such a large number of numerous circumstances where explicit rules or orders have been set down for utilization of assets. For instance, a huge trust might be given a gift of Rs 10 crores for an undertaking state accommodating vagrants and widows. The giver may need an affirmation that the assets gave have been properly utilized. It is conceivable this could end up being a flourishing ground for fakes and misappropriation of assets. Likewise, an emergency clinic may have been given assets for a particular ward with conditions. There could be deceptions and bogus reports. A business may have a distant site where certain exercises might be in progress. A chance of abuse of assets is additionally likely.

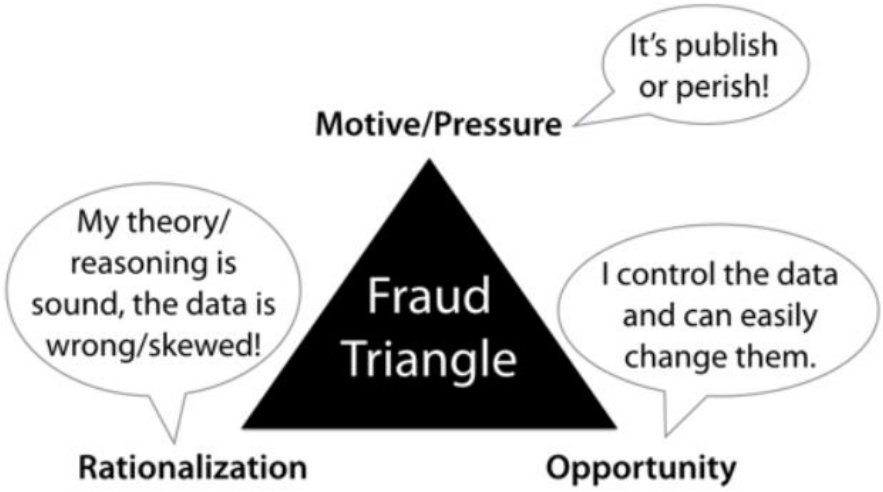
In every such circumstance, depicted above, it isn't essential that contracted bookkeepers could be evaluators. They could be CFOs, financiers, bookkeepers, part of the administration or even specialists. As needs be reference to the part directing a task has been summed up as an inspector (short type of extortion analyst) except if that task was explicitly done as a reviewer, CFO financier or something else. This specialized guide is proposed to give some material to any job that a contracted bookkeeper may end up in a circumstance of extortion location. While it would not be compulsory to apply SAP 4 or 5 in each circumstance, it is absolutely prudent to keep the arrangements set down in that to give a predictable nature of administration. It is likewise basic essential to recollect that there are no normalized misrepresentation identification techniques.

Every task must be completed in a way fitting to the particular needs. What this specialized guide gives is an intelligent and contemplated approach, data about specific strategies and procedures, certain speculations which have been tested and been compelling. They might possibly be compelling in different circumstances. Anyway, a part can attract upon this experience to refine, improve, or adjust these encounters in given circumstances.

In this setting it is applicable to cite the representation of Eklavya from our Puranas. When Dronachary communicated his powerlessness to show him bows and arrows, he was not disheartened. He made a sculpture of Dronacharya and gained from it. The good is that information abides inside ourselves. Whatever be the hindrances in every circumstance one can discover an answer from a limitless store of information existing inside. This similarity can't be pertinent more than in the field of crime scene investigation and extortion discovery. It is dependent upon the analyst to discover the arrangement inside himself, drawing upon his aptitude, experience, and judgment.

4.2 Fraud Triangle

Donald Cressey, a humanist, and criminologist during the 1940s, was one of the main people to work in the field of getting fraudsters and why they do what they do. Cressey stated "Burglary of the Nation," a composition on La Cosa Nostra-a progressively organized perspective on the Organizational Crime and he was broadly known for his examinations in sorted out wrongdoing. Cressey first increased open acknowledgment in this field while finishing his PhD thesis on thieves, at Indiana University. Cressey met almost 200 imprisoned people accused of misappropriation. From his exploration, Cressey created "The Fraud Triangle" which was a model he used to clarify what made a few people become fraudsters by investigating the conditions in which the subjects of his examination were brought into extortion.



Fraud Triangle (by Donald R. Cressey) adapted to Scientific Misconduct

The fraud triangle is used to explain the reason behind a fraud. However, what exactly is fraud?

Fraud refers to a deception that is intentional and caused by an employee or organization for personal gain. In other words, fraud is a deceitful activity used to gain an advantage or generate an illegal profit. Also, the illegal act benefits the perpetrator and harms other parties involved.

For example, an employee that pockets cash from the company's register is committing fraud. The employee would benefit from getting additional cash at the expense of the company.

Below, we discuss the components of the fraud triangle.

The Fraud Triangle – Opportunity

Opportunity refers to circumstances that allow fraud to occur. In the fraud triangle, it is the only component that a company exercises complete control over. Examples that provide opportunities for committing fraud include:

1. Weak internal controls

Internal controls are processes and procedures implemented to ensure the integrity of accounting and financial information. Weak internal controls such as poor separation of duties, lack of supervision, and poor documentation of processes give rise to opportunities for fraud.

2. Poor tone at the top

Tone at the top refers to upper management and the board of directors' commitment to being ethical, showing integrity, and being honest – a poor tone at the top results in a company that is more susceptible to fraud.

3. Inadequate accounting policies

Accounting policies refer to how items on the financial statements are recorded. Poor (inadequate) accounting policies may provide an opportunity for employees to manipulate numbers.

The Fraud Triangle – Incentive (Pressure)

Incentive, alternatively called pressure, refers to an employee's mindset towards committing fraud. Examples of things that provide incentives for committing fraud include:

1. Bonuses based on a financial metric

Common financial metrics used to assess the performance of an employee are revenues and net income. Bonuses that are based on a financial metric create pressure for employees to meet targets, which, in turn, may cause them to commit fraud to achieve the objective.

2. Investor and analyst expectations

The need to meet or exceed investor and analyst expectations to ensure stock prices are maintained or increased can create pressure to commit fraud.

3. Personal incentives

Personal incentives may include wanting to earn more money, the need to pay personal bills, a gambling addiction, etc.

The Fraud Triangle – Rationalization

Rationalization refers to an individual's justification for committing fraud. Examples of common rationalizations that fraud committers use include:

1. "They treated me wrong"

An individual may be spiteful towards their manager or employer and believe that committing fraud is a way of getting payback.

2. "Upper management is doing it as well"

A poor tone at the top may cause an individual to follow in the footsteps of those higher in the corporate hierarchy.

3. "There is no other solution"

An individual may believe that they might lose everything (for example, losing a job) unless they commit fraud.

4.3 The Fraud Diamond

Rationalization is not as intuitive as the pressure and opportunity. Pressure, the private financial need that may be cured by the fraud, and opportunity, the ability to carry out a fraud made possible by a break down in internal controls, are concepts that are easier to get your arms around. The Fraud Diamond seeks to further define and clarify the concept of rationalization.

As recently published by the ACFE and originally published by forensic accountant David T. Wolfe and professor Dana R. Hermanson, rationalizing wrongdoing remains a critical personal characteristic for fraudsters.

The Fraud Diamond extends thinking beyond rationalization to add a fourth model component — **capability** — that considers six other individual abilities and traits that are observable.

These capability factors include:

Having the right organizational position or function to take advantage of fraud opportunities.

Having the appropriate expertise to take advantage of fraud opportunities.

Having the confidence or ego to take advantage of fraud opportunities.

Being able to coerce others to participate in fraudulent activities.

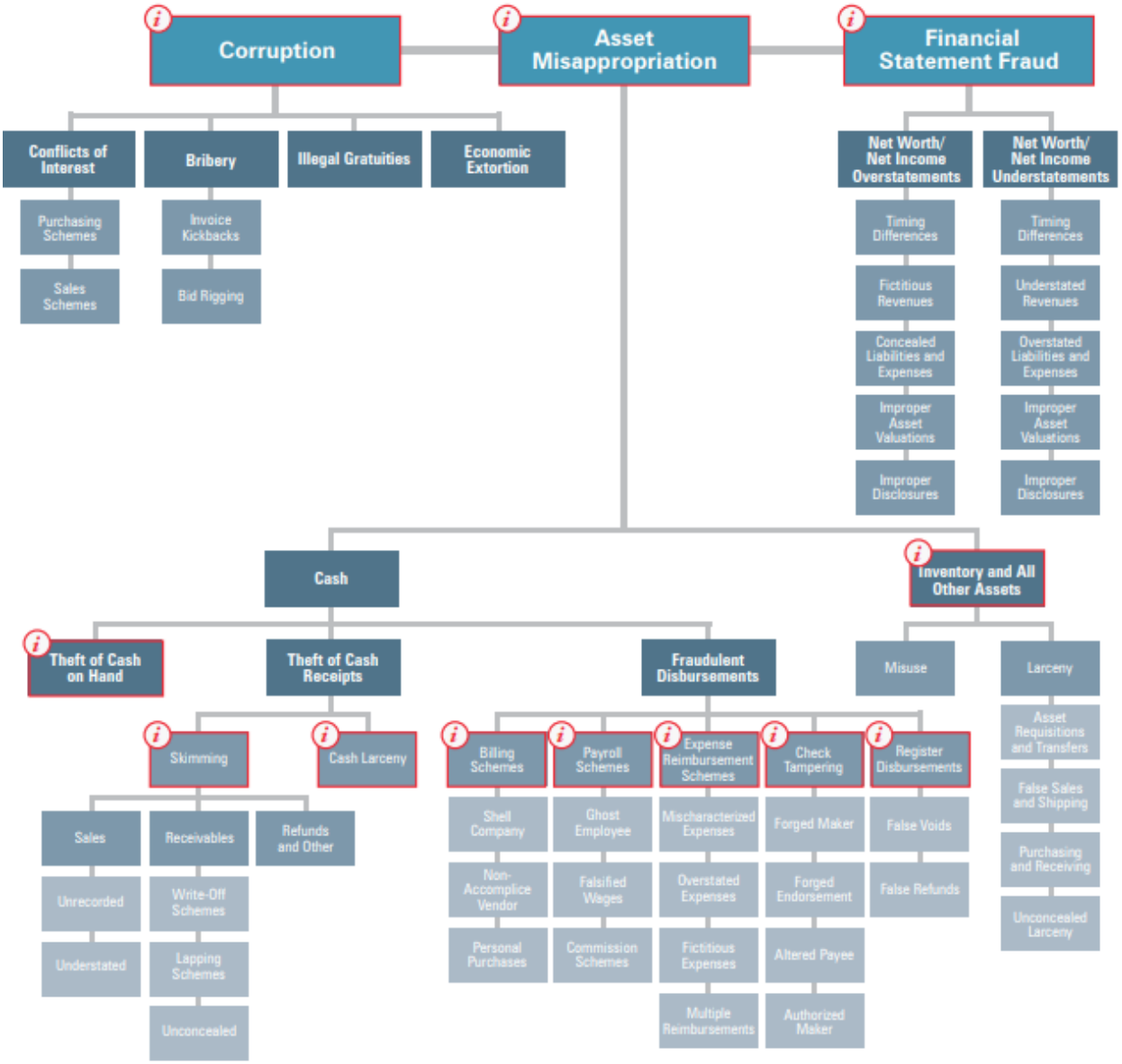
Being able to deal with the stress associated with committing fraud.

Being a good liar.

As organizations design fraud prevention programs, they certainly need to analyze the ability of individuals to "talk themselves into" a fraud, but the six points mentioned by Wolfe and Hermanson provide meaningful depth to that analysis. Anyone can talk themselves into doing the wrong thing in the right circumstances, but individuals with some or all of these six characteristics have a greater likelihood of committing fraud and may require additional safeguards.

4.4 The Fraud Tree

Word related misrepresentation and misuse order framework



1. Corruption

There are three types of corruption fraud: conflicts of interest, bribery, and extortion. Research shows that corruption is involved in around one third of all frauds.

- In a conflict of interest fraud, the fraudster exerts their influence to achieve a personal gain which detrimentally affects the company. The fraudster may not benefit financially, but rather receives an

undisclosed personal benefit as a result of the situation. For example, a manager may approve the expenses of an employee who is also a personal friend in order to maintain that friendship, even if the expenses are inaccurate.

- Bribery is when money (or something else of value) is offered in order to influence a situation.
- Extortion is the opposite of bribery, and happens when money is demanded (rather than offered) in order to secure a particular outcome.

2. Asset misappropriation

By far the most common frauds are those involving asset misappropriations, and there are many different types of fraud which fall into this category. The common feature is the theft of cash or other assets from the company, for example:

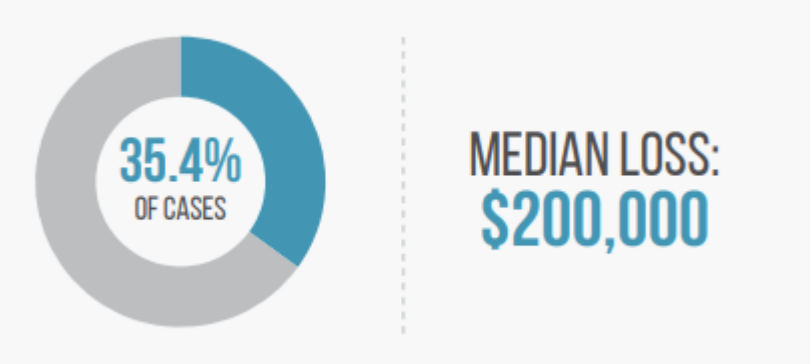
- Cash theft – the stealing of physical cash, for example petty cash, from the premises of a company.
- Fraudulent disbursements – company funds being used to make fraudulent payments. Common examples include billing frauds, where payments are made to a fictitious supplier, and payroll frauds, where payments are made to fictitious employees (often known as 'ghost employees').
- Inventory frauds – the theft of inventory from the company.
- Misuse of assets – employees using company assets for their own personal interest.

3. Financial statement fraud

This is also known as fraudulent financial reporting, and is a type of fraud that causes a material misstatement in the financial statements. It can include deliberate falsification of accounting records; omission of transactions, balances or disclosures from the financial statements; or the misapplication of financial reporting standards. This is often carried out with the intention of presenting the financial statements with a particular bias, for example concealing liabilities in order to improve any analysis of liquidity and gearing.

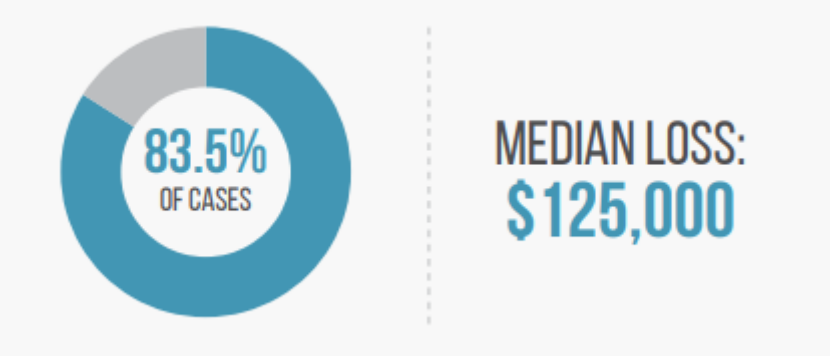
Defilement

A plan where a worker abuses their impact in a business exchange in a way that disregards their obligation to the business to increase an immediate or roundabout advantage



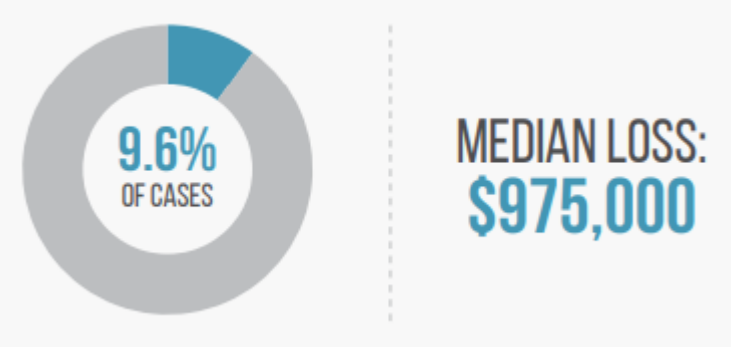
Resource Misappropriation

A plan wherein a worker takes or abuses the utilizing association's assets.



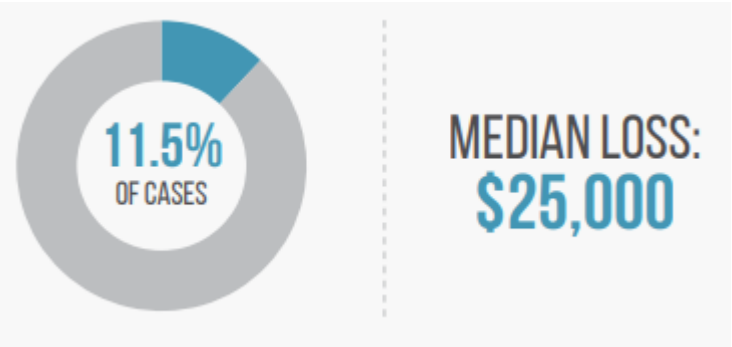
Fiscal report Fraud

A plan where a representative deliberately causes a misquote or exclusion of material data in the association's money related reports. Budget summary Fraud



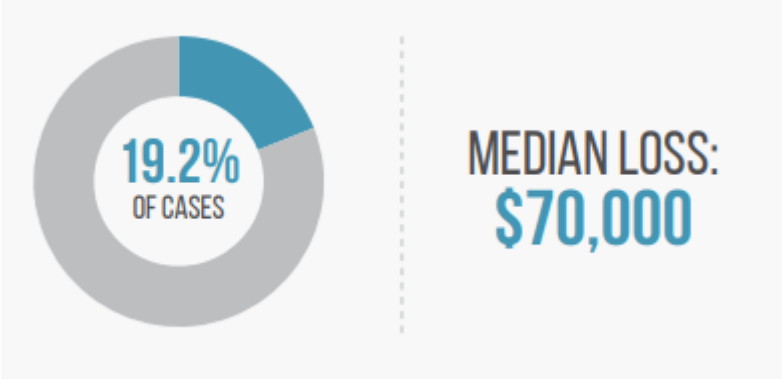
Robbery of Cash on Hand

Robbery of Cash on Hand A plan in which the culprit misuses money kept available at the casualty association's premises.



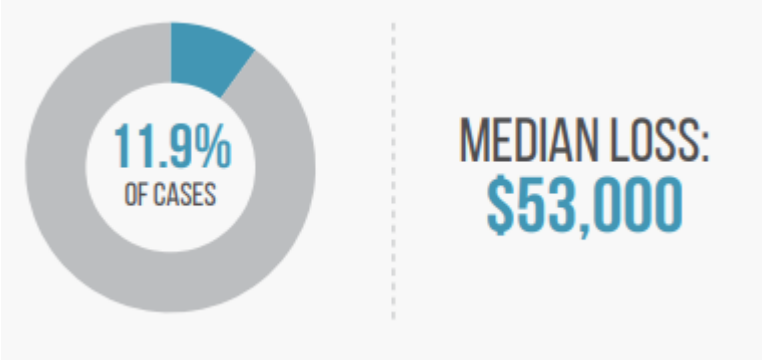
Stock and All Other Assets

A plan wherein a representative takes or abuses non-money resources of the casualty association.



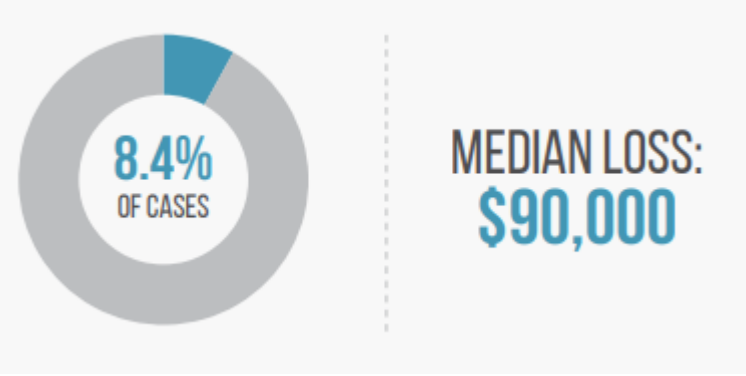
Skimming

A plan where an approaching instalment is taken from an association before it is recorded on the association's books and records



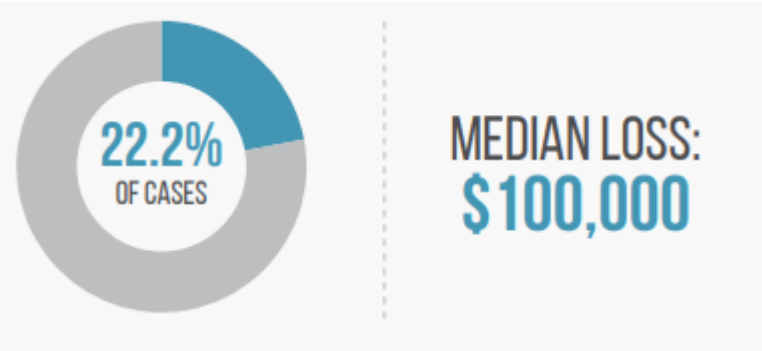
Money Larceny

A plan wherein an approaching instalment is taken from an association after it has been recorded on the association's books and records



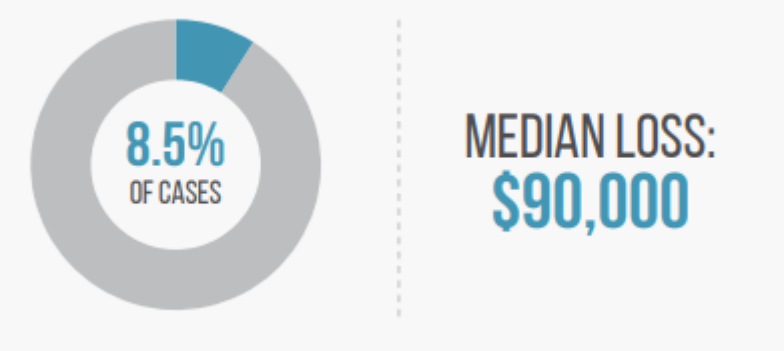
Charging Schemes

A plan where an individual makes their manager issue an instalment by submitting solicitations for imaginary merchandise or administrations, expanded solicitations, or solicitations for individual buys.



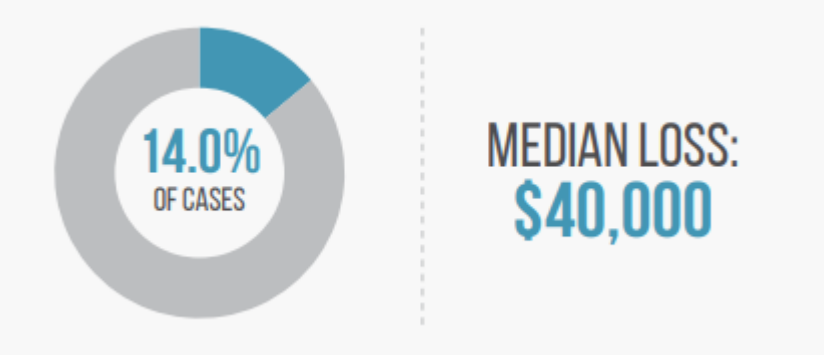
Finance Schemes

A plan wherein a worker makes their boss issue an instalment by making bogus cases for remuneration.



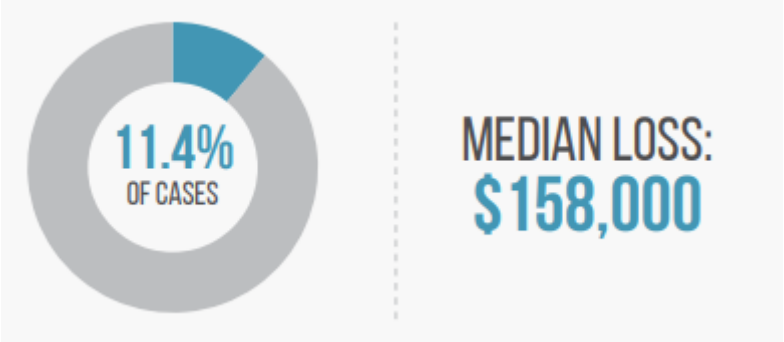
Cost Reimbursement Schemes

A plan wherein a worker makes a case for repayment of imaginary or expanded operational expense.



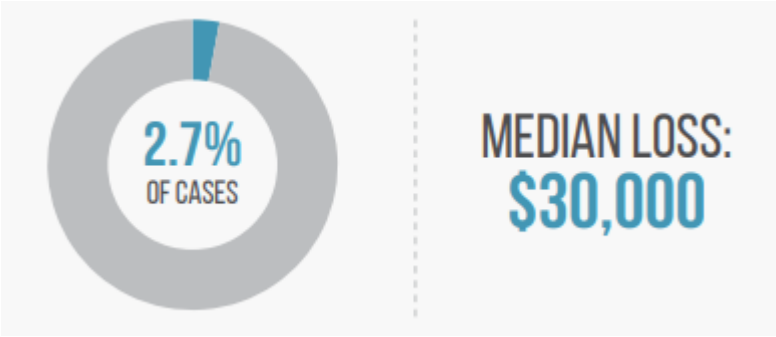
Check Tampering

A plan where an individual takes their boss' assets by blocking, manufacturing, or changing a check or electronic instalment drawn on one of the association's ledgers.



Register Disbursements

A plan wherein a worker makes bogus sections on a sales register to cover the false expulsion of money.



TECHNIQUES INVOLVED

Some of the techniques involved in Forensic Accounting to examine the frauds are:

i) **Benford's Law:** It is a mathematical tool and is one of the ways to determine whether a variable under the study is a case of unintentional errors or fraud on detecting any such phenomenon the variable under study is subjected to a detailed scrutiny.

ii) **Theory of Relative size factor (RSF):** It is a powerful test for detecting errors and measured as RSF = Largest Record in a Subset/Second Largest Record in a Subset . This test identifies subsets where the largest element is out of line with the other elements for that subset. It also highlights all unusual fluctuations which may be routed from fraud or genuine errors.

iii) **Computer Assisted Auditing Tools (CAATs):** These are computer programs that helps the auditor to perform various auditing procedures like testing details, identifying inconsistencies, sampling programs, redoing calculations etc

iv) **Data Mining Techniques:** It is a set of computer-assisted techniques like discovery, predictive modeling and deviation and link analysis designed to automatically mine large volumes of data for new and unexpected patterns.

v) **Ratio Analysis:** This analysis gives indications of the financial health of a company, data analysis ratios report on the fraud health by identifying possible symptoms of fraud

Additional Case Studies

Case Study-1 Bank Fraud Case

Facts

- The fraudster posed as banker's representative to various PSUs and Corporates for making fixed deposits with the bank.
- PSUs and Corporates believed that they would be anyhow be dealing with the Bank and transactions would happen via regular banking channels, hence remain unsuspecting and trusting.
- The fraudster would change his role when he visited the Banks and offered to bring huge deposits from companies. He acted as representative of PSUs and Corporates.
- In the Bank Fraud Case, upon transfer of Rs. 120 cr. by MNO Private Company to the Bank for creating of FD, the fraudster asked the banker to immediately transfer the amount to his accomplice's account KBP International via a FAX communication.
- Immediately, that transfer was followed by another trench of Rs. 80 cr. for FD creation.
- The Bank again received instructions to transfer the amount to the accomplice's account via FAX message.
- Meanwhile, MNO Private Company directly contacted Bank to enquire why FD Receipts were not yet delivered to the company. It was then, they and Bankers realized that fraud has been perpetrated on them.
- The last trench of transfer was saved, but the earlier transfer of money is yet to be recovered.
- In another Bank Fraud case, the Bank received bulk deposits from Corporates and PSUs between 30 January 2017 and 5 May 2017.
- The modus operandi was similar.
- The fraudster represented organizations and as banker to the other.
- Fraudster convinced the organizations to create FD with the Bank and collected the KYC documents, obtaining signatures on forms, filled them and helped them through the entire process.
- Fraudster then submitted forged documents with the Bank and obtained FDRs of Rs. 256.69 cr.
- The originals were collected by the fraudster as representative of the organizations.
- The FDRs were subsequently pledged with the Bank by the 'same signatories' against which overdraft facilities of Rs. 223.25 crore were obtained.
- The funds were then surreptitiously transferred out of bank.
- The Banks also lost their stocks by 5% and stock lost 3.5% upon these revelations.

- Bank FD scam gets bigger, 9 FIRs filed, estimated fraud at least Rs. 700 cr.
- Preliminary inquiry by EOW has been initiated in another 10 cases.
- Finance Ministry has ordered a forensic audit in these PSU Banks.
- Both the banks have complained to the CBI, which has also looking into the matter.

Red Flags

- Middle man
- Audience participation and
- Another involvements

Investigation Process

In addition to the usual,

- Document forensics: Paper and ink analysis, handwriting analysis, signature analysis.
- Following regulatory channels, analyzing the email and telephone call data records.
- Establishing the linkages between the fraudster's location while perpetrating fraud, communications with people within the organization and banks, etc.
- Mapping the procedural role of banking officials and the PSUs-Corporates officials with their job description and the standard operating practices / rules governing their duties.

Case study-2 *Satyam computer* (Satyam)

Satyam was the first major fraud of its kind, which shocked the country and led to tightening of regulations, reporting and governance mechanisms. The fraud had the same shock and awe effect like what Enron and Lehman brothers had in the USA. The enactment of strictest ever regulation, namely, Sarbanes and Oxley, was the outcome of these frauds and many countries followed with enactment of similar regulations.

Promoters of the company had devised ingenious methods to commit frauds with large scale dummy billings for services rendered to foreign clients. As a logical step forward, fake proceeds were shown to have been received in multiple bank accounts, opened in various countries. Many of these accounts were later found to be non-existing.

The company was consistently showing large bank balances in its financial statements, which were not consistent with other IT companies considering the size of its business. The whole of these operations was overseen by the promoter with the assistance of a separate staff working on this, what I would call a fraud factory.

At the closure of financials and to satisfy auditors, fake bank confirmations and statements were generated and produced as evidence of balances to auditors. The amount involved in the fraud was around USD 1 billion.

Surprisingly, Satyam received awards for excellence in corporate governance, conferred by some reputed organisations. Its promoter had over a period acquired respect of the industry and an overwhelming persona. In this background, sudden admission of fraud by the promoter, came as a rude shock to the country,

All said and done, Satyam had a sound business model and portfolio of large international clients. Government had to initiate an unprecedented rescue mission to save the company, by first dismissing the board members of the company, followed by the appointment of professionals as board members led by Deepak Parikh. Ultimately, the company was sold to Mahindra group and is now a major part of the successful technology business of the Group.

Case study-3 *PNB*

PNB was the first major banking fraud reported in the country, involving a massive amount of around Rs. 15000 crores. Fraud was committed by Nirav Modi and Mehul Choksi, (through Gitanjali Gems, a listed company owned by him). Both were in the business of importing rough diamonds and exporting polished diamonds.

Over a period, both had built retail chains of diamond business in India and at famous international destinations. Nirav was, particularly, PR and showmanship savvy.

At that time no one questioned the source of his funding. It was only after a few years, that this unprecedented fraud came to light, which shocked the nation as never before.

He was defrauding PNB and other bankers by opening LCs of large amounts without any underlying transactions (paper money in essence), with the connivance of a few junior level banking officials. He exploited an elementary deficiency in the IT systems of non-reconciliation of LCs opened with the underlying transactions. LCs opened were not recorded in the RTGS system as was the requirement applicable to all banks. Hence, existence of such LCs was not known till the time the fraud was unearthed.

Amounts involved are estimated to be around Rs 16000 crores (including dues of Mehul Choksi). Here again, there were multiple red flags, which were ignored by banks management and regulators, which could have unearthed the fraud much earlier. Periodic inspection reports of RBI, highlighting this deficiency, which were placed before the board, were not actioned, RBI also issued red alert to all banks several times, instructing banks to set right these system deficiencies (mainly RTGC and non-reconciliation). But these also went unattended.

Nirav and Mehul managed to fly out of India and currently India is trying hard in international courts to bring them back to India.

Case Study - 4 *ILFS*

ILFS fraud was the largest corporate fraud in India and triggered a showdown in the economy, as the company was a key vehicle for infrastructure development of the country. Fraud occurred, in spite of marquee shareholders like LIC, SBI etc., being the largest shareholders, having representatives on board. ILFS had the largest debt exposure of around Rs. 91000 crores (including Rs, 20000 crores invested by PF and pension funds),

Fraud was perpetrated mainly by:

- Diversion of borrowed funds to related entities of some of members of top management team

- Imprudent lending to parties who were not credit worth for ulterior motives
- Evergreening of loans by routing money from one group company to another through an unrelated party
- Over invoicing of project costs by vendors, accounting of fake expenses etc and difference being routed back to related entities of some of members of top management team
- Overstatement of profits by non- provisioning of loans, accounting of fake expense, inappropriate recognition of project revenue etc.
- The company had unprecedented number of subsidiaries and group companies, (346) which were used to route above transactions
- Non – disclosure of some of these companies as related parties
- Non-disclosure some of subsidiaries, associates, joint ventures

Most of the mutual funds, insurance companies and PF gratuity funds had invested large sums in its debt issuance, due to the high credit rating of the company. It was a case of negligence by reputed credit rating agencies that rating was not downgraded in spite of clear signs of financial stress in the company. Rating was downgraded abruptly to lowest level from the highest only after the company defaulted in its repayment obligations.

Surprisingly, this public interest entity, was run for years by the same top management team, who were treating ILFS as personal property. Their subordinates and even Board were so overawed by their overpowering persona that no one dared to challenge their decisions. Fraud was going on for years, but could not be detected till the damage was done.

Like Satyam, the government suspended the board and appointed eminent experts to the board chaired by reputed and seasoned banker, Uday Kotak. Currently the company is under resolution process and some of its infrastructure has been sold. However, progress has been slow. Hence, the extent and timing of recovery is uncertain.

UNIT-2
DATA ANALYSIS USING EXCEL AND CAAT TOOLS

CHAPTER

1

UNDERSTANDING DATA ANALYSIS

LEARNING OBJECTIVES

- To understand what is Data Analysis
- To understand the importance of Data Analysis
- To understand ETL in Data Analytics
- To understand various types of Data Analysis
- To understand the concept of Big Data
- To appreciate the difference between Data Analysis and Data Science
- To understand the Traditional and Modern approach to data analysis
- To Understand BADIR Framework to Data Analysis
- To know the challenges in Data analysis

1.1 INTRODUCTION

In this era, to get a better understanding of data to work upon it, we need to analyse it to know its exact impact or effect on our business. The sooner we derive the good and bad effects from it, the more beneficial it will be for the success of the business.

This Data analysis could be aimed at Business growth for which is commonly referred as business analytics when the objective is testing of controls becomes Assurance Analytics, but whenever used for locating anomalies and frauds comes within the ambit of Forensic Analysis, but ultimately it is the insights you get from analysis of Data. For the purpose of our discussion, we will use the generic term data Analysis in this chapter.

Despite the huge volume of data available, we cannot study each part of the data available, we derive our decision based on the data analysed which is merely 0.5% of the total data. While that may not seem like much, considering the amount of digital information we have at our fingertips, half a percent still accounts for a huge amount of data.

With huge data and less time, knowing how to collect, curate, organize and make sense of all this potentially business-boosting information can result to be a boon for the business.

Data helps you understand and improve business processes so you can reduce wasted money and time. Every company feels the effects of waste. It depletes resources, squanders time, and ultimately impacts the bottom line.

Data applies to everything, and everything creates data. Those companies that are most data-driven tend to be the most competitive and productive because having good quality information at your fingertips speeds up the decision-making process and identifies losses before they cause problems.

1.2 DATA ANALYSIS

Data Scientists and Analysts use data analytics techniques in their research, and businesses also use it to inform their decisions. Data analysis can help companies better understand their customers, evaluate their ad campaigns, personalize content, create content strategies and develop products. Data analysis is important in business to understand problems facing an organisation, and to explore data in meaningful ways. Data in itself is merely facts and figures. Data analysis organises, interprets, structures and presents the data into useful information that provides context for the data.

Data analysis is defined as "a process of cleaning, transforming, and modelling data to discover useful information for business decision-making. The purpose of data analysis is to extract useful information from data and taking the decision based upon the data analysis". Thus data analysis is a process that involves methods and techniques from taking raw data, mining data to gather relevant information required for the business's primary goals, drilling down into this information to get some resultants, facts, and figures for improvement of the organisation.

Not only from the perspective of growth, but security and fraud analysis are also desired to protect physical, financial, and intellectual assets from abuse by internal and external threats. Deterrence requires mechanisms that enable companies to quickly identify and anticipate future activity, and more important also identify potentially fraudulent activity, to quickly identify and track down the perpetrators behind it. Efficient data analysis capabilities will ensure better detection and prevention of fraud and other forms of fraudulent activity. Network-based methods for predicting fraud vulnerabilities that lead to alerts, as well as the use of advanced analytical and data mining techniques, will ensure timely responses triggered by automated alerts and mitigations.

Moreover, the integration and correlation of data across companies can provide a more comprehensive picture of a company's financial and financial health and assets. Cross-genre analysis on a data basis provides the possibility to identify weaknesses in fraud checks and investigations. Data management leads to improved risk management, better risk management, and better compliance with regulatory requirements.

ETL in Data Analytics

ETL stands for "extract, transform, and load." The process of ETL plays a key role in data integration strategies. ETL allows businesses to gather data from multiple sources and consolidate it into a single, centralized location. ETL also makes it possible for different types of data to work together. Businesses rely on the ETL process for a consolidated data view that can drive better business decisions.

Types of Data Analysis

Various types of Data Analysis techniques exist involving different technologies and business types. Broad types of data analysis are:

- Text Analysis
- Statistical Analysis
- Exploratory Analysis
- Diagnostic Analysis

- Predictive Analysis
- Prescriptive Analysis

Text Analysis

Text analysis, also known as **text mining**, is the process of recognizing patterns in text data using databases and data mining tools. Text Analysis is about parsing texts to extract machine-readable facts from them. The main objective of Text Analysis is to create structured data out of free text content like emails, social media conversations, survey responses, online reviews, and more.

You can think of it as slicing and dicing volumes of unstructured, heterogeneous documents into easier to manage data points in such a way to extract, examine, and ultimately unearth valuable insights. This is also very useful in Forensic Analysis.

An example could be **Word cloud** tools where you see the result in terms of the number of occurrences and gaining insights into the sentiments.

Statistical Analysis

SAS defines Statistical analysis as "*It's the science of collecting, exploring, and presenting large amounts of data to discover underlying patterns and trends*". Statistical Analysis shows "***What has happened***?" by using historical data in the form of charts and Dashboards. Statistical Analysis could be based on the whole population or a sample thereof and includes the collection, analysis, interpretation, presentation, and modeling of data. It can be categorised into - Descriptive Analysis and Inferential Analysis.

- **Descriptive Analysis (for describing the data)**

  Descriptive statistical analysis as the name implies helps in describing the data. It gets the summary of data ***for example mean, deviation, percentages, or frequency*** in a way that meaningful information can be interpreted from it. Using descriptive analysis, we do not get to a conclusion however we understand and profile the data i.e. we get to know the quantitative description of the data. Descriptive statistics thus enable a meaningful way to present the data. A descriptive analysis is an important first step for conducting statistical analyses.

- **Inferential Analysis (to generalize the population and draw inferences)**

  Inferential Statistics as the name implies is used to generalize the population drawing inference from the samples. Where the sample is drawn from the population itself. It is necessary that the samples should be representative of the population and should not be biased. The process of achieving these kinds of samples is termed as sampling. Inferential Statistics comes from the fact that the sampling naturally incurs sampling errors and is thus not expected to perfectly represent the population. In this type of Analysis, different conclusions can be drawn from the same data by selecting different samples.

Exploratory Data Analysis

It is a good practice to understand data first. Exploratory data analysis as the name implies is the exploration of data set to look at patterns. Exploratory Data Analysis refers to the critical process of performing initial investigations on data to discover patterns, spot anomalies, test hypotheses and to check assumptions with the help of summary statistics and graphical representations.

It is an analytical approach that focuses on identifying patterns in the data and figure out the unknown relationships. The purpose of Exploratory Data Analysis is to get check the missing data, find unknown relationships, and check hypotheses and assumptions.

Diagnostic analysis

Diagnostic Analysis as the name implies is about "***Why did it happen***?". It helps in determining why things are the way they are by finding the cause from the insight found in Statistical Analysis. This Analysis is useful to identify the root cause of the patterns in the data. If a new issue arises in a business process, then you can look into this Analysis to find similar patterns of that problem and probably locate the control weakness which caused the issue and could help in tackling similar issues.

Predictive Analysis

Predictive analysis as the name implies is used to predict future events. It is about "***What might happen***?" It is based upon the historical data and new data and behaviour, using the statistical algorithm and machine learning techniques determines the likelihood of future outcomes, trends. Businesses use predictive analytics to increase the competitive advantage and reduce the risk related to an unpredictable future. Assurance and forensics use it to compare Actuals with Predictions to look at anomalies and variances. An example is **forecasting**

Prescriptive Analysis

Prescriptive Analysis as the name implies tries to answer the question "***What should be done***"? It combines insight from all Analyses we discussed above to identify the best possible action for a situation. The objective of the prescriptive analysis is to provide advice that aims to find the optimal recommendation for a decision-making process. It takes inputs from Descriptive and Predictive analysis, Descriptive describes the data and predictive predicts the outcomes, prescriptive finds the best option. An example is **Simulation, Graph analysis.**

STEPS IN DATA ANALYSIS

There are a few steps involved in DATA ANALYSIS

- What data is required
- Collection of data
- Processing of Data
- Cleaning of Data
- Analysis
- Forming a Result

Let us elaborate on the steps involved for a better understanding of the Data Analysis process

What data is required

The data required for analysis is based on specific demand or establishing the needs of an organisation. Based on the requirements of those directing the analysis process, the data necessary as inputs to the analysis are identified. This is one of the most important data analytics techniques as it will shape the very foundations of your process. Specific variables required may be specified and obtained. Data may be numerical or otherwise based on the different requirements of Analysts.

Collection of data

The collection of data is the process of gathering information that is demanded or identified as data required. The focus is on ensuring accurate and optimum collection of data from a verified source. The decision to be taken depends upon if, data gathered is accurate. You should decide on your most valuable data sources and start collecting your Data.

Data collection sources may range from information in web pages to information derived specifically from database sources. This data may be unstructured and could have irrelevant information which need to be scrubbed, refined, cleaned, and processed to get the desired result.

Processing of Data

The collected data must be processed or structured in an organised form. This includes structuring the data as per requirement of various analytical tools. As an example, there could be a need that data be structured in rows and columns and may be inserted in a Data Model.

Cleaning of Data

The processed and organized data may be having errors, duplications, and unnecessary content. Trimming the informational fat is one of the most crucial methods of data analysis as it will allow you to focus your analytical efforts and squeeze every drop of value from the remaining 'lean' information. These errors can be prevented and corrected through Data Cleaning.

Analysing Data

This processed, clean and organised Data will be used for the analysis. You may use various data analysis techniques available to understand, interpret, and derive conclusions based on the requirements. By integrating the right technology for your statistical method data analysis, you can save ample time and effort while allowing you to enjoy the maximum value from your business's most valuable insights.

Various Statistical Data Models like Correlation, Multivariate Analysis is often wont to identify the relations among the data variables. These models that are descriptive of the data help simplify analysis and communicate results.

Forming a Result

The results of the data analysis would be reported in a format as needed by the users to support their decision making. The users might give their feedback which again could lead to more analysis.

The data analysts can choose among various data visualization techniques, such as tables and charts, which may help in communicating the message easily and much conveniently and efficiently to the users.

1.3 BIG DATA

Big data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis. But it's not the amount of data that's important. It's what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves. Big Data refers to the large amounts of data which is pouring in from various data sources and has different formats. Even previously there was huge data which were being stored in databases, but because of the varied nature of this Data, the traditional relational database systems are incapable of handling this Data. Big Data is much more than a

collection of datasets with different formats, it is an important asset which can be used to obtain enumerable benefits. Big data is data that contains greater variety arriving in increasing volumes and with ever-higher velocity. This is known as the three Vs. Put simply, big data is larger, more complex data sets, especially from new data sources. These data sets are so voluminous that traditional data processing software just can't manage them. But these massive volumes of data can be used to address business problems you wouldn't have been able to tackle before.

The Three Vs of Big Data

Volume	The amount of data matters. With big data, you'll have to process high volumes of low-density, unstructured data. This can be data of unknown value, such as Twitter data feeds, clickstreams on a webpage or a mobile app, or sensor-enabled equipment. For some organizations, this might be tens of terabytes of data. For others, it may be hundreds of petabytes.
Velocity	Velocity is the fast rate at which data is received and (perhaps) acted on. Normally, the highest velocity of data streams directly into memory versus being written to disk. Some internet-enabled smart products operate in real time or near real time and will require real-time evaluation and action.
Variety	Variety refers to the many types of data that are available. Traditional data types were structured and fit neatly in a relational database. With the rise of big data, data comes in new unstructured data types. Unstructured and semistructured data types, such as text, audio, and video, require additional preprocessing to derive meaning and support metadata.

1.4 DATA SCIENCE VS. DATA ANALYTICS

Data is everywhere. We live in a data-driven world that revolves around facts and figures. In fact, the amount of digital data that exists is growing at a very tremendous rate, increasing manifold yearly. According to some researches, it results that in the coming years, about 2 megabytes of new information will be created every second for every human being on the planet, which makes it extremely important to know the basics of the field at least. After all, here is where our future lies.

When we talk about data processing, Data Science vs Big Data vs Data Analytics are the terms that you might think of and there has always been confusion between them.

Let's first start with understanding what these concepts are.

DATA SCIENCE

It is the combination of statistical, mathematical, problem-solving tools available with us. It is the scientific ability to look at things differently. It also involves solving a problem in various ways to arrive at the solution and on the other hand, it involves designing and construct new processes for data modelling and production using various scientific tools. . Data science is a multidisciplinary field focused on finding actionable insights from large sets of raw and structured data. The field primarily fixates on unearthing answers to the things we don't know we don't know. Data science experts use several different techniques to obtain answers, incorporating computer science, predictive analytics, statistics, and machine learning to parse through massive datasets in an effort to establish solutions to problems that haven't been thought of yet.

DATA ANALYSIS

It is defined as the science of examining raw data and processing such data to conclude into some information. It involves applying some defined process to derive information and conclude some results based on it. It is used by major industries to allow organizations and companies to make better decisions. The focus of Data Analytics lies in the process of deriving conclusions that are solely based on the knowledge of the researcher. Data analytics focuses on processing and performing statistical analysis on existing datasets. Analysts concentrate on creating methods to capture, process, and organize data to uncover actionable insights for current problems, and establishing the best way to present this data. More simply, the field of data and analytics is directed toward solving problems for questions we know we don't know the answers to. More importantly, it's based on producing results that can lead to immediate improvements.

1.5 TRADITIONAL APPROACH VS MODERN APPROACH TO DATA ANALYSIS

The traditional approach to data analysis isolates analysis as a specialized reporting activity where data analysts are more working with "clean" or structured data since it is easier to compile, store, organise, and analyse. Traditional data analysis uses more of *descriptive and exploratory analysis* rather than *predictive analysis* to reveal performance results and discover patterns. In many cases, it is a long-drawn-out process involving processing cleaning and structuring data, many a time based on samples with tight assumptions about data, and by the time analysis is done it might be out of date and maybe not relevant. Results are often they are in the form of reports which might not answer all the questions. The traditional approach may also be using tools that also rely on many text-based commands to painstakingly analyse data and build charts.

The modern approach is more about the ability to handle unstructured or "dirty" data and interpret and understand data at the speed of thought. This requires analytical engines to enable manage this heterogeneous and distributed data and provide results that can be optimized to solve a business problem. It uses more of Concepts of Machine Learning (ML), pattern recognition, and Big data techniques. With the help of these modern tools, you can find answers to problems as fast as you think of them while using massive and diverse data sets. With these Modern Tools, the insight-to-decision process is swift. From software installation to access to complex data sets and publishing interactive dashboards, the modern approach saves time at every step of the workflow. These Tools help create a visualisation of your data with simple drag and drop actions.

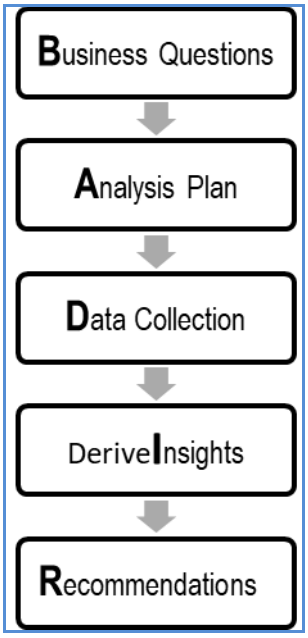
Another aspect of the Modern approach is to use some framework for Data Analysis like Six Sigma Approach, BADIR, and others.

1.6 BADIR FRAMEWORK

B.A.D.I.R is an acronym for the proprietary framework developed by the authors of *Behind Every Good Decision*, Priyanka Jain and Puneet Sharma.

Data Science and Decision Science, progress through these 5 steps and it is believed that analytics fail when any of these steps are not followed or skipped altogether.

This proprietary Data-to-Decisions framework is being widely used in organizations which involves the 5 different major steps defined below:



- Business Questions – identify the business pain points
- Analysis Plan – plan and models are selected with other preparations
- Data Collection – collect data necessary for the analytics planned
- Derive Insights – discover insights from steps 1 to 3
- Recommendations – provides suggestions/recommendations to the management

STEP 1: Business Question

In this we identify business pains and problems, which organization needs to resolve, we may start with asking questions like What, Who, Where, When, Why and How.

Thus, you must understand what you are trying to solve. It is suggested to define a problem by taking root cause analysis. Understanding the context, the impacted segment, and potential reasons as understood by the business may provide a quicker path to resolving the problem. The most important question is "*WHY*" once you know Why then ask other questions for understanding the scope of the problem. This can not only be a business question, but also the Assurance or Forensic question, ultimately it means what is the objective of data analysis.

STEP 2: Analysis Plan

Now when you have defined the problems of the Business, they would be large in scope for implementation, so it is better to break them into smaller achievable parts in order to narrow the objectives to make sure goals are manageable. These goals should follow the S.M.A.R.T methodology – Specific, Measurable, Attainable, Relevant, and Time-bound. It's important to only collect relevant data

Some common analytics methodologies are used to help solve business problems, perhaps the most popular of all analytical models is SWOT – Strengths Weaknesses Opportunities and Threats.

Finally, the analysis model and implementation plan will be concluded in this phase. Also, draft the project plan as it will tie together all of the information gathered and shall serve as the road map for the process execution.

STEP 3: Data Collection

This is the resultant step of the above two steps and is quite often described as the starting point of the process. Efforts put in the first 2 steps shall be the resultant of this step, which may result in less successful returns on the analysis if the first two steps of planning are not done properly. it is needed to collect relevant information (data) within the organizations from different systems or outside the organization such as open data in the market or from government bodies. It may also be noted that analysis does not work if our agenda is not set to the best of findings. The main parts of this step involve:

- Collection of the data following the data specification from the previous step.
- Cleaning data timely and validating it with our requirements to avoid the GIGO (garbage in, garbage out) syndrome, the data needs to be cleaned for usability and validated for accuracy.

STEP 4: Insights

This step is the execution portion of the process where data is being processed accordingly to the methodology prescribed for it. This process helps validate patterns in the data if there is a real problem and if there are unusual patterns in key variables. Then further additional testing is applied to prove or disprove hypotheses. Finally, the findings tested are presented to the user of 0such information. Moreover, it is important to make sure the flow and underlying processes on the insight discovery.

STEP 5: Recommendations

This may be the most crucial step as it is required to take action for the betterment of organisation to develop a positive impact on it. The main purpose of the process is to turn data into insights and then insights into actions. High-level recommendations will be presented to the top managers and values should be demonstrated throughout the presentation. Usually, the recommendations are done by presenting to an audience. Mainly following things need to be achieved

- Engage the audience by presenting a short and summarised set of recommendations.
- Be perceived as an effective business partner by presenting credible recommendations.
- Solve the business problem for which such a process is undertaken by deriving some actions that create impact.

1.7 CHALLENGE IN DATA ANALYTICS

The amount of data being collected

In this digitalized world, a large amount of data is produced every minute. The amount of data produced in every minute makes it challenging to store, manage, utilize, and analyse it. Even large business enterprises are struggling to find out the ways to make this huge amount of data usage.

The amount of data produced is growing each day tremendously. Simply storing huge amounts of data will not be particularly useful, and that is why companies are exploring big data analysis tools that can help them deal with such massive data on a large scale. An organization may call for information on every incident and interaction that takes

place daily, leaving analysts with thousands of interdependent and independent data sets. So, there is a need for a data system that automatically collects and organizes information

Data Validation and Data Cleaning

The second major Challenge is the Data quality problems, with these data is coming from across different platforms the issue is trying to clean data and data validation across these multiple, disjointed sources. Many surveys have found that many of the departments in a large business are not able to use data effectively due to redundancies and data complexity. According to an IBM study, poor data quality costs 3.1 trillion dollars per year in the US.

Validating data is the process of identifying invalid values in your data, this would require knowledge of requirements for data.

Data cleaning is a process that recognizes, corrects, deletes, and stores data from a database or any other data source.

Data cleaning is different from data validation in that validation means data is rejected from the platform at entry and is performed at the same time whereas Data cleaning is performed on group of data at a later stage.

Data quality affects the analysis, you would have heard the GIGO, Garbage in Garbage out, quality of analysis, insights will depend of the quality of data. Data cleaning plays a very important role in quality of Data.

There are various types of inconsistency data issues for example:

- Duplicate data
- Conflicting Data
- Incomplete Data
- Invalid Data

These issues could be caused due to improper synchronisation of data, software bugs in application., Poor application controls, information concealment by the user.

In all the above cases data needs to be cleaned through various data cleaning tools that are available, but Data quality can also be improved with better data validations since data validations will correct data at source and will not allow erroneous data to enter the systems.

Excel is a wonderful tool that can be used to validate and clean the data, in the next chapters this topic will be discussed in detail.

1.8 SUMMARY

Many companies rely on the skills of data analysts to uncover actionable trends or sometimes, red flags; in this chapter, you learned the basics of Data Analysis.

You have learned the differences in terms like Data Science, Big data, and Data analysis as well as the various steps in data analysis.

You have also learnt about various types of data analysis answering various questions like What happened? Why did it happen? What might happen? What should be done? You learned the differences in the traditional and modern

approaches to data analysis. BADIR is a proprietary Data-to-Decisions framework that enables Data analysis outcomes successful and meaningful.

Finally, you understood the challenges involved in data analysis

References

[1] Piyanka Jain, Puneet Sharma, 'Behind Every Good Decision, McGraw-Hill Education,2014

[2] Charles Wheelan, 'Naked Statistics: Stripping the Dread From Data", W. W. Norton & Company,2014

[3] U Dinesh Kumar, "Business Analytics: The Science of Data", Wiley, 2017

[4] Anil Maheshwari, "Data Analytics Made Accessible", Kindle, 2020

CHAPTER

2 EXCEL BASICS AND CUSTOMISATIONS

LEARNING OBJECTIVES

- To understand basic Screen overview
- To understand the Quick Access Toolbar
- To understand Data Types with New rich Data Types
- To How to Customise Options in excel and Trust centre
- To understand the use of Go To Special Feature
- To understand the Benefits of Paste Special
- To manage Worksheets, rows, and Columns in Excel

2.1 INTRODUCTION

Excel is the world's most generally utilized spreadsheet programming and part of the Microsoft Office suite. Other spreadsheet programs like Google Sheets, OpenOffice calc, etc. are available, however, Excel is by a wide margin the most well-known and has been the world norm for a long time as far as spreadsheets are concerned. No matter which career path you choose, you will likely need to use Excel to accomplish your daily professional work.

Taking a straightforward view, Excel is an application that permits you to enter quantitative information into an electronic spreadsheet to apply one or numerous numerical calculations. These calculations eventually convert that quantitative information into data. Using Excel, you can analyse large amounts of data and report those results in a meaningful way

Much of the appeal of Excel is since it very versatile and evolving as per the requirements. Some of the major uses of Excel are:

Calculations: Create Financials, tabulate Numbers, what-if analysis, and perform just about any type of financial analysis.

Text manipulation: Clean data which is text.

Creating Tables: Use the row-and-column layout to tabulate data efficiently.

Creating charts: Create a wide variety of highly customizable charts.

Accessing data from varied sources: Link or Import data from a wide variety of sources.

Creating visualisations: Summarize a large amount of business information in a dashboard

Creating graphics and diagrams: Use Shapes and SmartArt to create professional-looking diagrams.

Automating tasks: Using Macros, many repetitive tasks can be automated.

2.2 EXCEL WORKBOOK AND WORKSHEET

The task you do within Excel is carried out in a workbook file. You may have numerous workbooks open which you require, and each one appears in a separate window. By default, Excel workbooks use an. xlsx file extension, though there are other extensions like .xlsm (Macro enabled workbooks) or .xlsb (Excel binary workbook) etc.

A workbook can be thought of a book containing many sheets of paper and worksheet is akin to a single sheet of paper in that book.

Thus, each workbook can have more than one worksheet, the worksheet is the name given to the large grid where you perform your calculations. Every worksheet has 16,384 columns (A - XFD) and 1,048,576 rows. This means that every worksheet contains 17,17,98,69,184 individual cells. This cannot be changed. The exact number of cells displayed at any one time depends on the size of the screen, screen resolution, and video driver. Each intersection of a row and a column defines a cell and this cell can be identified by its unique address. Each cell can contain a value, a formula, or text.

The worksheet can also hold charts, images, and diagrams. Each worksheet in a workbook is accessible by clicking the tab at the bottom of the workbook window.

Excel is a massive application with a multitude of features and hundreds of ribbon (menu) commands. It is very easy to get lost once you open Excel. So, one of the basic survival skills is to understand how to navigate Excel and access the features you are looking for.

This Figure shows you some important parts of the screen:

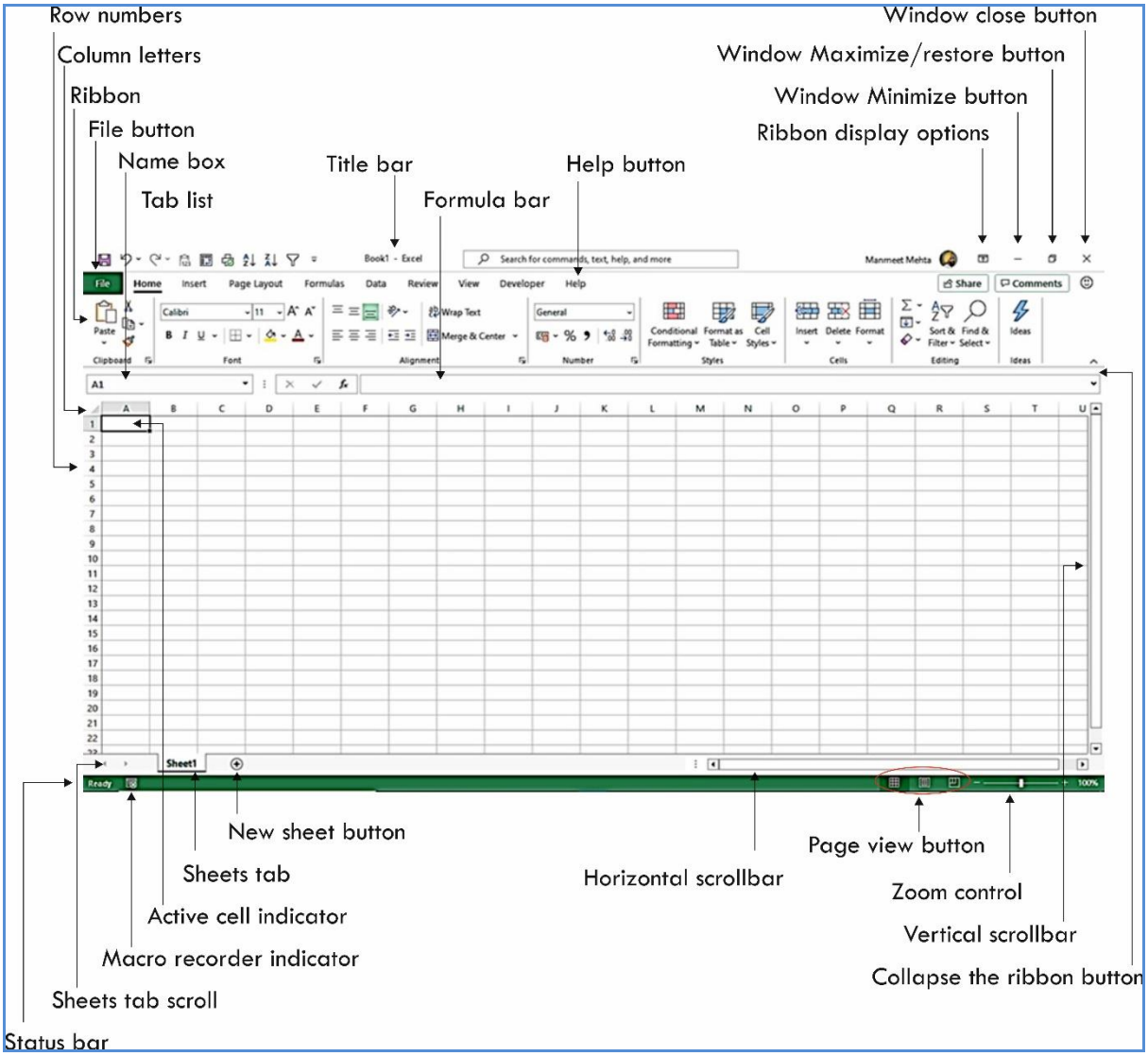


Figure 2.2.1 Parts of an Excel screen

Broadly we can categorise into 5 parts

- **Quick Access Toolbar:** This is a place where all the important tools can be placed.
- **Ribbon:** This is an expanded menu.
- **Formula bar:** This is where any calculations or formulas you write will appear.
- **Spreadsheet Grid:** This is where all your numbers, data, charts & drawings will go.
- **Status bar:** This tells us what is going on with Excel at any time.

2.2.1 PARTS OF THE EXCEL SCREEN

Title bar

This displays the name of the program and the name of the current workbook. It also holds the Quick Access toolbar (on the left) and some control buttons that you can use to modify the window (on the right).

Quick Access Toolbar

As the name suggests it is the place to keep the items that you not only need to access quickly but want to be immediately available regardless of which of the Ribbon's tabs you are working on. When you start Excel for the very first time, it has only a couple of icons (Save, Undo, Redo), you can, however, add any feature of Excel to the Quick Access Toolbar through adding more commands

Tell Me

This is a text field where you can enter words and phrases about what you want to do next and quickly get to feature you want to use or actions you want to perform. You can also use Tell Me to find help about what you are looking for or to use Smart Lookup to research or define the term you entered.

Window Close button

Click this button to close the active workbook window.

Window Maximize/Restore button

This button helps toggle to increase the workbook window's size to fill the entire screen.

Minimize button

Click this button to minimize the workbook window.

The Ribbon

The ribbon is like an expanded menu. It is designed to help you quickly find the commands that you need to complete a task. Commands are organized in logical groups, which are collected under Tabs. Each tab relates to a type of activity, such as Home, Insert, Formulas, Page Layout & Data. To reduce clutter, some Tabs are context-sensitive, shown only when needed. For example, the Pivot Table tab is shown only when a PivotTable is selected.

Tab list

Use these commands to display a different Ribbon, like a menu.

Collapse the Ribbon button

Click this button to temporarily hide the Ribbon. And pin it to make the Ribbon remain visible.

File Menu

Opens the backstage, here you will find the basic commands such as new, open, save, print, export etc.

Help button

Click this button to display the Excel Help system window.

Active cell indicator:

This dark outline indicates the currently active cell (one of the 17,179,869,184 cells on each worksheet). Each cell is represented with a Column and Row address like A1 to XFD1048576

Name box

This box displays the active cell address or the name of the selected cell, range, or object.

Formula Bar

A place where you can enter or view formulas or text.

Expand Formula Bar Button

This button allows you to expand the formula bar. This is helpful when you have either a long formula or a large piece of text in a cell.

Column letters

Letters range from A to XFD — one for each of the 16,384 columns in the worksheet. You can click a column heading to select an entire column of cells or drag a column border to change its width.

Row numbers

Numbers range from 1 to 1,048,576 — one for each row in the worksheet. You can click a row number to select an entire row of cells

Worksheet Navigation Tabs

By default, every workbook starts with 1 sheet. A workbook can have any number of sheets, and each sheet has its name displayed in a sheet tab.

Sheet tab scroll buttons

Use these buttons to scroll the sheet tabs to display tabs that are not visible. You can also right-click to get a list of sheets.

Insert Worksheet Button

Click the Insert New Worksheet button to insert a new worksheet in your workbook.

Horizontal/Vertical Scroll

Allows you to scroll vertically/horizontally in the worksheet.

Status bar

This bar displays various messages, as well as the shows summary information about the range of cells selected. Right-click the status bar to change the information displayed.

Macro recorder indicator

Click to start recording a VBA macro. The icon changes while your actions are being recorded. Click again to stop recording.

Normal View

This is the "normal view" for working on a spreadsheet in Excel.

Page Layout View

View the document as it will appear on the printed page.

Page Break Preview

View a preview of where pages will break when the document is printed.

Zoom Level

Allows you to quickly zoom in or zoom out of the worksheet

2.2.2 NAVIGATING IN A WORKSHEET

There are various ways to navigate the cells in a worksheet. Each worksheet consists of rows (numbered 1 through 1,048,576) and columns (labeled A through XFD). Column labeling is Column A to column Z followed by column AA to column AZ and so on until Column XFD.

The cell address is the intersection of a row and a column, and each cell has a unique address made up of its column letter and row number. For example, the address of the upper-left cell is A1. The address of the cell at the lower right of a worksheet is XFD1048576.

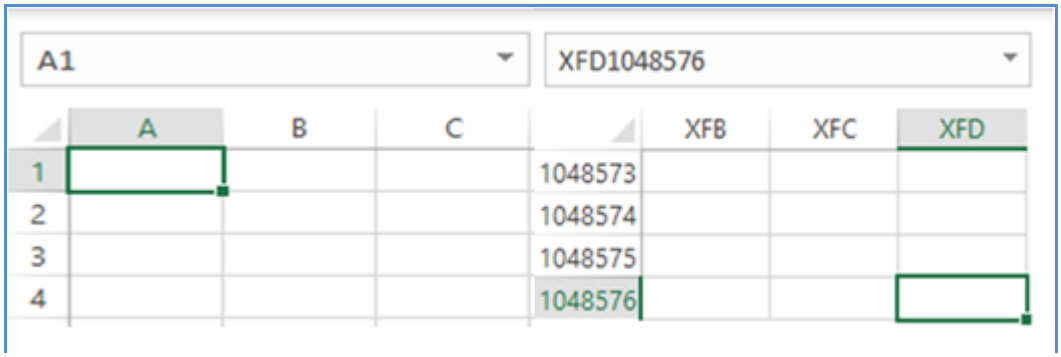


Figure 2.2.2 Cells A1 to XFD1048576

At any given time, one cell is the **active cell**. The active cell is the cell that accepts keyboard input, and its contents can be edited. You can identify the active cell by its darker border, as shown in Figure1.3. Its address appears in the **Name box**. Depending on the technique that you use to navigate through a workbook, you may or may not change the active cell when you navigate.

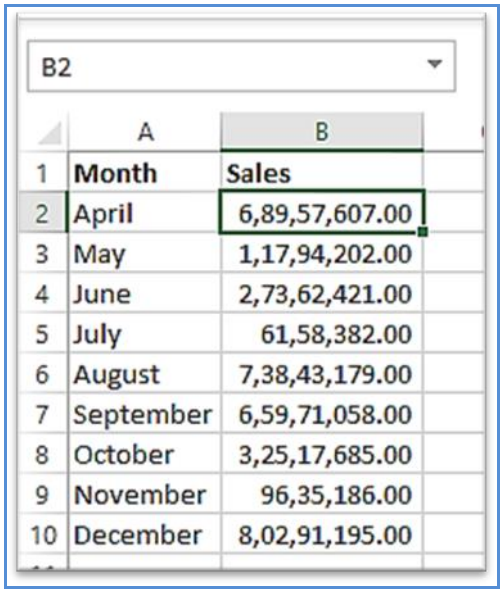


Figure 2.2.3 Active Cell

The active cell is the cell with the dark border — in this case, cell B2

You can use the navigational keys on your keyboard to move around a worksheet. The down arrow moves the active cell down one row, the right arrow moves it one column to the right, and so on. PgUp and PgDn move the active cell up or down one full window. (The actual number of rows moved depends on the number of rows displayed in the window.)

Key	Action
Up arrow (↑)	Moves the active cell up one row
Down arrow (↓)	Moves the active cell down one row
Left arrow (←) or Shift+Tab	Moves the active cell one column to the left
Right arrow (→)	Moves the active cell one column to the right
PgUp	Moves the active cell up one screen
PgDn	Moves the active cell down one screen
Alt+PgDn	Moves the active cell right one screen
Alt+PgUp	Moves the active cell left one screen
Ctrl+Backspace	Scrolls the screen so that the active cell is visible

Using a mouse, you can change the active cell by just click another cell, and it becomes the active cell. If the cell that you want to activate is not visible in the workbook window, you can use the scrollbars to scroll the

window in any direction. To scroll one cell, click either of the arrows on the scrollbar. To scroll by a complete screen, click either side of the scrollbar's scroll box.

2.2.3 EXCEL RIBBON

Microsoft Excel ribbon is the row of tabs and icons at the top of the Excel window that allows you to quickly find, understand, and use commands for completing a certain task.

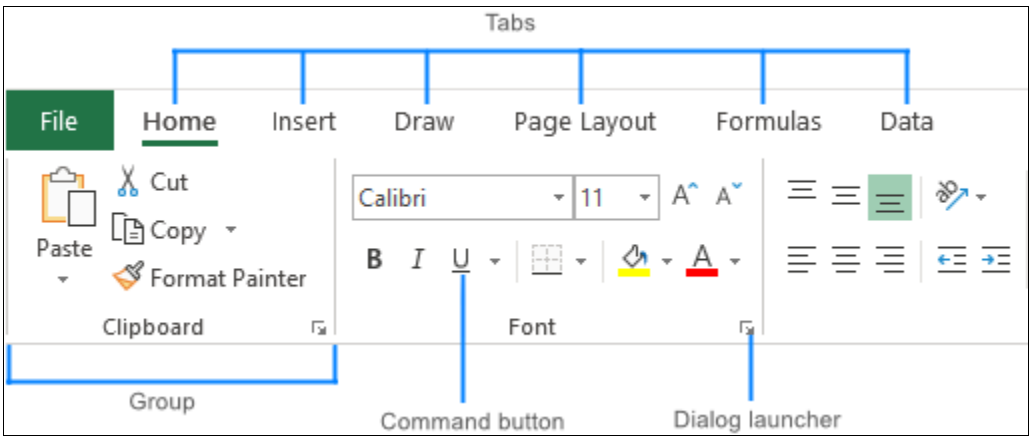


Figure 2.2.4 Excel Ribbon

The ribbon in Excel is made up of four basic components: tabs, groups, dialog launchers, and command buttons.

RIBBON TABS contain multiple commands logically subdivided into groups.

RIBBON GROUP is a set of closely related commands normally performed as part of a larger task.

DIALOG LAUNCHER is a small arrow in the lower-right corner of a group that brings up more related commands. Dialog launchers appear in groups that contain more commands than available space.

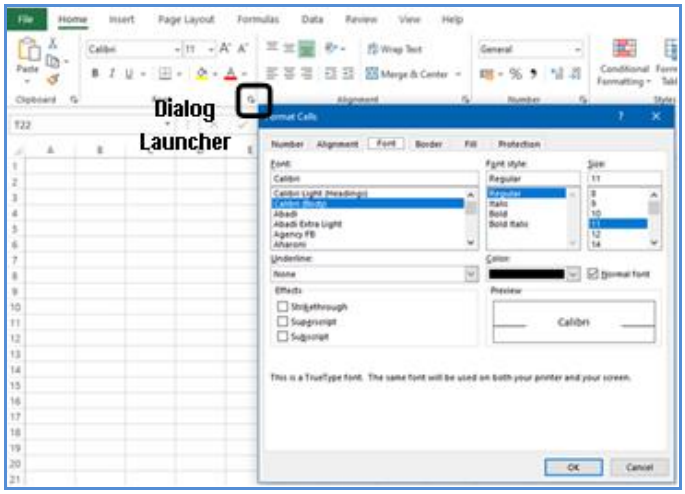


Figure 2.2.5 Dialog Launcher

COMMAND BUTTON is the button you click to perform a particular action. It could be one of the following:

Simple Button: You click, and it does what is required. An example is Cut, you click, and it is cut immediately.

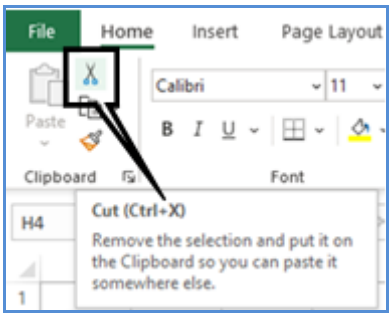


Figure 2.2.6 Simple Button

Toggle Button: This kind of a button shows has two states if it is activated has a different colour, so when you click it you can toggle between activated not activated. An example is Wrap Text, if it is coloured you can tell that text wrap is activated and you can deactivate by clicking on the button again.

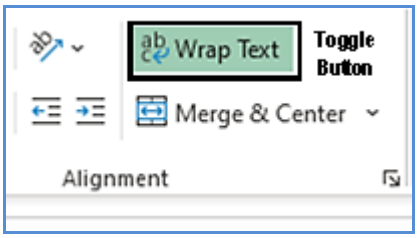


Figure 2.2.7 Toggle Button

Simple Drop Down: In this Ribbon could have a small drop-down arrow, which when clicked gives additional commands. An example is Insert, which when clicked gives more options.

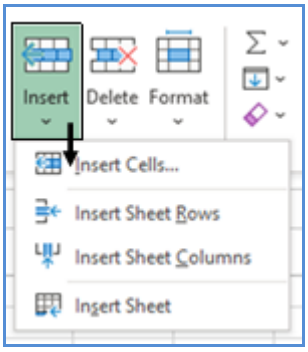


Figure 2.2.8 Drop Down Button

Split Button: This is a combination of a button with a small drop-down arrow beside it, If you click on the button command gets executed, if you click on the dropdown gives additional commands. An example is Underline, which also has a dropdown beside it which when clicked gives more options.

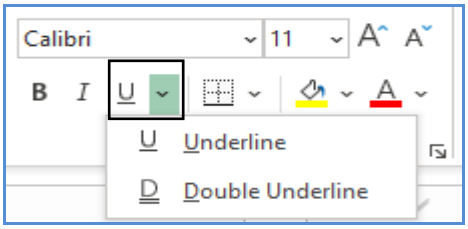


Figure 2.2.9 Split Button

2.2.4 RIBBON TABS

The commands available in the Ribbon vary, depending upon which tab is selected. The Ribbon is arranged into groups of related commands. The appearance of the commands on the Ribbon varies, depending on the width of the Excel window, some of the commands may be grouped and need to be expanded if the width is too narrow. The standard Excel ribbon contains the following tabs, from left to right:

File – allows you to jump into the backstage view that contains the essential file-related commands and Excel options.

Home – contains the most frequently used commands such as copying and pasting, formatting, etc.

Insert – is used for adding different objects in a worksheet such as images, charts, PivotTables, hyperlinks, Sparklines, headers, and footers.

Draw –It lets you draw with a digital pen, mouse, or finger. This tab is available in Excel 2013 and later, but it is not visible by default.

Page Layout – provides tools to manage the worksheet's appearance, both onscreen and printed. These tools control theme settings, gridlines, page margins, object aligning, and print area.

Formulas – contains tools for inserting functions, defining names, and controlling the calculation options.

Data – holds the commands for managing the worksheet data as well as connecting to external data.

Review – allows you to check spelling, track changes, add comments and notes, protect worksheets, and workbooks.

View – provides commands for switching between worksheet views, freezing panes, viewing, and arranging multiple windows.

Help – only appears in Excel 2019 and Office 365. This tab provides quick access to the Help Task Pane and allows you to contact Microsoft support, send feedback, suggest a feature, and get quick access to training videos.

Developer – provides access to advanced features such as VBA macros, ActiveX and Form controls, and XML commands. This tab is hidden by default and you must enable it first.

2.2.5 CONTEXTUAL RIBBON TABS

In addition to the constant tabs described above, the Excel ribbon also has **context-sensitive tabs**, alias Tool Tabs, which show up only when you select a certain item such as a table, chart, shape, or picture. For example, if you select a Table, the Table Design tab will appear.

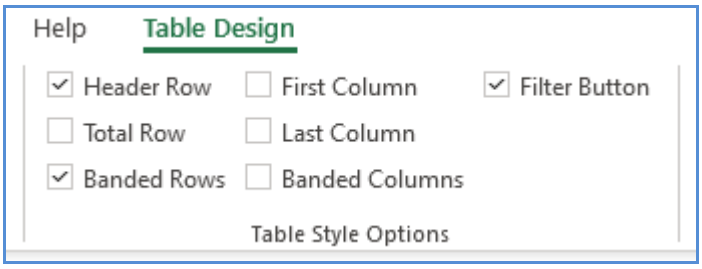


Figure 2.2.10 Contextual ribbon

2.2.6 HIDING THE RIBBON IN EXCEL

You can hide the Ribbon can either be hidden or made visible to toggle the Ribbon's visibility, press **Ctrl+F1**

You can hide the ribbon completely by clicking the Ribbon Display Options button at the upper-right corner of the Excel window, and then clicking Auto-hide Ribbon.

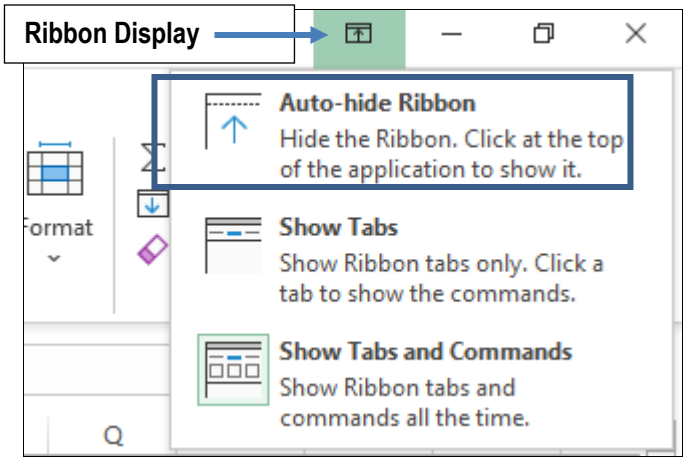


Figure 2.2.11 Hide Ribbon

2.2.7 CUSTOMISING THE RIBBON

You can personalise the ribbon according to your needs, so you know exactly where everything is located.

Method 1: File > Options > Customize the Ribbon

For this, you need to navigate to the Customize Ribbon window under Excel Options.

Method 2: Right Click on Ribbon

And the shortest path to it is to right-click on the ribbon and select Customize the Ribbon

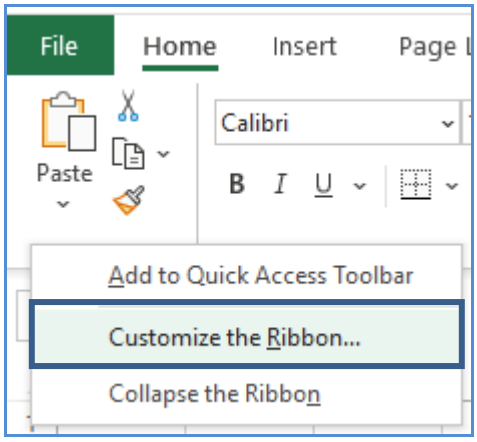


Figure 2.2.12 Customising Ribbon

From **Customise the Ribbon**, you can add your tabs with any commands you choose, change the order of tabs and groups, show, hide, rename tabs, etc.

What you can customise:

- The show, hide and rename tabs.
- Rearrange tabs, groups, and custom commands in the order you want.
- Create a new tab with your own commands.
- Add and remove groups on existing tabs.
- Export or import your personalized ribbon.

What you cannot customise:

- You cannot change built-in commands, their names & icons.
- You cannot resize the Ribbon of the text size of commands.

2.2.8 CREATE A NEW TAB FOR THE RIBBON

You can add and create your own tab to the Excel ribbon.

- In the **Customize the Ribbon** window, under the list of tabs, click the **New Tab** button.
- This will create a new tab named **New Tab (Custom).**
- Rename the Tab by clicking on **Rename.**

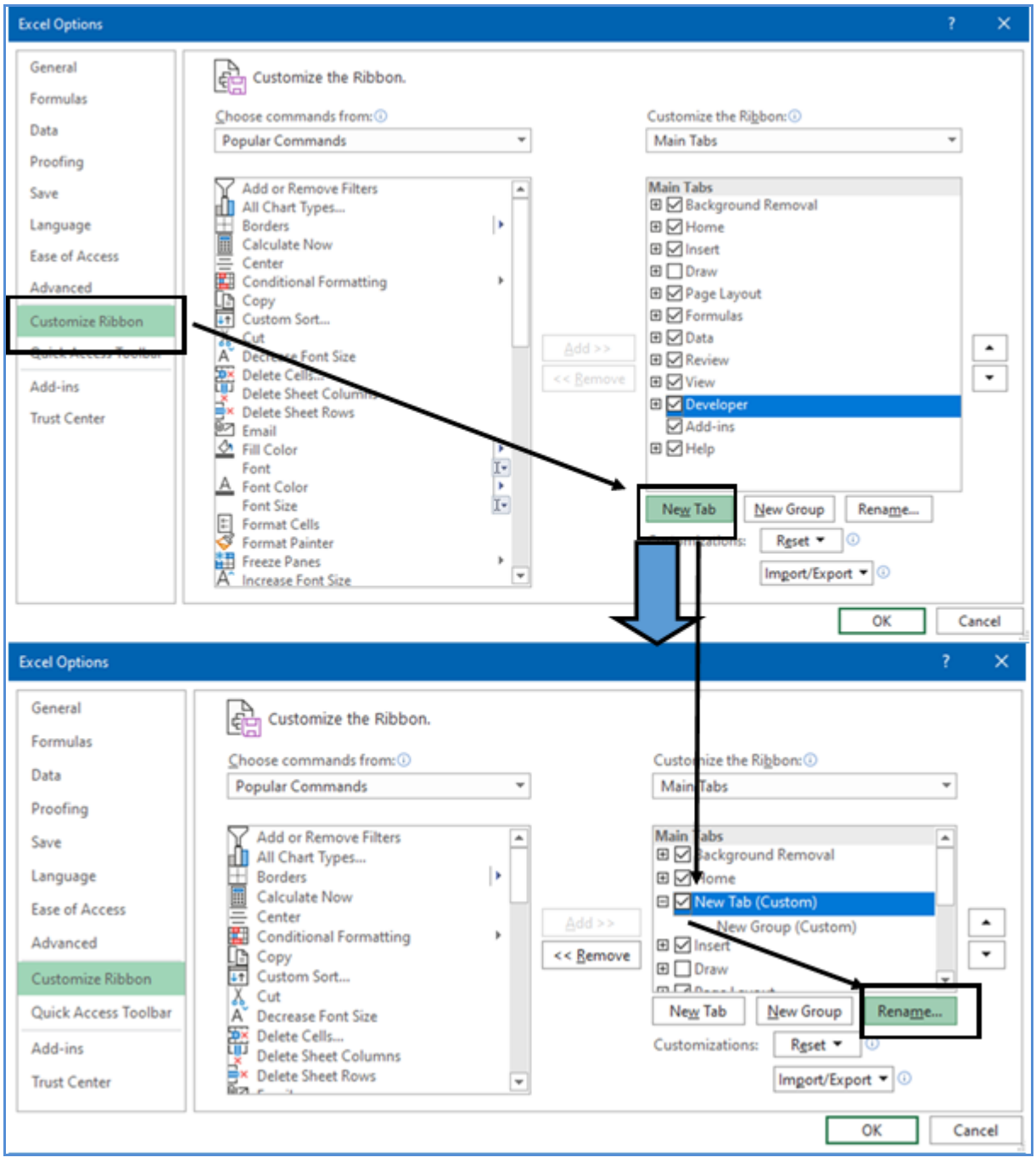


Figure 2.2.13 Creating a new Tab

2.2.9 ADD A CUSTOM GROUP TO A RIBBON TAB

You can also add a new group to either a default or custom tab, as follows:

- In the right pane of the Customize the Ribbon window, select the tab to which you would like to add a new group.
- Click the New Group button. This adds a custom group, named **New Group (Custom),** at the bottom of the list of groups.
- To create a new group in a specific location, **select the group** after which the new group is to appear.

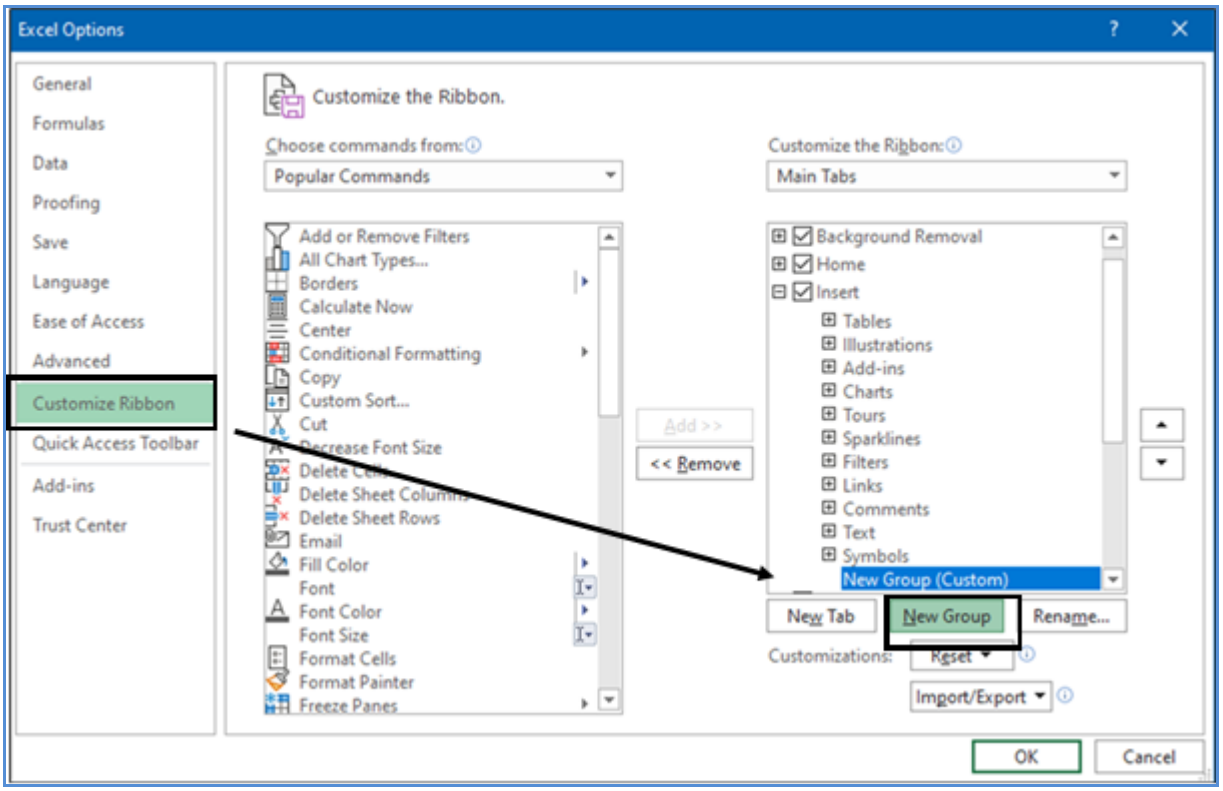


Figure 2.2.14 Create Group

You can rename this group through **Rename** option, also select an icon as shown below:

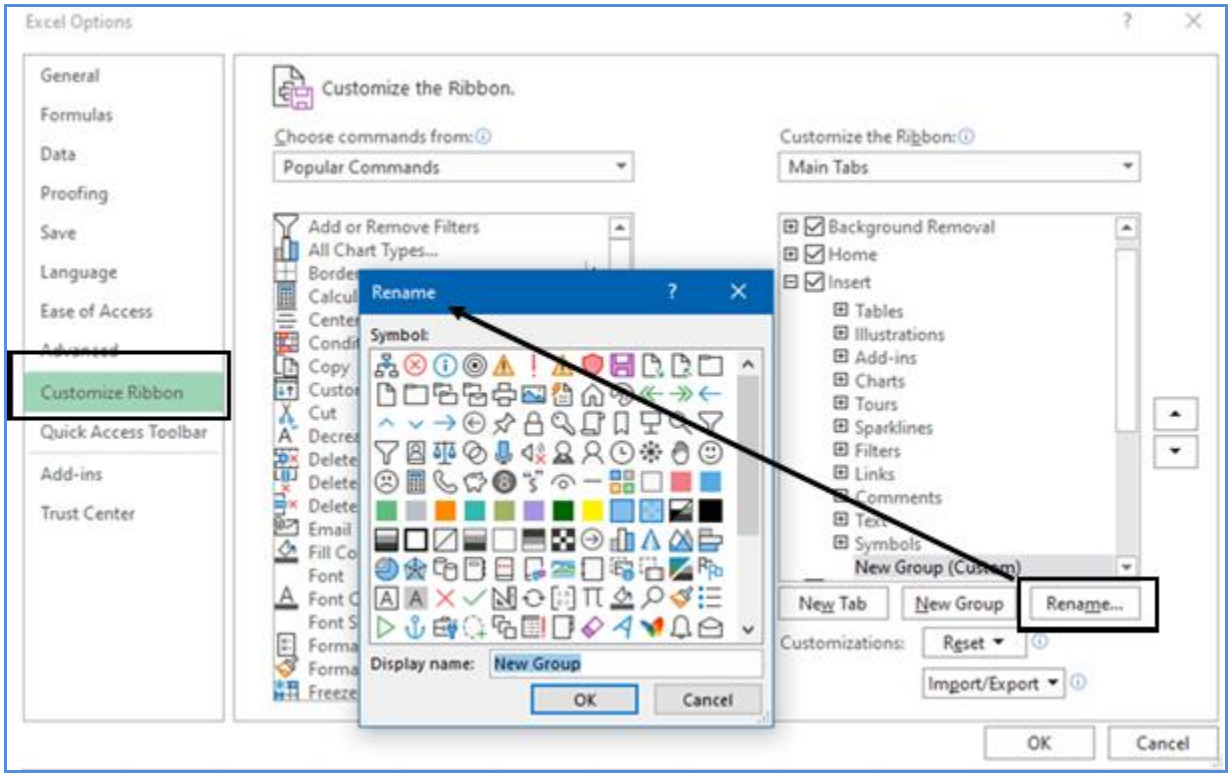


Figure 2.2.15 Rename group

ADD A COMMAND BUTTON TO EXCEL RIBBON

Some Commands like **Commands not in Ribbon** can only be added to custom groups as follows:

- Create a custom group on an inbuilt or custom tab.
- In the list under **Customize the Ribbon**, select the target customer group.
- In the Choose commands from the drop-down list on the left, select the list from which you want to add commands, As an example **Commands Not in the Ribbon**.
- In the list of commands on the left, **click the command** you want to add.
- Click the **Add** button.
- Click **OK** to save the changes.
- As an example, we have added a Share Workbook (Legacy) command to Ribbon

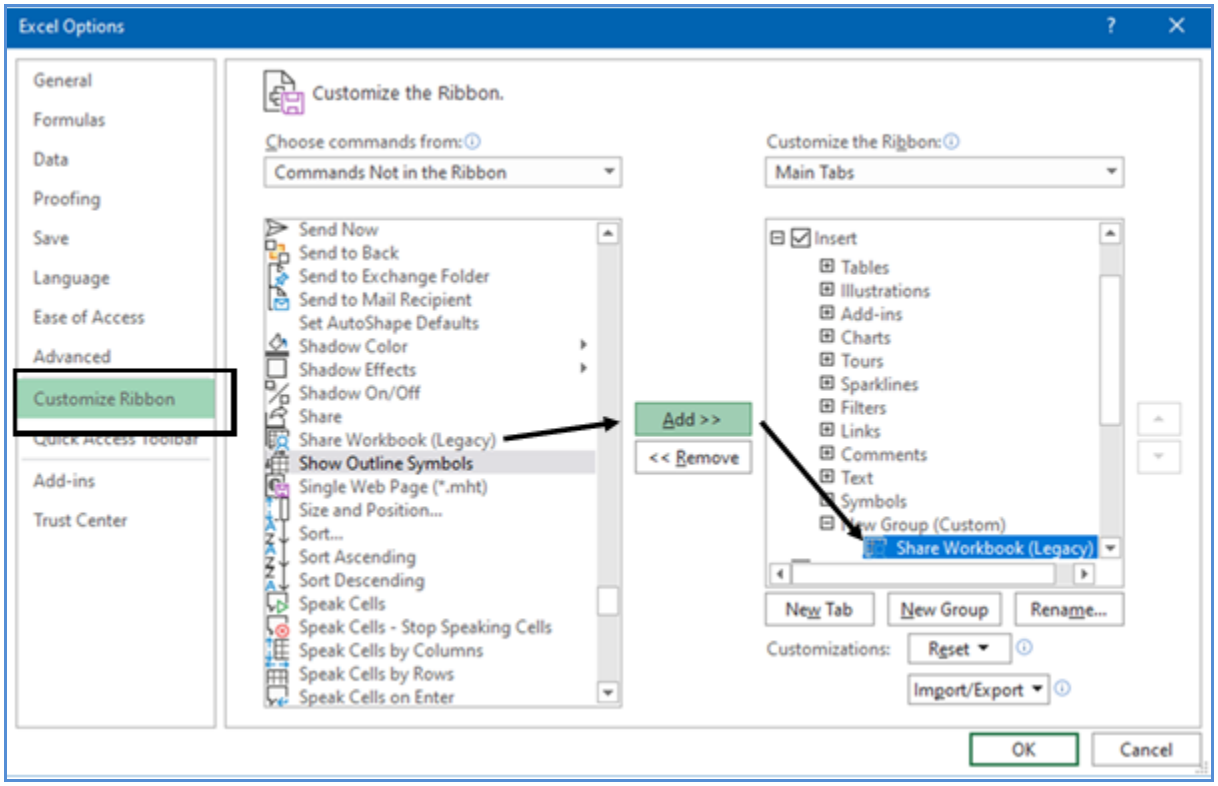


Figure 2.2.16 Add Command Button

2.2.10 RENAME RIBBON TABS, GROUPS, AND COMMANDS

You can give your own names to custom tabs and groups that you create, also you can rename the built-in tabs and groups.

Spoiler: You cannot change names of the inbuilt commands, only commands added to custom groups can be renamed.

- On the right side of the **Customize the Ribbon** window, click on the item you want to rename.
- Click the **Rename** button below the list.
- In the **Display name** box, type the name you want, and click **OK**.
- Click OK to close the Excel Options window and view your changes.

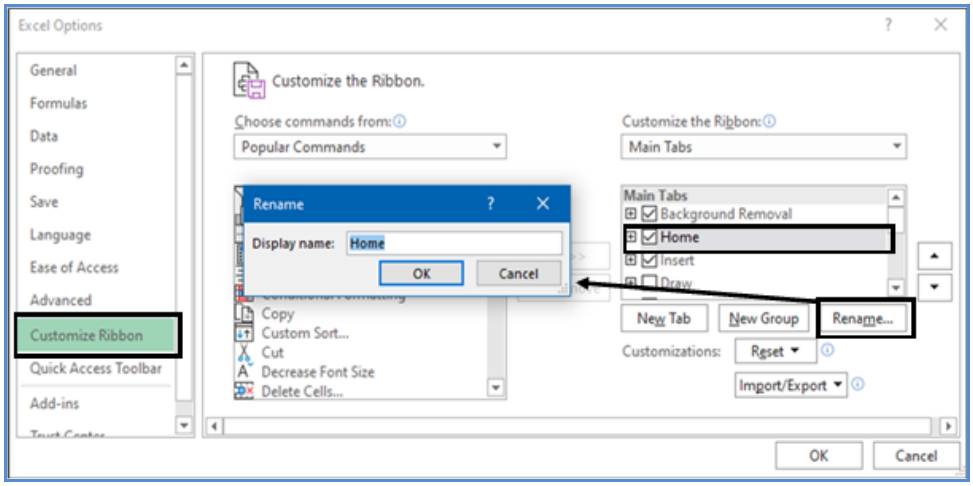


Figure 2.2.17 Rename Ribbon Tabs

2.2.11 MOVE RIBBON TABS, GROUPS, AND COMMANDS

You can move tabs and groups on the Excel Ribbon in the most convenient places. However, the build-in commands cannot be moved, you can only change the order of commands in custom groups.

- In the list under **Customize the Ribbon**, click on the tab, group, or command in a custom group that you wish to move.
- Click the **Up or Down arrow** to move the selected item left or right on the ribbon, respectively.
- When the desired order is set, click **OK** to save the changes.

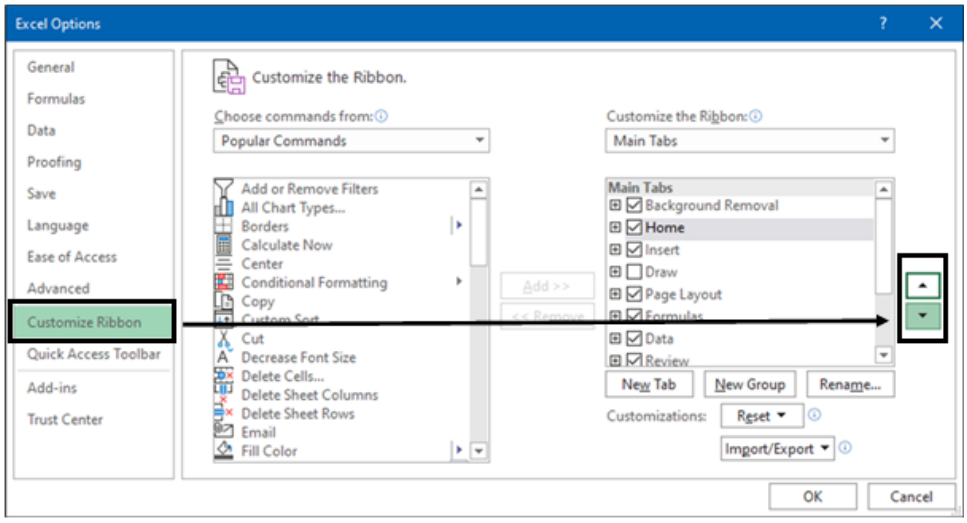


Figure 2.2.18 Moving Ribbon Tabs

2.2.12 HIDE AND SHOW TABS ON THE RIBBON

You can hide from view any tabs that you never use. Similarly, you can show whenever you want.

- To hide a ribbon tab, simply **uncheck its box** in the list of tabs under **Customize the Ribbon**,
- Then click **OK**.
- To show a ribbon tab, **select the box** next to it, and click OK.
- You can show the Developer tab, which is not visible in Excel by default.

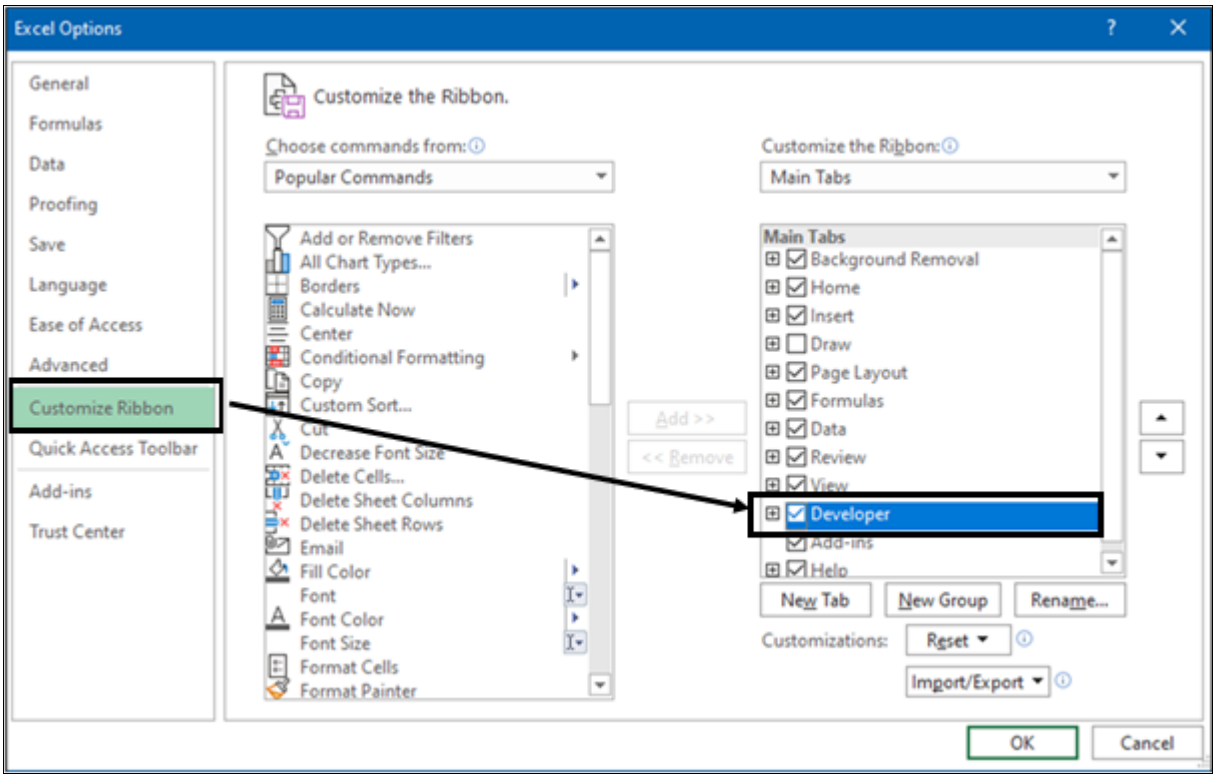


Figure 2.2.19 Hide and show Tabs

2.2.13 EXPORT AND IMPORT A CUSTOM RIBBON

You can export your Customised Ribbon settings to another PC or share your ribbon customizations with someone else.

Export a custom ribbon:

On the computer where you customized the ribbon, open the **Customize the Ribbon** window, click Import/Export, then click **Export all customizations**, and save the ***Excel Customizations.exportedUI*** file to some folder.

Import a custom ribbon:

On another computer, open the **Customize the Ribbon** window, click Import/Export, select **Import customization file**, and browse for the customizations file that you saved.

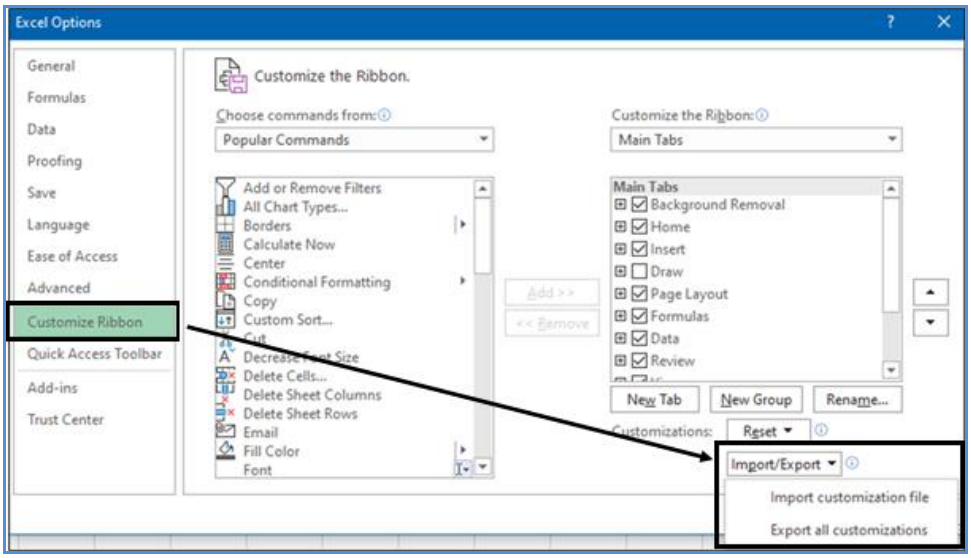


Figure 2.2.20 Import/Export Customisation

Warning: When you import a customized ribbon to a specific PC, all prior ribbon customizations on that PC are lost.

2.3 QUICK ACCESS TOOLBAR

The **Quick Access Toolbar (QAT)** is a small customizable toolbar at the top of the Excel window where you can keep the items that you not only need to access quickly but want to be immediately available independent of which of the Ribbon's tabs you are working on.

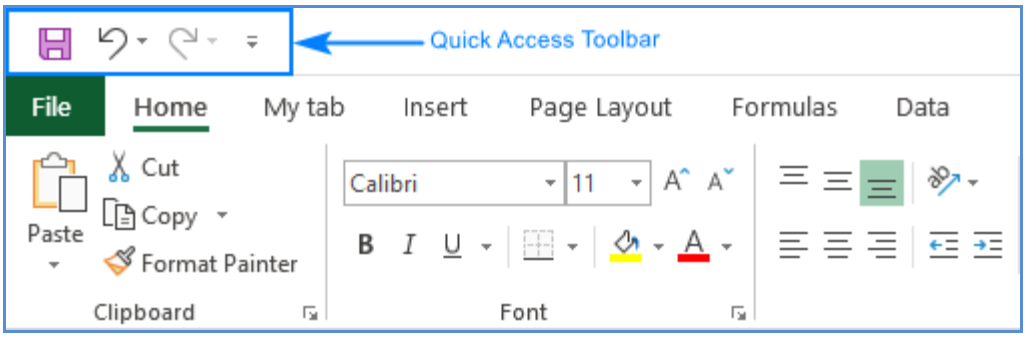


Figure 2.3.1 Quick Access Toolbar

The QAT has beside it a drop-down menu containing a **predefined set of the default commands**, which can be displayed or hidden. Additionally, it consists of a choice to add your own commands.

2.3.1 CUSTOMIZING YOUR QUICK ACCESS TOOLBAR

Apart from default options in the Quick Access Toolbar. You can **customize** your Quick Access toolbar.

2.3.2 MOVE QUICK ACCESS TOOLBAR BELOW OR ABOVE THE RIBBON

Typically, the Quick Access toolbar appears on the left side of the title bar, above the Ribbon. Alternatively, you can display the Quick Access toolbar **below the Ribbon**; just right-click the Quick Access toolbar and choose **Show Quick Access Toolbar below the Ribbon**.

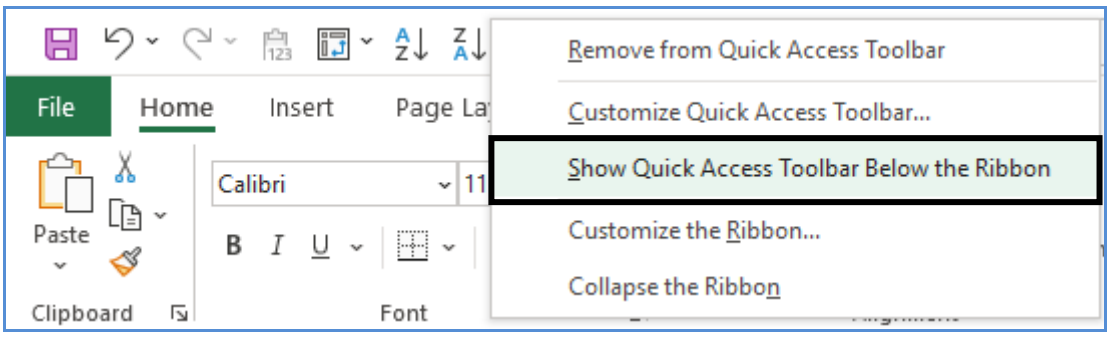


Figure 2.3.2 Move QAT

By default, the Quick Access Toolbar contains three tools: Save, Undo, and Redo. You can customize the Quick Access toolbar by adding other commands that you use often.

What can be customised:

- Add your commands.
- Change the order of commands
- Move the QAT
- Add Macros to QAT
- Import Export your customisations

What cannot be customised:

- Only command Icons can be displayed.
- QAT Buttons cannot be resized
- QAT is only in one line.

2.3.3 HOW TO CUSTOMISE QAT

There are primarily 3 methods

Method 1- File > Options > Quick Access Toolbar (QAT)

Method 2- Right Click on Quick Access Toolbar

Right-click the Quick Access toolbar and choose to Customize the Quick Access Toolbar.

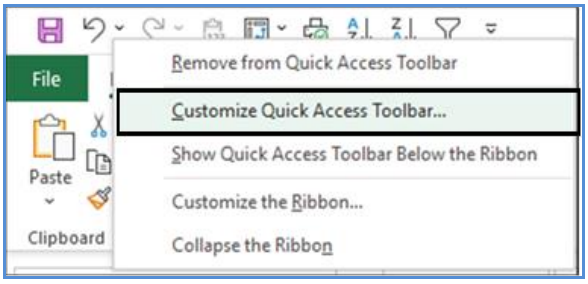


Figure 2.3.3 Customise QAT

Method 3- Click the down arrow beside QAT

If you click the down arrow to the right of the Quick Access toolbar, you see a **drop-down menu** with some additional commands that you might want to place in your Quick Access toolbar.

Apart from commands in Ribbon excel also has many **commands which are not available in the ribbon** which you can access by adding them to your Quick Access toolbar.

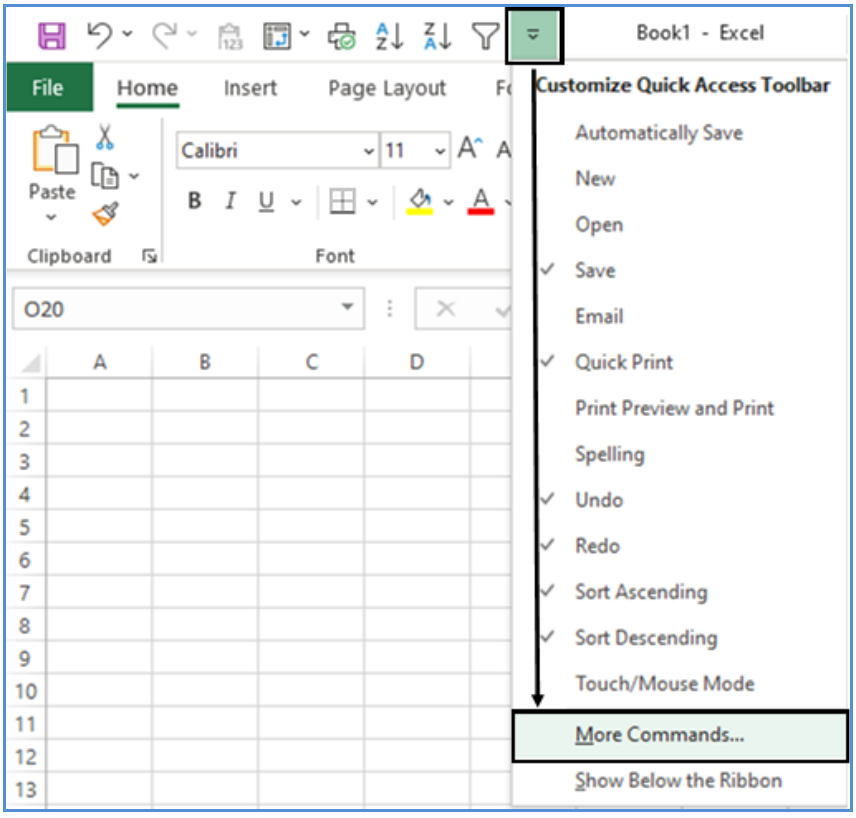


Figure 2.3.4 Customise QAT

You see the **Excel Options** dialog box, shown in Figure 2.3.5

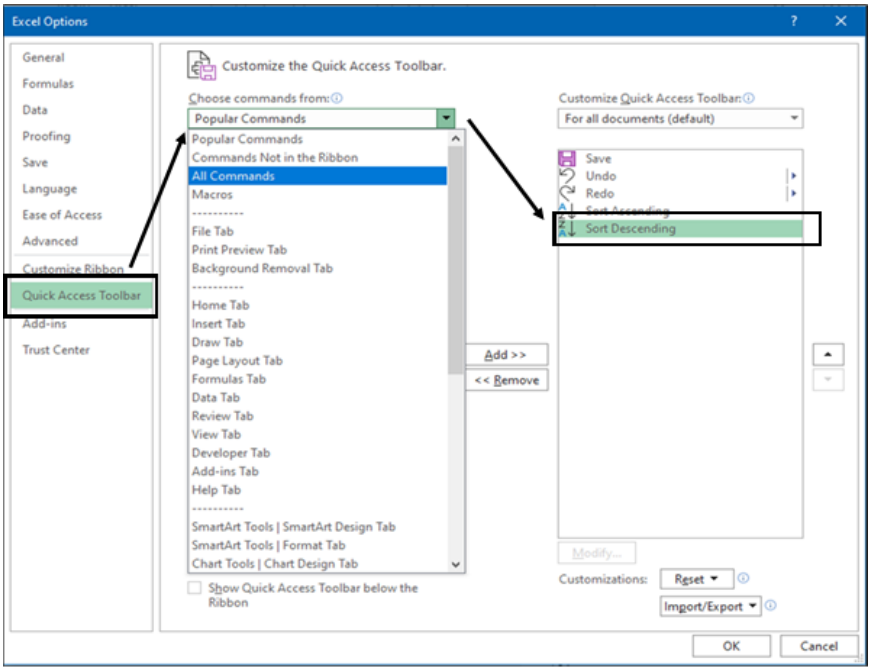


Figure 2.3.5 Customise QAT

2.3.4 ADD A COMMAND TO QUICK ACCESS TOOLBAR

On the left pane in the above Figure, you can see Choose commands from when you click the dropdown, you can see many options, you also have options like "**Commands Not in Ribbon"**

Once you select an option you can click on add to include in Quick Access Toolbar (QAT)

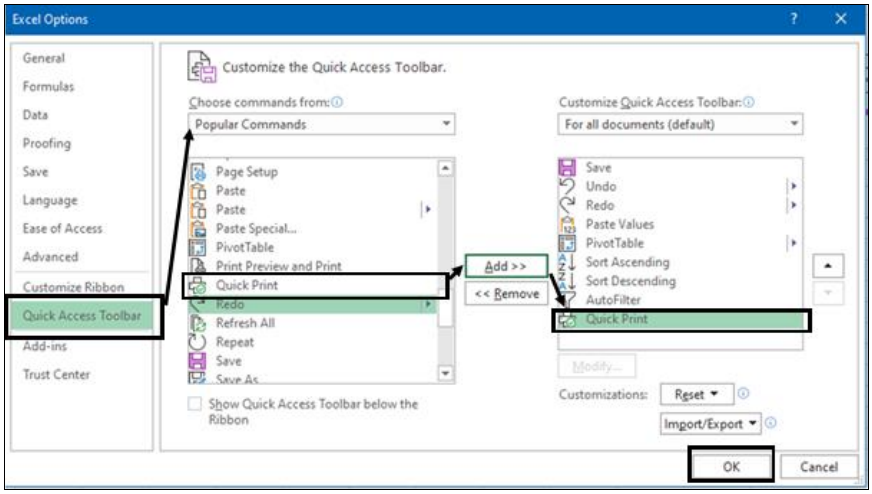


Figure 2.3.6 Adding a Command to QAT

Another method to add a command from the Ribbon to your Quick Access toolbar, right-click the command and choose to Add to Quick Access Toolbar.

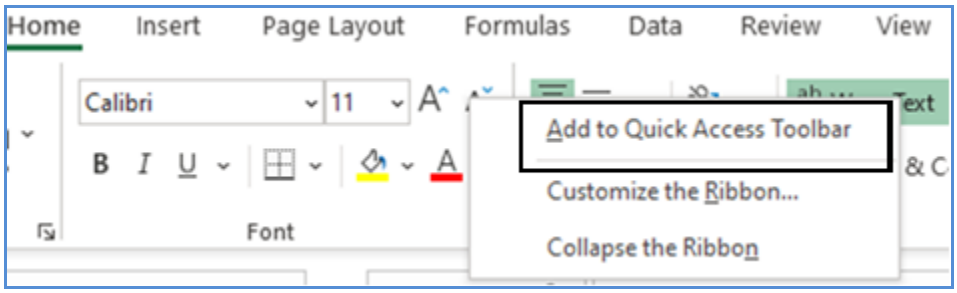


Figure 2.3.7 Adding a Command to QAT

You can also enable command from the predefined list.

To enable a currently hidden command from the predefined list, you can click the Customize Quick Access Toolbar button **(the down arrow).** In the list of the displayed commands, click the one you wish to enable.

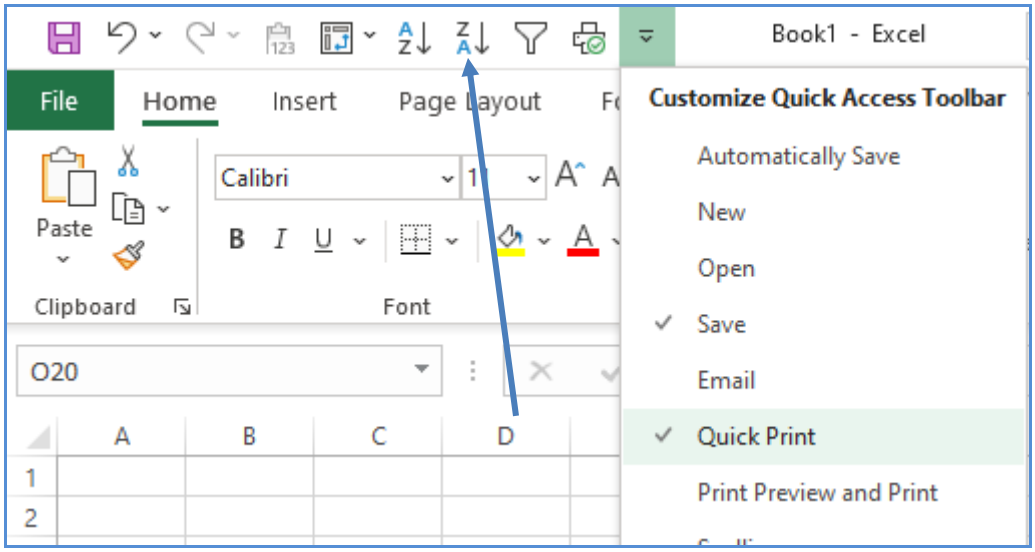


Figure 2.3.8 Enable Command from list

2.3.5 ADD MACROS TO QUICK ACCESS TOOLBAR

You can **add Macros** to the Quick Access Toolbar as follows

- Open the Customize the Quick Access Toolbar window.
- In the Choose commands from the drop-down list on the left, **select Macros**.
- In the list of macros, **select** the one you wish to add to the Quick Access Toolbar.
- Click the **Add button**.

- Click OK to save the changes and close the dialog box.

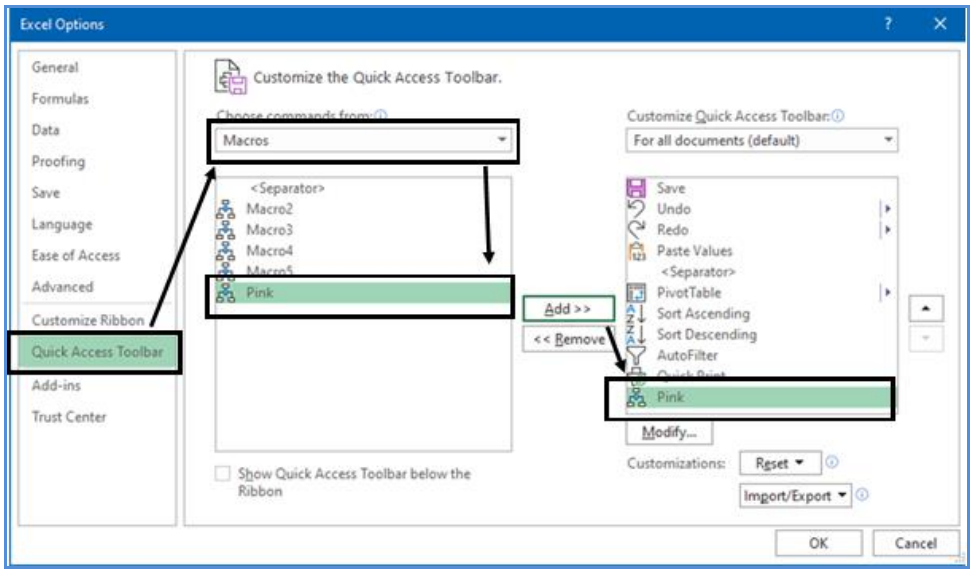


Figure 2.3.9 Adding a Macro to QAT

2.3.6 HOW TO REMOVE A COMMAND FROM QUICK ACCESS TOOLBAR

You can also remove either a default or custom command from the Quick Access Toolbar, just right-click QAT and pick "**Remove from Quick Access Toolbar** from the menu

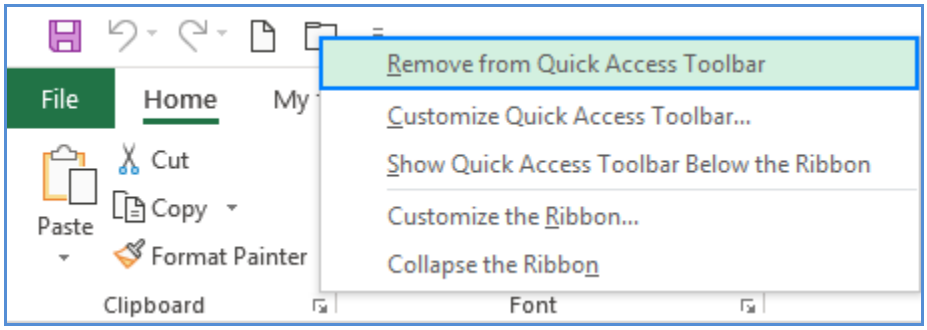


Figure 2.3.10 Remove from QAT

Another method to remove it by selecting the command in the **Customise Quick Access Toolbar** window, and then click the **Remove** button.

2.3.6 REARRANGE COMMANDS ON QUICK ACCESS TOOLBAR

You can **change the order** of the QAT commands:

Open the Customize the Quick Access Toolbar window.

Under Customize Quick Access Toolbar on the right, select the command that you want to move, and click the **Move Up or Move Down arrow**.

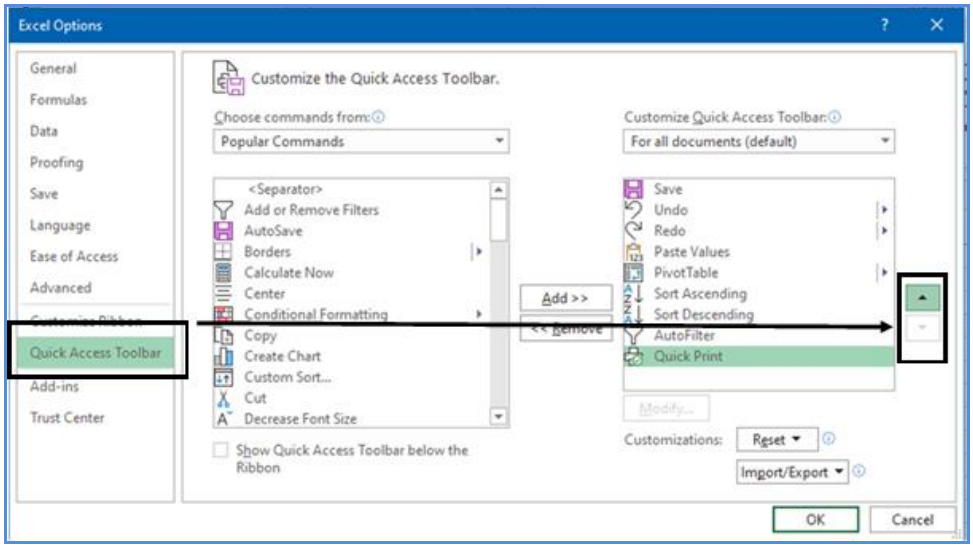


Figure 2.3.11 Moving commands in QAT

2.3.7 GROUP COMMANDS ON QUICK ACCESS TOOLBAR

There could be many commands on the Quick Access toolbar. You can sub-divide them into logical groups, for example, Home and Data Commands, you can group commands by adding a separator.

- Open the **Customize the Quick Access Toolbar** dialog window.
- In the Choose commands from the drop-down list on the left, pick **Popular Commands**.
- In the list of commands on the left, select **<Separator>** and click Add.
- Click the **Move Up or Move Down arrow** to position the separator where needed.
- Click OK to save the changes.

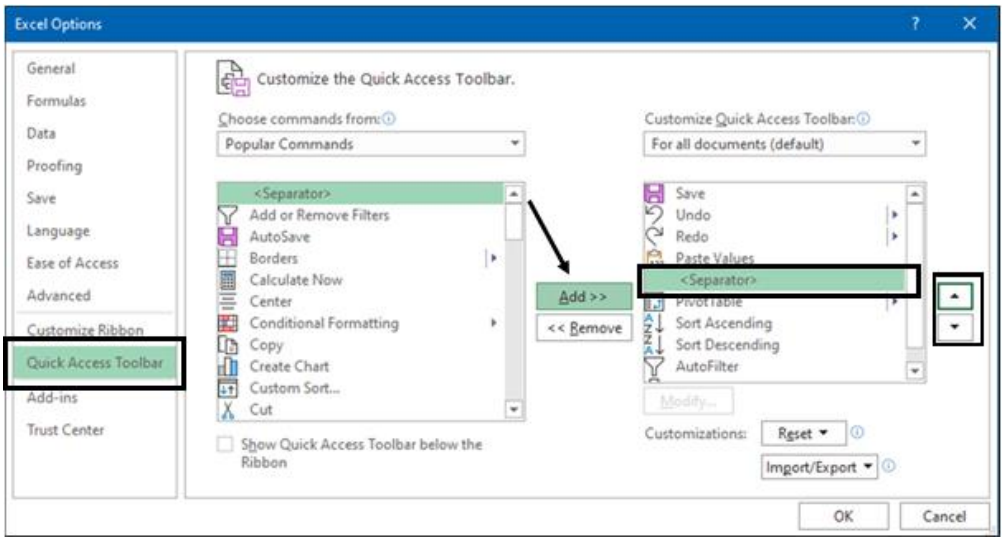


Figure 2.3.12 Group Commands in QAT

2.3.8 RESET QUICK ACCESS TOOLBAR TO THE DEFAULT SETTINGS

You can discard all your customizations and revert the QAT back to its original setup as follows:

- Open the **Customize the Quick Access Toolbar** window.
- Click the **Reset** button, and then click **Reset only Quick Access Toolbar**.

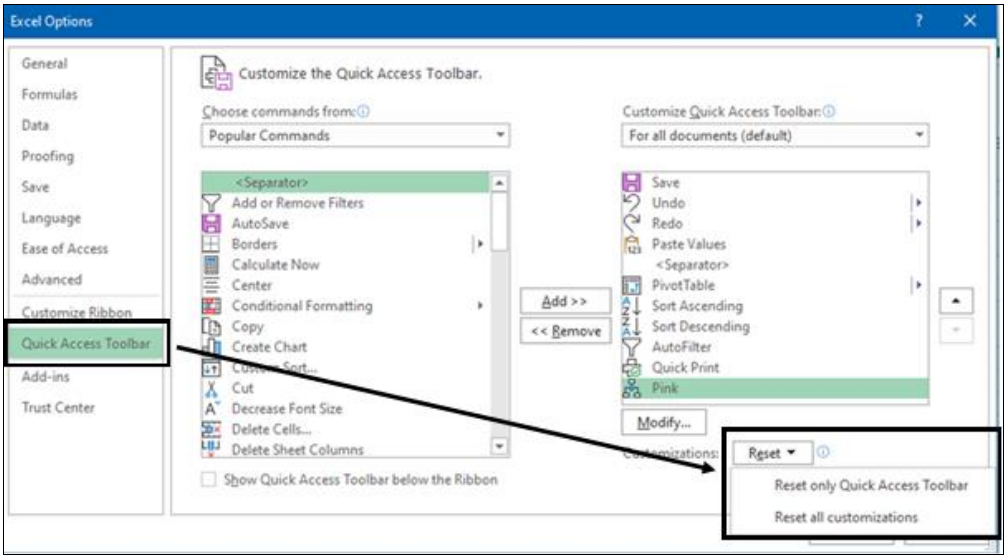


Figure 2.3.13 Reset QAT

2.3.9 EXPORT AND IMPORT A CUSTOM QUICK ACCESS TOOLBAR

You can save your Quick Access Toolbar and ribbon customizations into a file that can be imported later.

Through this, you can keep your Excel interface looking the same on all the computers that you and share it with others.

Export a customized QAT:

In the Customize the Quick Access Toolbar window, click Import/Export, then click **Export all customisations**, and save the customisations file to some folder.

Import a customized QAT:

In the Customize the Quick Access Toolbar window, click Import/Export, select **Import customisation file**, and browse for the customisations file that you saved earlier.

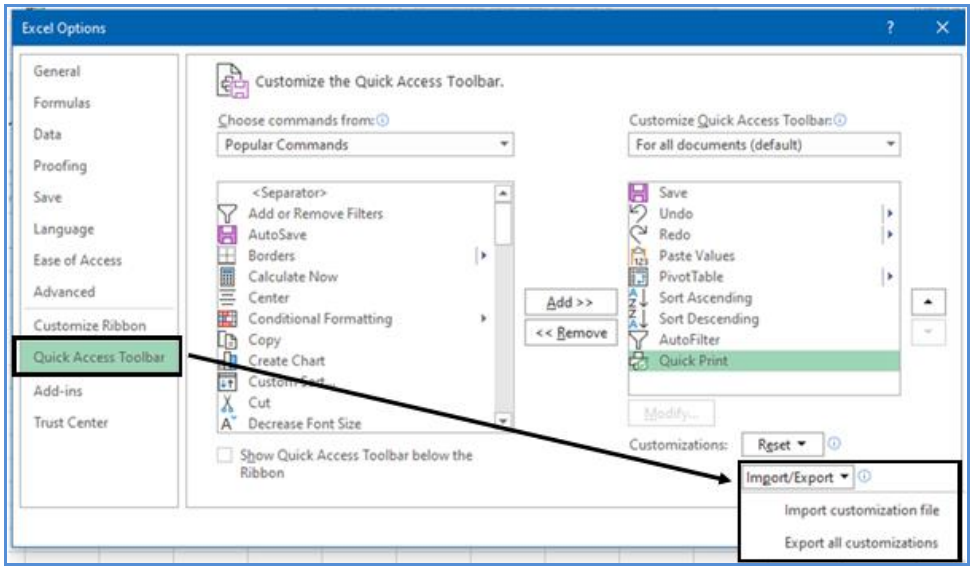


Figure 2.3.14 Import/ export Customisation

Warning: When you import a customized QAT to a specific PC, all prior ribbon customizations on that PC are lost.

2.4 DATA TYPES IN EXCEL

Excel workbook can contain may worksheets and each worksheet can contain Billions of cells, each of these cells can contain primarily three types of data in cells: labels (Text), values (Numbers), and formulas.

A worksheet can also contain charts images etc which are not in cells, but on a Virtual draw layer on top of every sheet.

Values (Numbers)

Numeric values speak of an amount of some sort: Net Profit, Turnover, Dates, even time are all numbers in excel. Excel aligns Numbers to the right side of the cell.

Excel's numbers are precise up to 15 digits it cannot handle numbers larger than that. Numbers that are entered that contain more than 15 digits will be truncated. That is, digits will be truncated off the right-hand side and replaced with zeros.

These 15 Digits are pretty enough in most circumstances, but issues do crop up in Bank account Numbers which are often larger than 15 digits or credit card numbers which are often 16 digits, in these cases, it should be entered as text.

Labels (Text)

Text are descriptive pieces of information, such as names, years, or other identifying statistics, and they usually include letters of the alphabet, numerical characters, symbols such as % and $, as well as spaces and tabs are all valid text. Excel aligns text to the left side of the cell.

Formulas

Formulas are instructions for Excel to perform calculations. Excel aligns formulas to the left side of the cell and starts with an **"="** sign. Cells could have logical values (TRUE or FALSE, also called **Boolean values**)

2.4.1 DATE AND TIME

Dates and times are also **stored as number types** by Excel. Dates are stored as the **number of days since the date 1/1/1900**. In other words, January 1, 1900, is considered by Excel to be 1. Therefore, 1/2/1900 would be stored as 2, 1/3/1900 as 3, and so on. Remember that Excel does not recognize dates BEFORE 1/1/1900.

Excel treats **Times as fractions of days**. Since a day is 24 hours in length, then 1/4th of a day (0.25) would be 6 hours. Since each day begins at 0 hours and ends 24 hours later, Excel would store the date and time for 6:00 AM on 1/1/1900 as 1.25. As an example, if it is 09:00 AM on July 1st, 2020. Excel would store the current date and time as 44013.375. That is, forty thousand Four-hundred Thirteen and (roughly) 38/100ths days since 1/1/1900.

2.4.2 VALUES VS FORMATS

It is important to differentiate the concept of value and format in excel. Since the excel window dresses a number to be displayed differently. The way Excel displays a given numerical value within a cell is known as formatting. As an example, let us consider the following:

Value in Cell	Format	Displayed in Cell
1.00	General	1.00
1.00	Number (4 digit decimal precision)	1.0000
1.00	Currency	₹ 1.00
1.00	Percent	100%
1.00	Date	01-01-1900
0.75	Fraction	3/4
0.75	Time	18:00:00

*The **value** of a number is **stored in the cell**. **Formatting** determines **how the number is displayed**, and what level of precision is displayed*. When using the value of the cell in calculations, the true value is used, not the displayed value

2.4.3 NEW DATA TYPES IN EXCEL

Traditionally in Excel, one cell could only contain one piece of data.

You can now have multiple data fields inside one cell. The new data types which Excel has introduced are really **connections to an online data source** that provides more information about the data.

At present, there are two data types available in Excel: **Stocks** and **Geography**. Many more are being released.

Figure 2.4.1 New Data types

2.4.4 STOCKS

The new stock data type of our cell might display the name of a company but will also contain information like the current share price, trading volume, market capitalization, employee headcount, and year of incorporation for that company.

- You will need **to select the range of cells** with your text data.
- Go to the Data tab.
- Click on the **Stock data** type.
- This will convert the plain text into a **rich data** type.

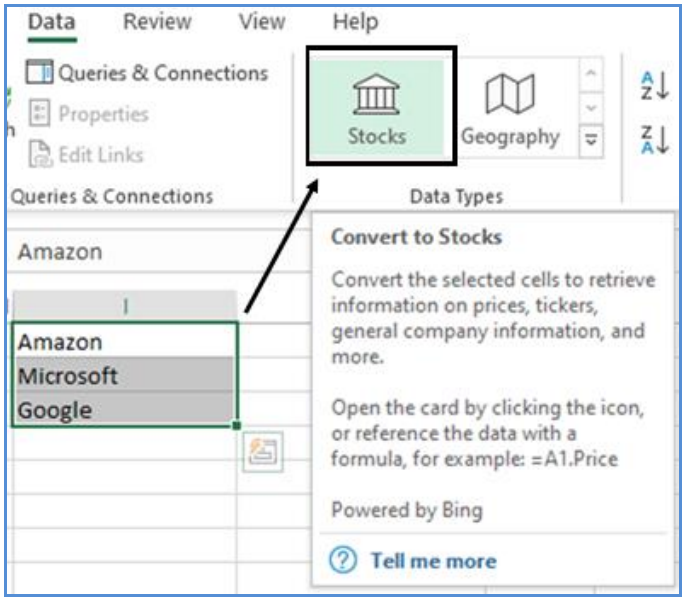


Figure 2.4.2 Stock Data Type

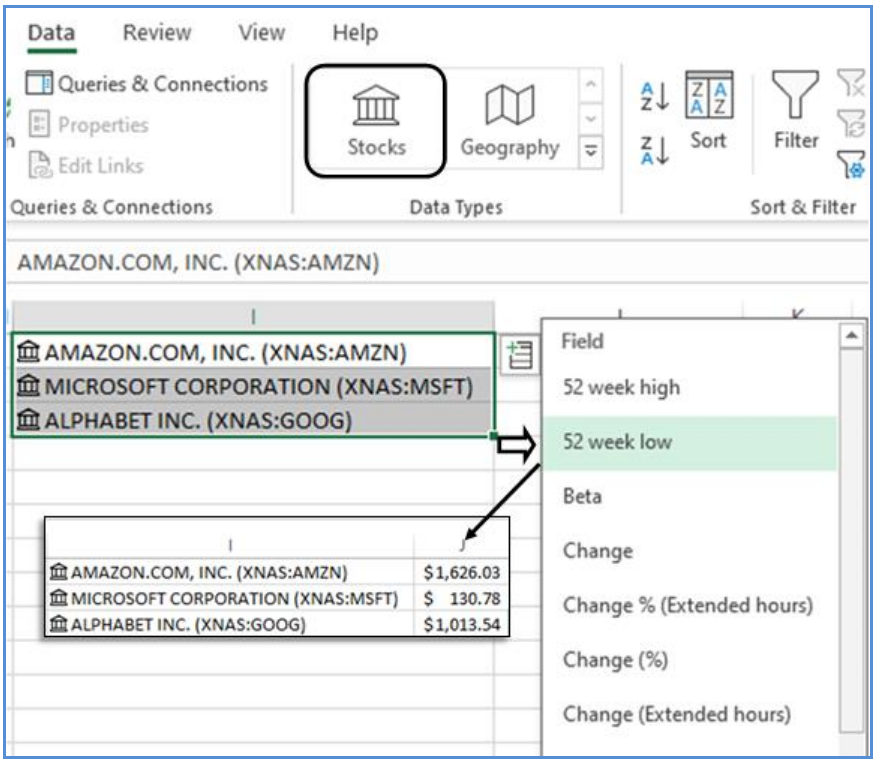


Figure 2.4.3 Extracts the relevant data

Currency & Cryptocurrency Data

With the stock data type, you can not only obtain stocks from a company name but can also be used for **currency and cryptocurrency conversions.**

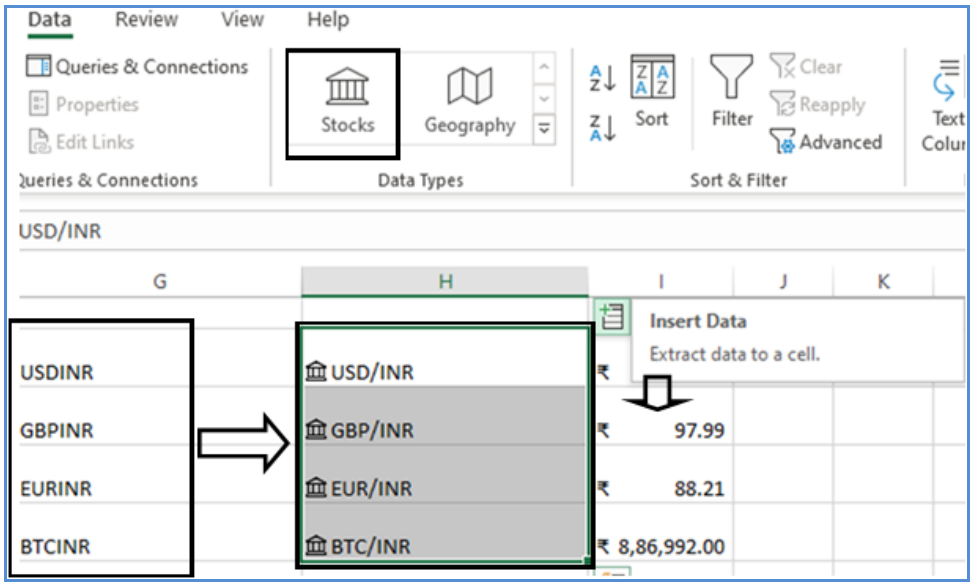


Figure 2.4.4 Currency data through Stock

2.4.5 GEOGRAPHY

The geography data type supports most types of geographies. Countries, states, provinces, counties, and cities. display the name of a company but will also contain information like the Area, Population, Latitude, Longitude, and various other parameters for that Geography.

- You will need to select the range of cells with your text data.
- Go to the Data tab.
- Click on the **Geography** data type.
- This will convert the plain text into a **rich data type**.
- Click on **insert** and you can insert any desired parameter.

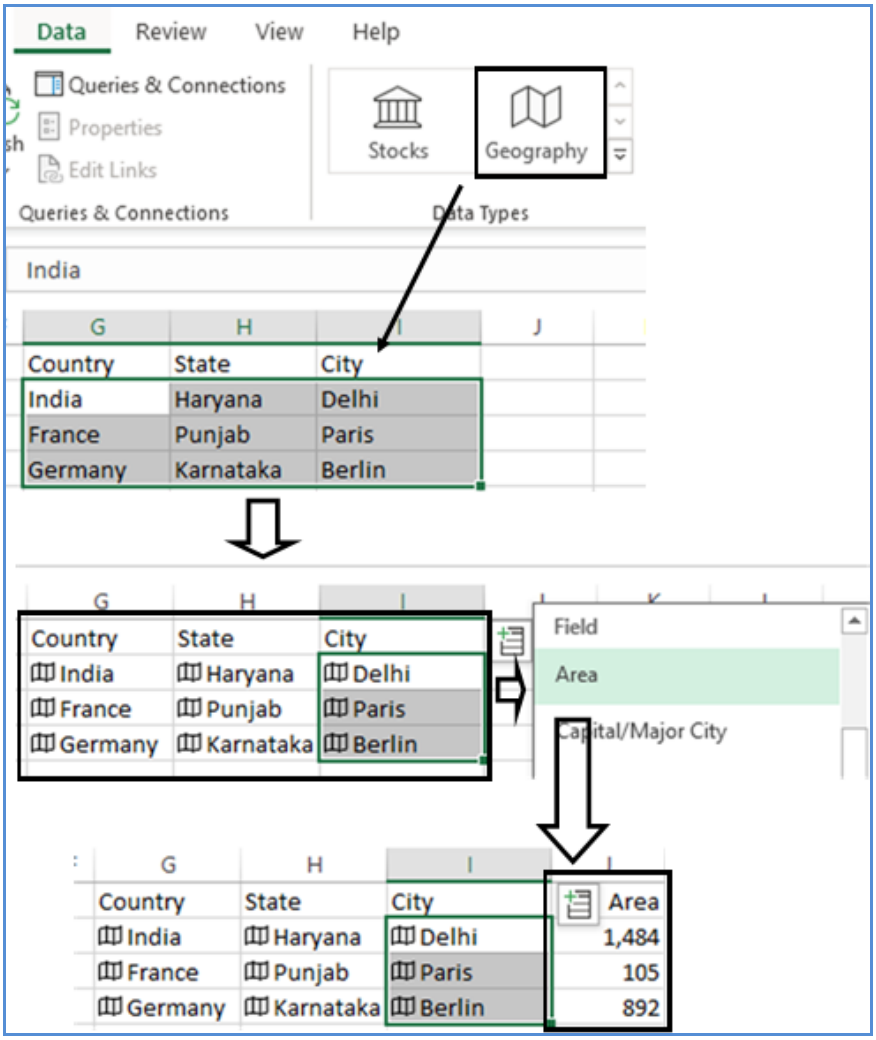


Figure 2.4.5 Geography data type can extract any geographical data

2.5 EXCEL OPTIONS AND CUSTOMISATIONS

Excel is an out of the box global product. When it comes to Excel, the default settings work fine in most of cases. But if need be, you can easily customize Excel through Excel Options.

File > Options

Open the File menu and select Options from the left navigation pane to open the Excel Options dialog box.

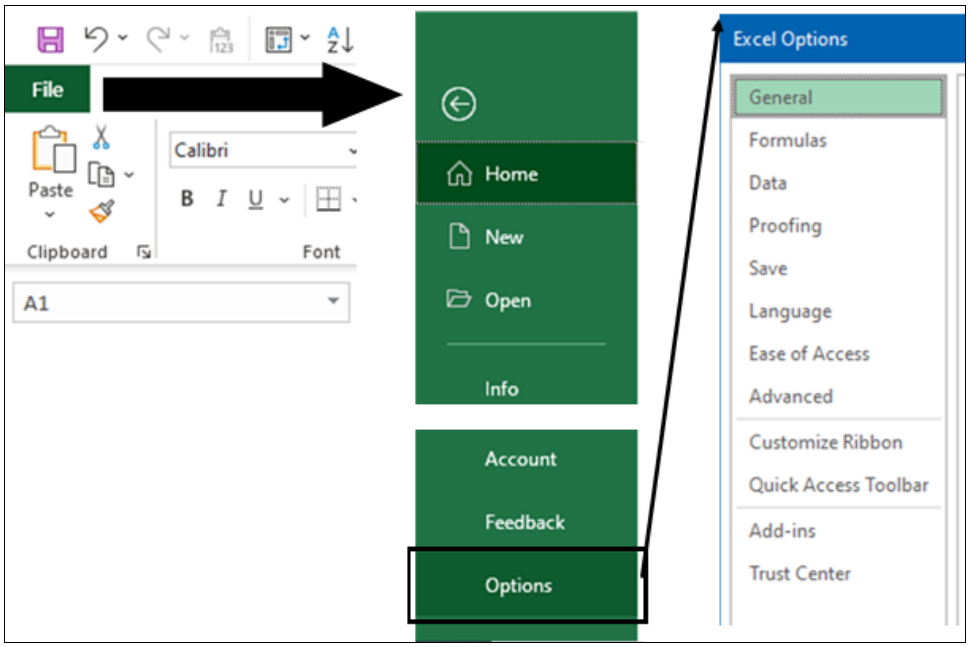


Figure 2.5.1 Excel options

The dialog box has categories for General, Formulas, Data, Proofing, Save, Language, Ease of Access, Advanced, Customize Ribbon, Quick Access Toolbar, Add-Ins, and Trust Center.

General	The most used settings, such as user interface settings, the default font for new workbooks, number of sheets in a new workbook, customer name, and Start screen.
Formulas	All options for controlling calculation, error checking rules, and formula settings
Data	The data category offers the new Edit Default Layout for pivot tables, several other pivot table options, and then a series of checkboxes to bring back the legacy Get Data categories.
Proofing	Spell-check options and a link to the AutoCorrect dialog box.
Save	The default method for saving, Auto Recovery settings, legacy colours, and web server options.
Language	Choose the editing language, ToolTip language, and Help language.
Ease of Access	Options available are Provide Feedback with Sound, Provide Feedback with Animation, Screen Tip Style, and the default document font size.
Advanced	All options that Microsoft considers advanced, spread among 14 headings.
Customize Ribbon	Icons to customize the ribbon.
Quick Access Toolbar	Icons to customize the Quick Access Toolbar (QAT).
Add-Ins	A list of available and installed add-ins. New add-ins can be installed from the button at the bottom of this category.

HELP WITH SETTINGS

Many settings appear with a small ⓘ icon. If you hover the mouse near this icon, Excel displays **a super ToolTip** for the setting. The ToolTip explains what happens when you choose the setting. It also provides some tips about what you need to be aware of when you turn on the setting.

As an example, In figure 2.5.2, The Tooltip tells about Mini Toolbar

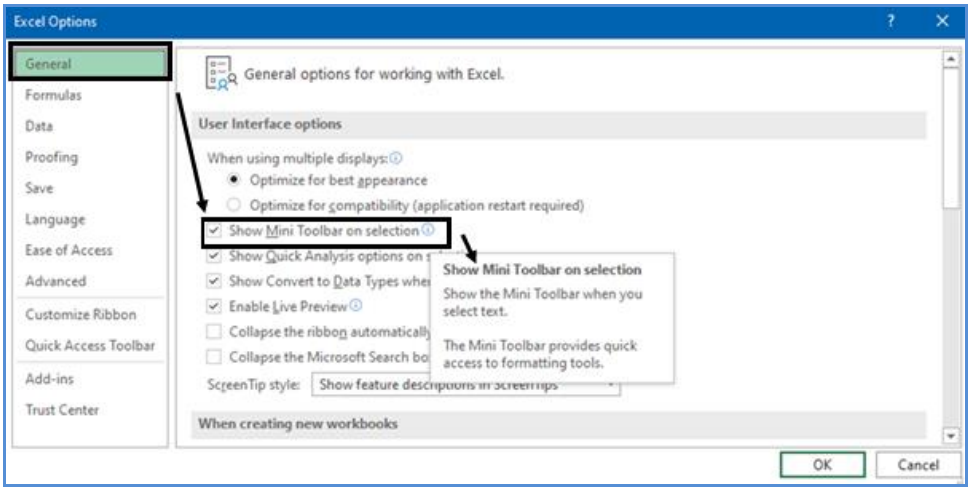


Figure 2.5.2 Help with settings

Some of the useful Excel options are as follows:

2.5.1 CHANGE THE DEFAULT FONT TYPE AND SIZE AND NUMBER OF WORKSHEETS

By default, Excel uses Calibri (body font) font size 11. You can, however, select a different Default font type and size as follows:

File> Excel Options> General.

- In General options, within **'When creating new workbooks'** section, change the default font type and font size.
- You can also change the default number of sheets in workbook **'Include this many sheets'** to 1

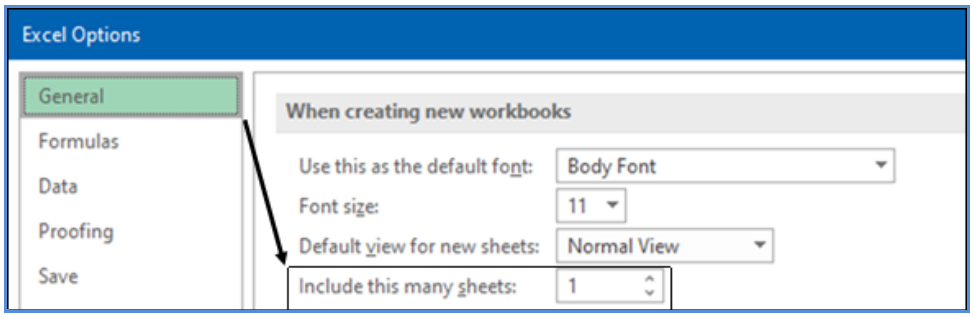


Figure 2.5.3 Change the number of sheets

2.5.2 AUTOMATICALLY INSERT DECIMAL POINTS DURING DATA ENTRY AND ENABLE FILL HANDLE

It can be a very handy option for you if you do a lot of data entry with decimals.

When you enter numeric values with 2 decimal points, instead of typing 123.45, you can simply type 12345, and Excel would automatically insert 2 decimal places.

File> Excel Options> Advanced.

- In Editing Options, select '**Automatically insert a decimal point'** and specify the places.
- Select '**Enable fill handle and cell drag-and-drop'**

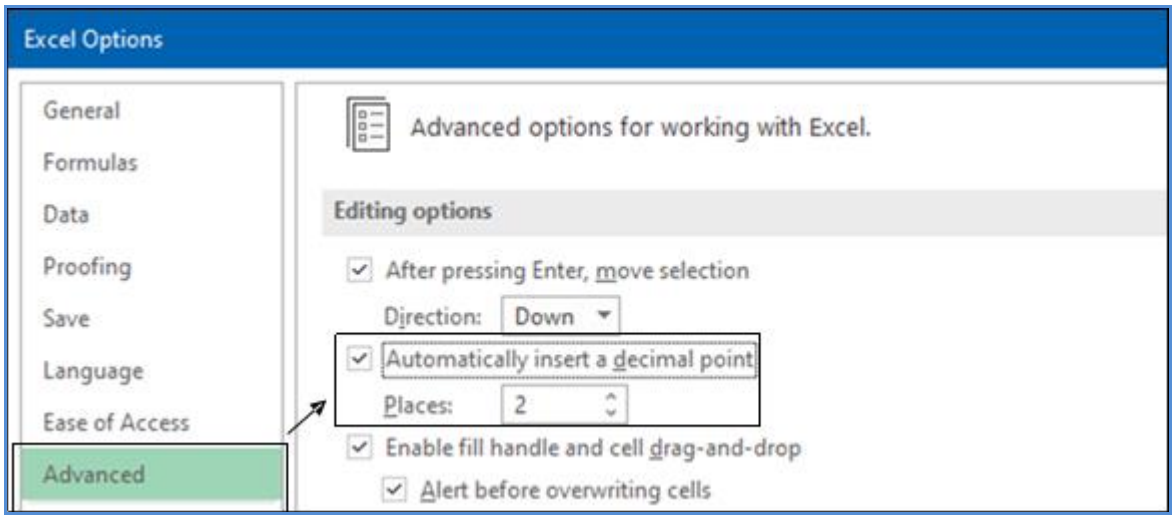


Figure 2.5.4 Enable fill handle

2.5.3 AUTOMATICALLY COMPLETE ABBREVIATIONS

There is a feature in Excel where you can specify shorthand text and Excel will automatically complete it for you. For example, whenever you type ICAI, Excel would automatically replace it with **The Institute of Chartered Accountants of India**.

- **File> Excel Options> Proofing**
- Click on **AutoCorrect Options** button
- In the AutoCorrect dialog box, type what you want to Replace (in Replace field) with what you want to replace it with (in With field)
- Click on Add then OK

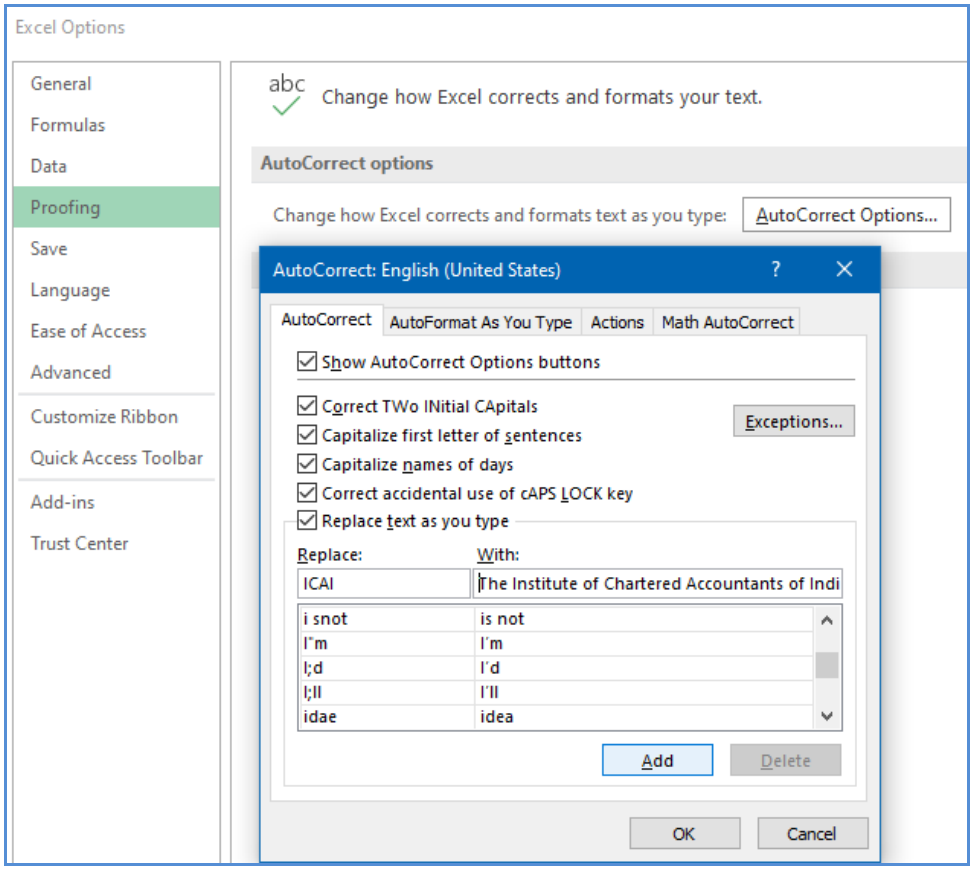


Figure 2.5.5 Auto complete abbreviations

2.5.4 USING AUTORECOVER OPTIONS AND DEFAULT FILE LOCATION

Excel periodically saves a copy of your work every 10 minutes, you change that time If your computer crashes, the recovery pane offers to let you open the last AutoRecovered version of the file. This feature is sure to save you from retyping data that might have otherwise been lost. You can also specify the default location where your file is to be saved.

- **File> Excel Options> Save**
- Select **Save Autorecover** and specify the value 5
- Select **Save to Computer by default** and specify the location.

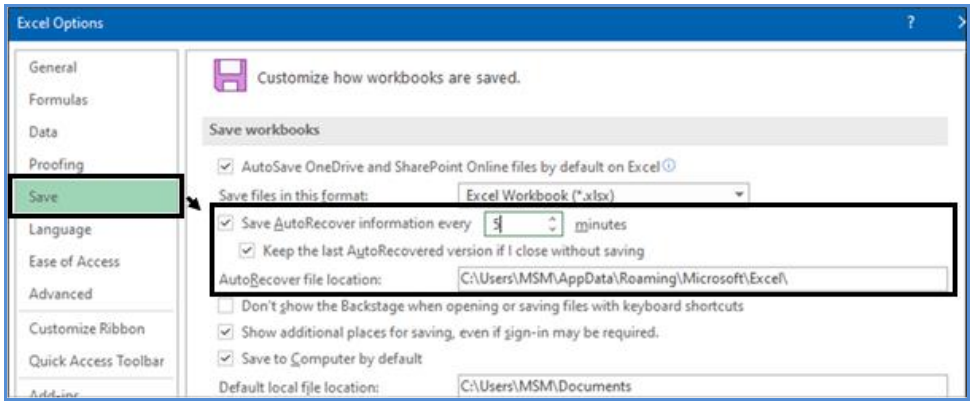


Figure 2.5.6 AutoRecover Options

2.5.5 CREATE CUSTOM LIST

You can create your own custom lists, Details have been explained in Topic Sort

TRUST CENTER

2.5.6 PROTECTED VIEW FOR FILES ORIGINATING FROM THE INTERNET

The files from the Internet or Outlook initially open in protected mode. This mode gives you a chance to look at the workbook and formulas without having anything malicious happen. If you only want to view or print the workbook, the protected mode works great.

After you click **Enable Editing**, Excel will skip protected mode the next time you open the file.

This protected mode can be enabled or disabled as follows

File> Excel Options>Trust center>Protected View

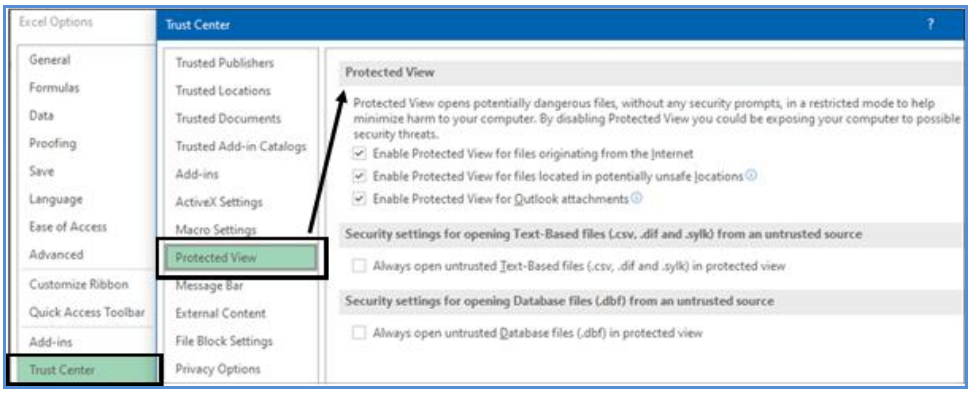


Figure 2.5.7 Protected view

2.5.7 TRUSTED DOCUMENT SETTINGS

By default, Excel warns you about all sorts of things. If you open a workbook with macros, links, external data connections, a message bar appears above the worksheet to let you know that Excel disabled those “threats.”

If you declare a folder on your hard drive to be a trusted folder, you can open those documents without Excel warning you about the items.

File> Excel Options>Trust center>Trusted Locations

- Click on **‘Add a new location’** specify the path.
- You can also make the subfolders trusted.

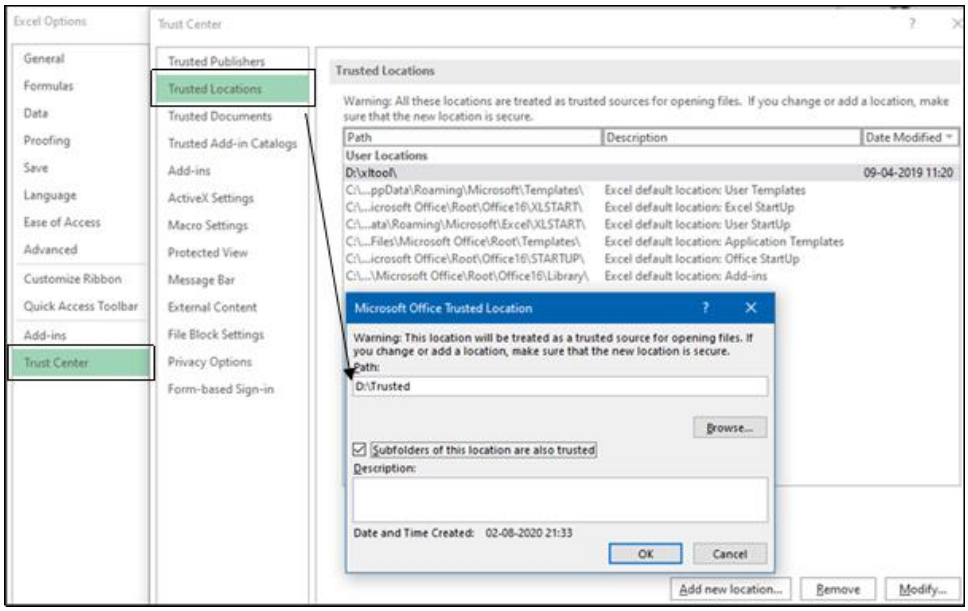


Figure 2.5.8 Trusted Location

2.6 FIND AND SELECT BLANKS, VISIBLE CELLS, NUMBERS, ERRORS

You often receive an Excel file from a client or import it from an external database. You are not sure about the integrity of data thus it is important to check the data to make sure there are no Blanks or Formula errors, or Numbers stored as text and so forth. In a small dataset It could be still easy, but if you have a huge file containing hundreds or even thousands of rows, pinpointing them manually is next to impossible.

All this is possible since Excel provides a dedicated dialog box to access special groups of cells, called "**Go To Special**". Go To Special is a tool within Microsoft Excel that enables you to quickly select cells of a specified type within your Excel worksheet.

There are various methods to access it

Method 1

1. On the **Home** tab, in the **Editing** group, click **Find & Select**.

2. Click **Go To Special**.

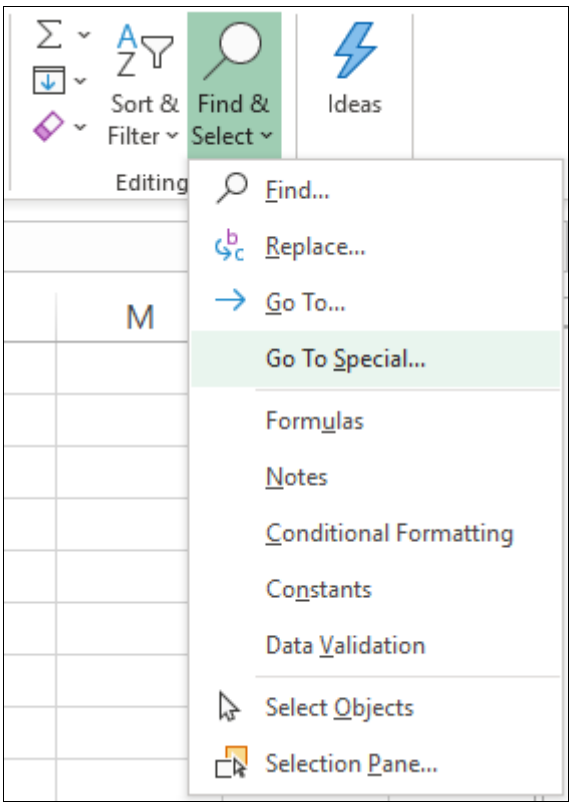


Figure 2.6.1 Go to Special on Home

Method 2

1. To access this dialog with the keyboard, type **Control + G**,
2. Click the Special button (or use **Alt + S**) on Windows.

Method 3

1. Press Function **key F5**
2. Click on **Special** button

Go To Special has various selection features

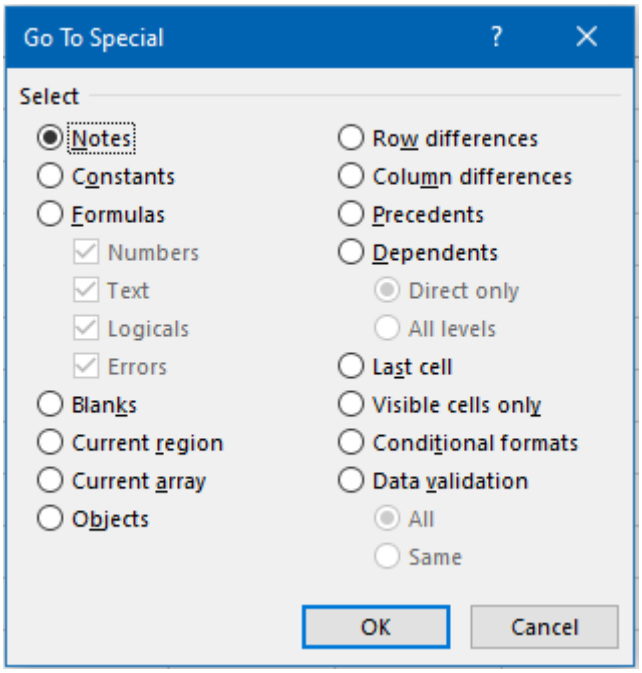


Figure 2.6.2 Go to Special Dialog Box

2.6.1 NOTES

Action: Selects all cells with Notes

Audit Tip: A quick way of finding all cells with Notes, particularly useful if you want to clear all Notes from your worksheet

2.6.2 CONSTANTS

Action: Selects all cells containing constants

Options:

Numbers: Selects all cells with constants that are numbers

Text: Selects all cells with constants that are text

Logical: Selects all cells with constants that are logical (TRUE or FALSE)

Audit Tip: The number constants in your spreadsheet should all be inputs. Highlighting all constants is a great way of checking the structure of your spreadsheet and you can keep it in a different colour to highlight cells where inputs need to be made.

Selecting text could also indicate Numbers stored as Text.

When auditing – **select all constants** and change the **fill colour**. This instantly gives you visibility of your model inputs and flags any inconsistencies.

2.6.3 FORMULAS

Action: Selects all cells containing formulas

Options:

Numbers: Selects all cells with formulas that return numbers

Text: Selects all cells with formulas that return text

Logicals: Selects all cells with formulas that return logicals (TRUE or FALSE)

Errors: Select cells which contain formula errors

Audit Tip: Highlighting all the formulas within your spreadsheet can help in protecting or hiding the formulas. It can also be used to highlight formula errors in the sheet.

2.6.4 BLANKS

Action: Selects all blank cells

Audit Tip: A quick way to select all blank cells. This is useful if you want to quickly format all blank cells or as a way of identifying cells that look blank but contain a constant or formula (i.e. with white on white formatting).

2.6.5 VISIBLE CELLS ONLY

Action: Selects cells that are not hidden (& therefore are visible)

Audit Tip: Useful if you only want to select the non-hidden cells and Copy them somewhere else for example in the filter.

CASE STUDY: In an audit, your principal CA Chandiwala gives you a division wise quarterly sales report as shown in Fig received from a client. You are to find inconsistencies in terms of: are blank cells really blank, any numbers stored as text, any errors in formulas?

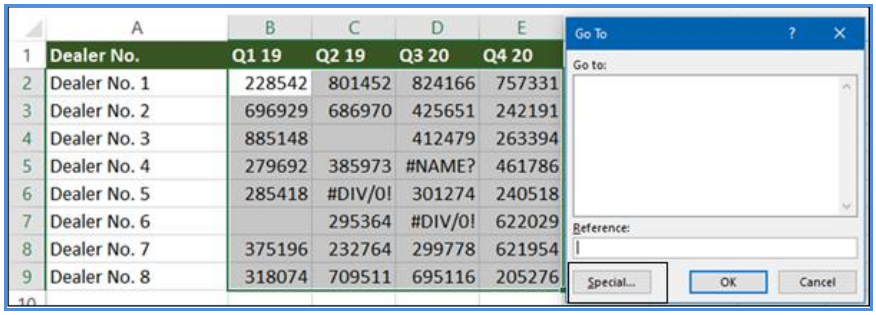
	A	B	C	D	E
1	**Dealer No.**	**Q1 19**	**Q2 19**	**Q3 20**	**Q4 20**
2	Dealer No. 1	228542	801452	824166	757331
3	Dealer No. 2	696929	686970	425651	242191
4	Dealer No. 3	885148		412479	263394
5	Dealer No. 4	279692	385973	#NAME?	461786
6	Dealer No. 5	285418	#DIV/0!	301274	240518
7	Dealer No. 6		295364	#DIV/0!	622029
8	Dealer No. 7	375196	232764	299778	621954
9	Dealer No. 8	318074	709511	695116	205276

Figure 2.6.3 Quarterly sales report

Strategy:

We can analyse the file using '**Go To Special'** as follows:

1. Select the range where you want to highlight. To select all cells with data, click the upper-left cell and press **Ctrl + Shift + End** to extend the selection to the last used cell.
2. Press **F5**
3. Click **Special**



	A	B	C	D	E
1	Dealer No.	Q1 19	Q2 19	Q3 20	Q4 20
2	Dealer No. 1	228542	801452	824166	757331
3	Dealer No. 2	696929	686970	425651	242191
4	Dealer No. 3	885148		412479	263394
5	Dealer No. 4	279692	385973	#NAME?	461786
6	Dealer No. 5	285418	#DIV/0!	301274	240518
7	Dealer No. 6		295364	#DIV/0!	622029
8	Dealer No. 7	375196	232764	299778	621954
9	Dealer No. 8	318074	709511	695116	205276

Figure 2.6.4 Press F5 Go To Special

Finding Blanks

4. In the **Go To Special** dialog box, select Blanks and click OK.
5. This will **select all empty cells** in the range.

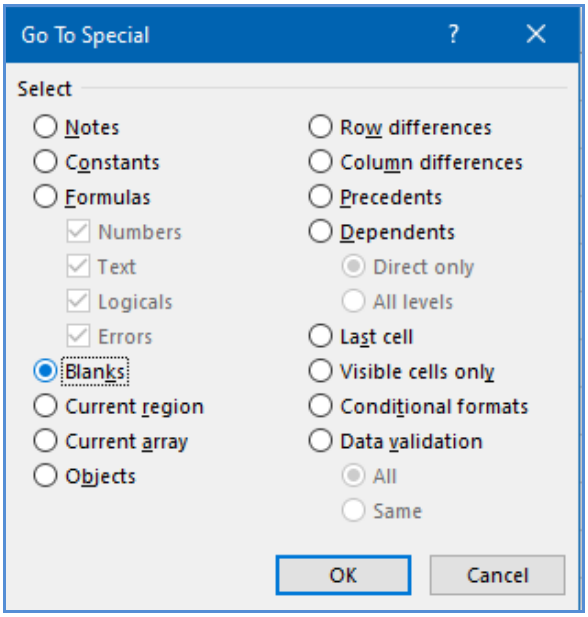


Figure 2.6.5 Select Blanks

	A	B	C	D	E
1	**Dealer No.**	**Q1 19**	**Q2 19**	**Q3 20**	**Q4 20**
2	Dealer No. 1	228542	801452	824166	757331
3	Dealer No. 2	696929	686970	425651	242191
4	Dealer No. 3	885148		412479	263394
5	Dealer No. 4	279692	385973	#NAME?	461786
6	Dealer No. 5	285418	#DIV/0!	301274	240518
7	Dealer No. 6		295364	#DIV/0!	622029
8	Dealer No. 7	375196	232764	299778	621954
9	Dealer No. 8	318074	709511	695116	205276

Figure 2.6.6 Blanks Highlighted

6. If you observe **cell C4** seems to be **blank** but it is **not showing as blank**.
7. If we check then we find someone has **formatted the number in white** so will not appear in printing but will affect the result of any operation like sum

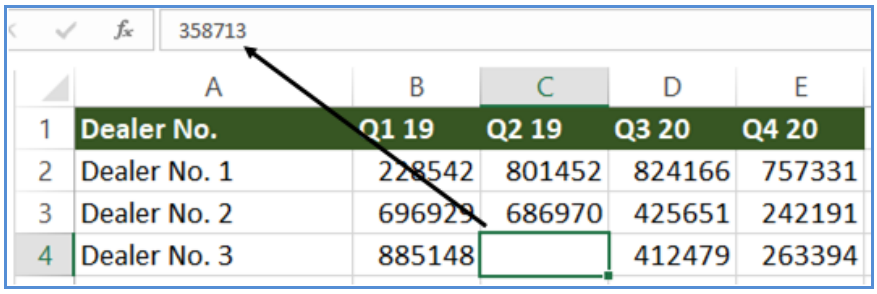


Figure 2.6.7 Cell has a value

Numbers stored as text

1. Again, **highlight the range > Press F5 or CTRL +G >Special**
2. In the **Go To Special** dialog box, select **Constants** and under options select **Text**
3. Click OK.
4. This will select all **Text** cells in the range which in this case are Numbers stored as Text
5. If you observe in the formula bar, you can see a **quote before number** indicating that the number has been stored as text

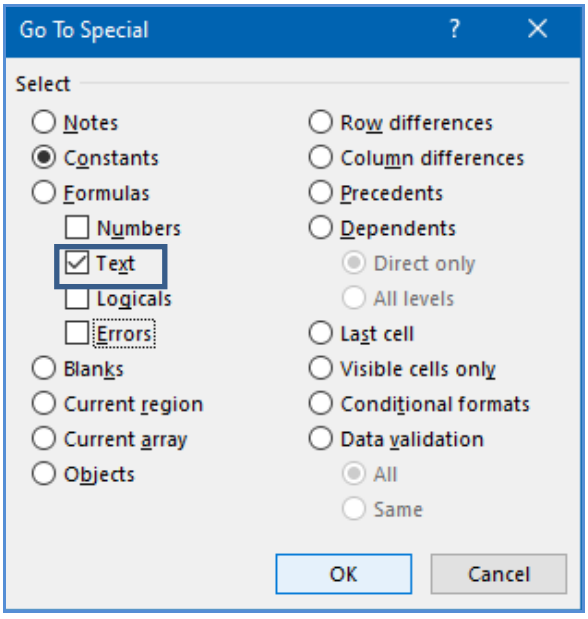
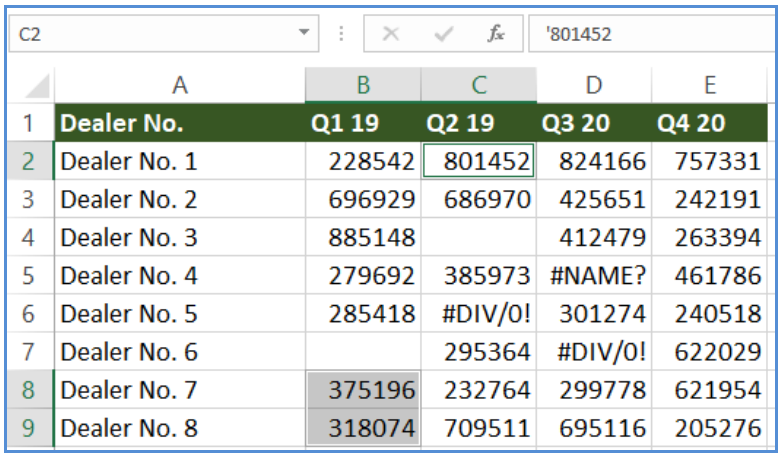


Figure 2.6.8 Go to Special Text



	A	B	C	D	E
1	**Dealer No.**	**Q1 19**	**Q2 19**	**Q3 20**	**Q4 20**
2	Dealer No. 1	228542	801452	824166	757331
3	Dealer No. 2	696929	686970	425651	242191
4	Dealer No. 3	885148		412479	263394
5	Dealer No. 4	279692	385973	#NAME?	461786
6	Dealer No. 5	285418	#DIV/0!	301274	240518
7	Dealer No. 6		295364	#DIV/0!	622029
8	Dealer No. 7	375196	232764	299778	621954
9	Dealer No. 8	318074	709511	695116	205276

Figure 2.6.9 Text Highlighted

Errors in Formula

1. Again **highlight the range > Press F5 or CTRL +G >Special**
2. In the **Go To Special** dialog box, select **Formulas** and under options select **Errors**
3. Click OK.
4. This will select all Formula Error cells in the range.

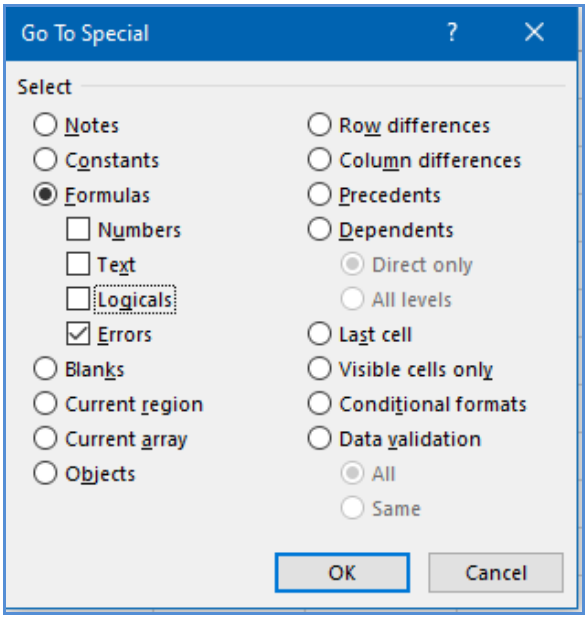


Figure 2.6.10 Go To Special Errors

	A	B	C	D	E
1	**Dealer No.**	**Q1 19**	**Q2 19**	**Q3 20**	**Q4 20**
2	Dealer No. 1	228542	801452	824166	757331
3	Dealer No. 2	696929	686970	425651	242191
4	Dealer No. 3	885148		412479	263394
5	Dealer No. 4	279692	385973	#NAME?	461786
6	Dealer No. 5	285418	#DIV/0!	301274	240518
7	Dealer No. 6		295364	#DIV/0!	622029
8	Dealer No. 7	375196	232764	299778	621954
9	Dealer No. 8	318074	709511	695116	205276

Figure 2.6.11 Errors Highlighted

Gist: We have analysed the Excel data to find inconsistencies using the feature **Go To Special**.

Commands learnt: Find and select > Go to Special

Food for thought: Go To Special

- **Go To Special** only selects cells in the current worksheet rather than the whole workbook.
- **Go To Special** searches within the selected range, if you want to select the entire worksheet to ensure that only one cell is selected.

2.7 SMART COPY PASTE IN EXCEL

Often while working in Excel, you would need to copy and paste information from one cell to another, or one range to another. You would all be familiar with Copy using a shortcut **CTRL + C** and pasting at another location using **CTRL + V**. In Excel you can Copy a cell to another location, Copy a cell to a range ie the same cell is copied to all cells in the range, also copy a range to another range if they are the same size.

The normal Copy operation can be done by using Ribbon command **Home >Clipboard >Copy** transfers a range of cells to the clipboard for pasting to a different location **Home >Clipboard >Paste**

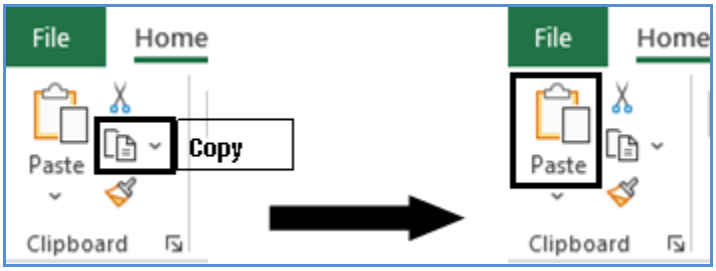


Figure 2.7.1 Copy Paste

When you copy a cell or range, Excel surrounds the copied area with a **moving border**. If that border remains moving, the copied information is available for pasting. If you press Esc to cancel the animated border, Excel removes the information from the Clipboard. Also, once the information is pasted the information is removed from Clipboard. In case you want to

Smart Tip: Smart Copy Paste

However, if you click the Copy button more than once before you click the Paste button, Excel may automatically display the Office Clipboard taskbar.

Or

You can click on the dialog launcher on the bottom right of the **Home >Clipboard** group to toggle the Clipboard task pane on and off.

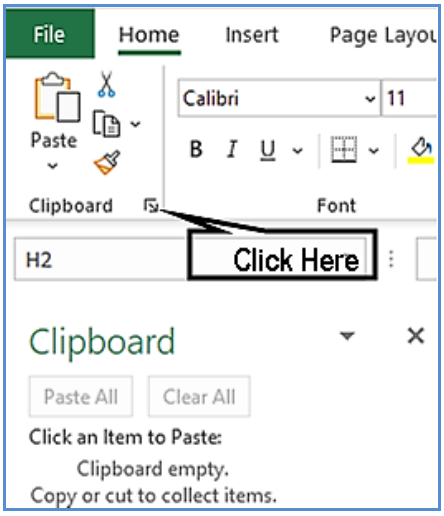


Figure 2.7.2 Activate Clipboard

When you copy information to the Office Clipboard, it gets appended to the information to the Office Clipboard instead of replacing what is already there. With multiple items stored on the Office Clipboard, you can then paste the items either individually or as a group.

Case Study: In an audit, your principal CA Chandiwala gives you a workbook containing four sheets for different quarters which he wants you to append into a single sheet so that that annual data can be analysed.

	A	B	C	D
1	DATE	INV_NO	CUSTOMER	AMOUNT
2	24-04-2019	6	Sun Way	45,460.00
3	29-05-2019	34	Sun Way	81,068.00
4	05-06-2019	41	Sun Way	7,493.00
5	06-06-2019	42	Sun Way	43,877.00
6	09-06-2019	45	Sun Way	24,659.00
7	29-06-2019	69	Sun Way	22,906.00
8	30-06-2019	70	Sun Way	8,954.00
9				

Q1 Q2 Q3 Q4

Figure 2.7.3 Sheet with 4 Quarters

Strategy:

We can Append the sheets into one annual data as follows:

1. Click on the dialog launcher on the bottom right of the **Home >Clipboard** to activate Clipboard

2. Copy the data from Sheet Q1, you will see it shows in the clipboard.
3. Similarly, we will keep on copying the data from Q2, Q3, and Q4 and the **data appears on Clipboard**.
4. Now if you want to copy all of them in one go. You have an option of **Paste all**.

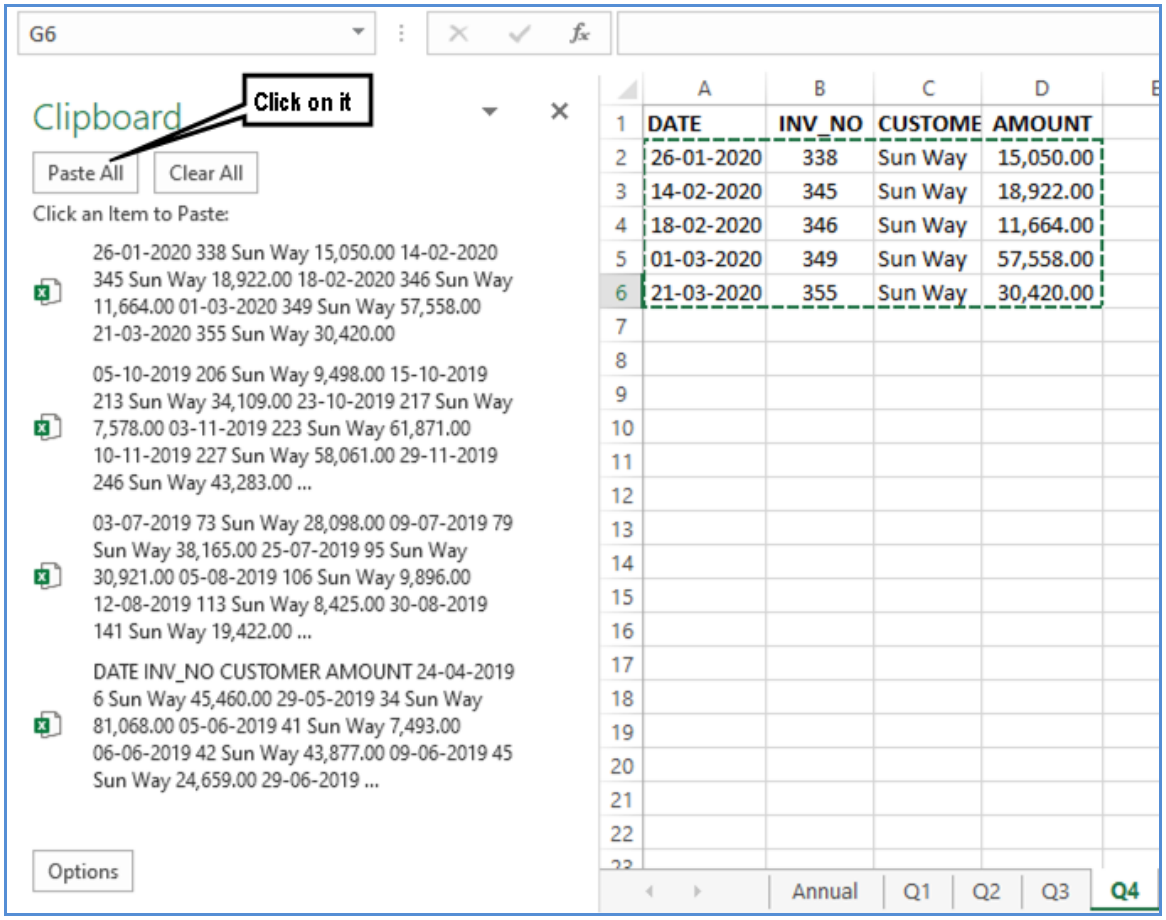


Figure 2.7.4 Copied to Clipboard

5. The whole years' data is pasted at once

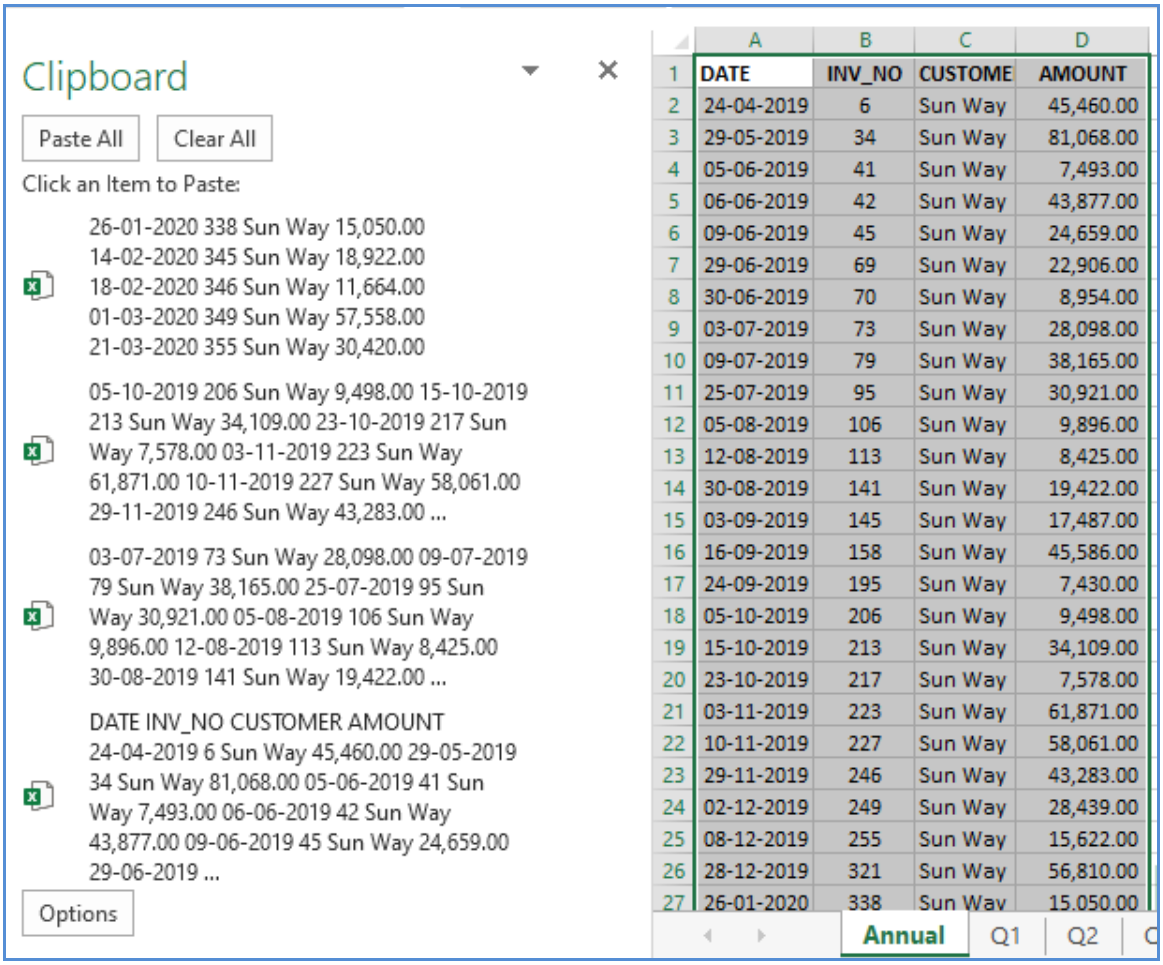


Figure 2.7.5 Data of 4 Quarters pasted in one go

Gist: We have appended the data from various ranges into one sheet using **Paste all** feature in **Clipboard**.

Commands learnt: Home > Clipboard

2.7.1 PASTE SPECIAL

In Excel **Paste Special** is a feature through which apart from pasting a whole cell, you can paste only a certain attribute such as value, formula, formatting, column width, or note.

Excel's Paste Special offers a wide range of options to paste

- only **specific elements** of the copied cells or
- perform a mathematical operation with the copied data.

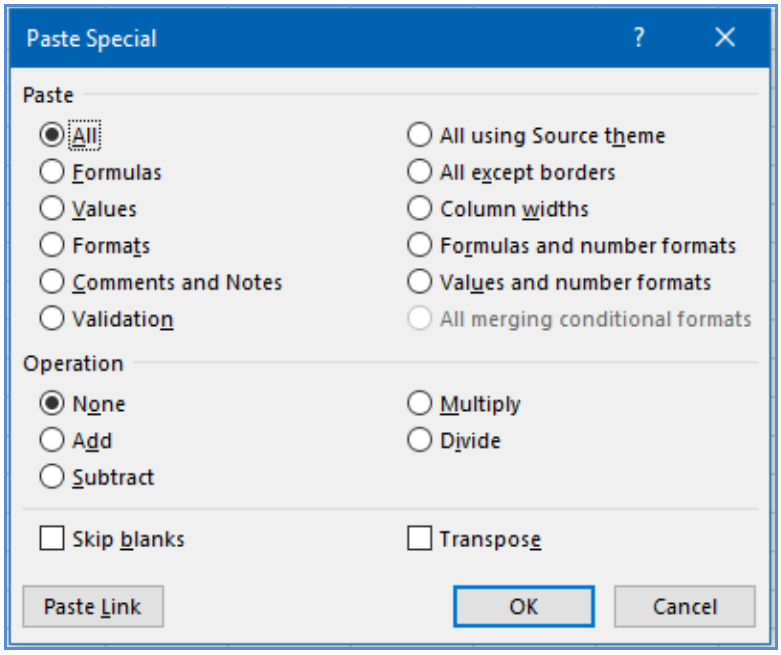


Figure 2.7.6 Paste Special

You can use the Paste Special through following steps

- Copy the source cell or a range of cells **CTRL +C**
- Select the destination cell(s).
- Open the **Paste Special** dialog.
- Select the desired **paste option**,
- Click OK or press the Enter key.

There are various methods to open **Paste Special** Dialog

Method 1

- Home> ClipBoard> Paste
- Click on **Dropdown**
- Click **Paste special**

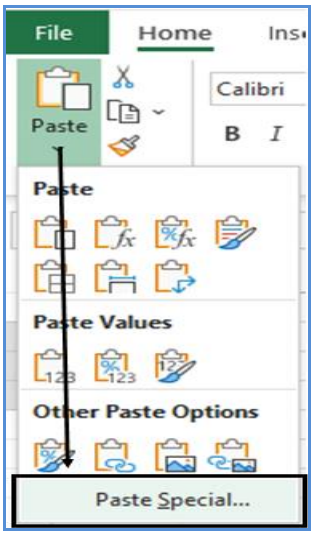


Figure 2.7.7 Paste Special on Home Tab

Method 2

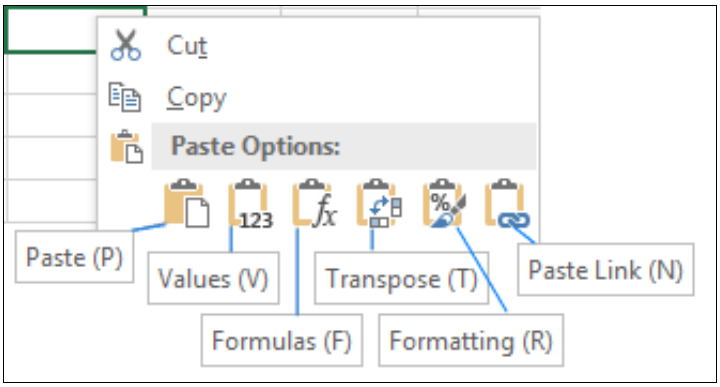
- **Right-click** a cell where you want to paste the copied data,
- Click Paste Special in the context menu.
- 

Figure 2.7.8 Paste Special through right-click

- Hover over the Paste Special item in the context menu,
- A fly-out menu will show up offering 14 more paste options

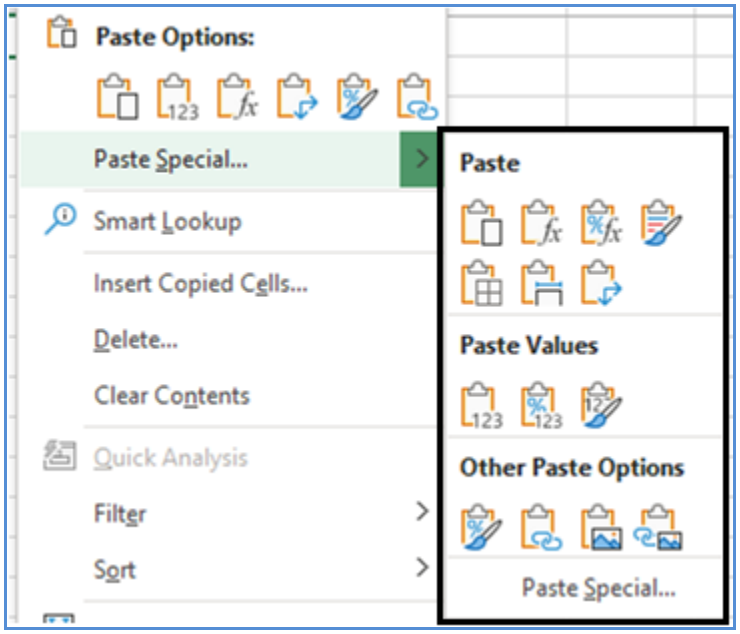


Figure 2.7.9 Paste Special

Method 3

The fastest way is to use the shortcut **ALT+E+S**

2.7.2 HOW TO USE PASTE SPECIAL

There are various ways in which Paste special can be used

Values

You can Paste Values rather than underlying formula

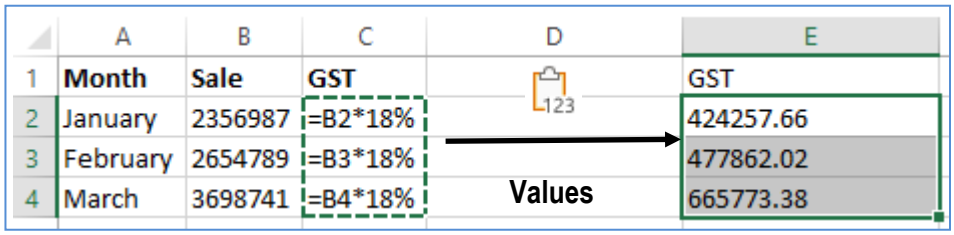


Figure 2.7.10 Paste special values

Comments & Notes

You can copy only Comments or Notes

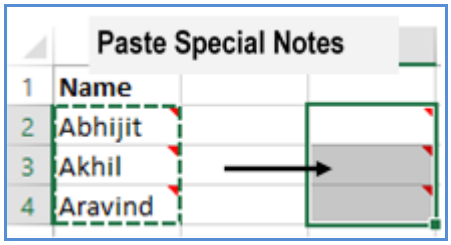


Figure 2.7.11 paste Special Comments

Transpose

You can change columns to rows in Excel using Paste Transpose Option

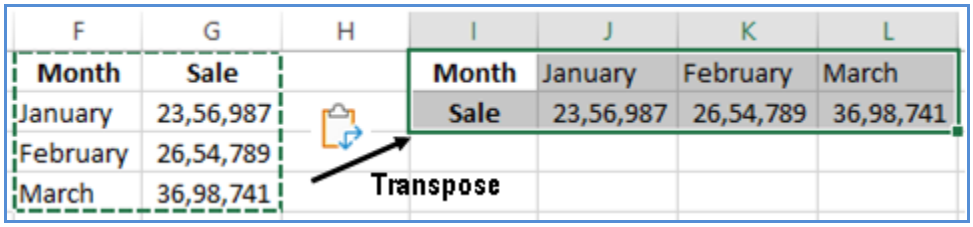


Figure 2.7.12 Paste Special Transpose

Column Width

You can Paste only column width; it is a quick way to have the destination cell with the same column width.

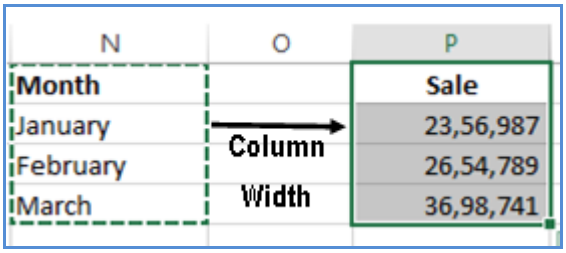


Figure 2.7.13 Paste Special Column Width

Paste and add/subtract/multiply/divide at a time

You can paste as well as perform operations like add, subtract, multiply or Divide on all destination cells at the same time.

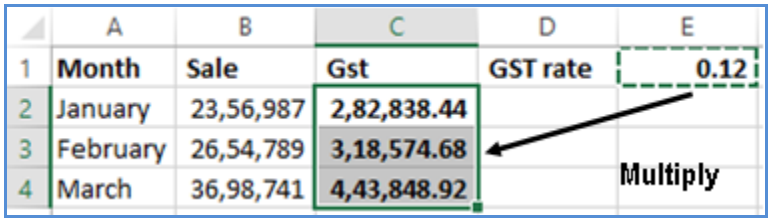


Figure 2.7.14 Paste Special Multiply

Case Study: In an investigation, your principal CA Chandiwala gives you a workbook containing Data consisting of Name, emailed, Sales. You have to calculate GST and Total sales which you have to compare with total GST deposited. Since the data has been extracted in ASCII format and imported to Excel, the Numbers are stored as text, performing any calculations is a challenge. You are to perform the following:

(i) Remove the hyperlinks in the email id of the customer.

(ii) Convert Numbers stored as text to Numbers.

(iii) Calculate GST and Total sales

Strategy:

We can use the power of **Paste Special** for all the above tasks:

Paste + Multiply(*)can be used to remove all hyperlinks in your worksheet in one go

1. Select cells **F1** which is a **Blank**
2. Press **Ctrl + C** to copy them.
3. Select the Email, cells B2:B8
4. Press the paste special shortcut **(ALT + ES),**
5. Select **Multiply** under Operations, or press M.
6. This will multiply each of the amounts in column B by value in F1.
7. Click Enter resulting in **removing all Hyperlinks**.

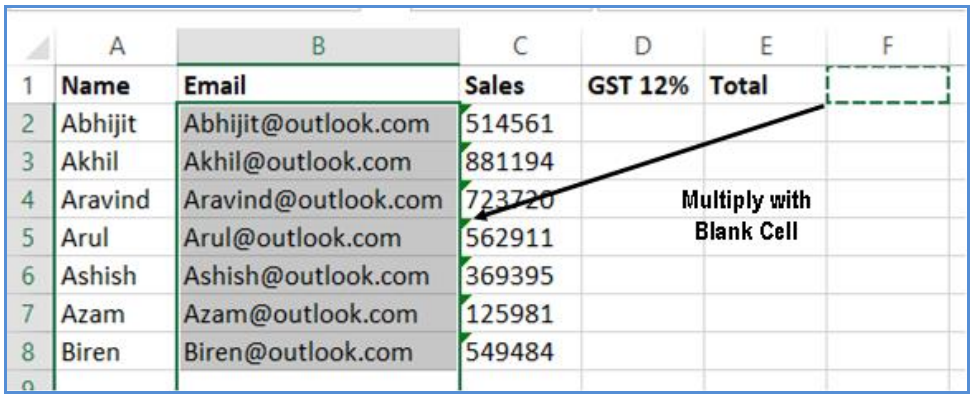


Figure 2.7.15 Paste Special Multiply

Paste + Multiply can be used to convert Numbers stored as text in a range in one go

1. Select cells **F1** which has a value 1
2. Press **Ctrl + C** to copy them.
3. Select the Email, cellsC2:C8
4. Press the paste special shortcut **(ALT + ES),**
5. Select **Multiply** under Operations, or press M.
6. This will multiply each of the amounts in column B by a value 1.
7. Click Enter resulting in converting all **Numbers stored as text to Numbers**.

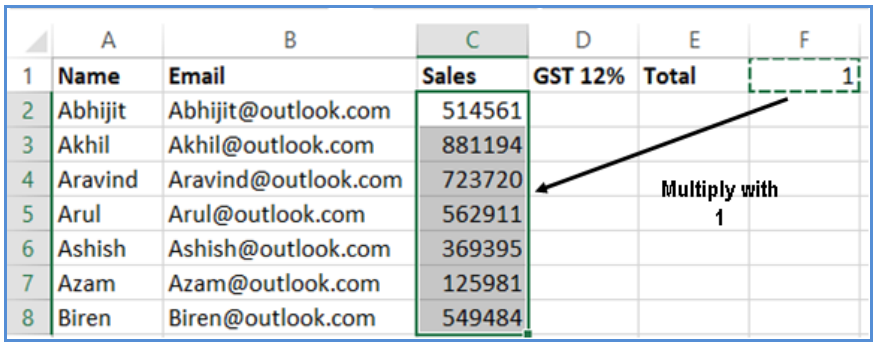


Figure 2.7.16 Paste Special Multiply

Paste + Multiply can be used to Multiply all the cells in the destination with the copied value.

1. Select cells **F1** which has a value .12 ie GST Rate 12%
2. Press **Ctrl + C** to copy them.
3. Select the Email, cells D2:D8
4. Press the paste special shortcut **(ALT + ES),**

5. Select **Multiply** under Operations, or press M.
6. This will multiply each of the amounts in column B by a value of 0.12.
7. Click Enter resulting in **converting all Values to GST**.

	A	B	C	D	E	F
1	**Name**	**Email**	**Sales**	**GST 12%**	**Total**	0.12
2	Abhijit	Abhijit@outlook.com	5,14,561	61747.32	5,14,561	
3	Akhil	Akhil@outlook.com	8,81,194	105743.3	8,81,194	
4	Aravind	Aravind@outlook.com	7,23,720	86846.4	7,23,720	**Multiply**
5	Arul	Arul@outlook.com	5,62,911	67549.32	5,62,911	
6	Ashish	Ashish@outlook.com	3,69,395	44327.4	3,69,395	
7	Azam	Azam@outlook.com	1,25,981	15117.72	1,25,981	
8	Biren	Biren@outlook.com	5,49,484	65938.08	5,49,484	

Figure 2.7.17 Paste Special multiply

1. Similarly Multiplying with 1.12 can give you total Sales

	A	B	C	D	E	F
1	**Name**	**Email**	**Sales**	**GST 12%**	**Total**	1.12
2	Abhijit	Abhijit@outlook.com	5,14,561	61747.32	576308.3	
3	Akhil	Akhil@outlook.com	8,81,194	105743.3	986937.3	
4	Aravind	Aravind@outlook.com	7,23,720	86846.4	810566.4	
5	Arul	Arul@outlook.com	5,62,911	67549.32	630460.3	
6	Ashish	Ashish@outlook.com	3,69,395	44327.4	413722.4	**Multiply**
7	Azam	Azam@outlook.com	1,25,981	15117.72	141098.7	
8	Biren	Biren@outlook.com	5,49,484	65938.08	615422.1	

Figure 2.7.18 Paste Special

Gist: We have removed Hyperlinks, Converted Numbers stored as Text to Numbers and Calculated Percentage Values using **Paste Special** feature in **Clipboard**.

Commands learnt: Home > Clipboard> Paste Special

2.8 MANAGING WORKSHEETS

Excel offers a convenient way to manage data. As an example, as an Accountant, you might have on the table with your firm's yearly sales, another with expenses, and third analyzing profitability and another one with Balance sheet figures and probably one with Forecasts. If you create these tables in different spreadsheets, you must copy the information you want the sheets to share from one location to another, all without misplacing a number or making a mistake. Excel offers a neat way to do it through Workbooks.

In Excel, each file is called a workbook, and each workbook can contain one or more worksheets.

A workbook can be thought of a book containing many sheets of paper and worksheet is akin to a single sheet of paper in that book.

2.8.1 2.8.1 ADDING AND REMOVING WORKSHEETS

When you start a fresh workbook in Excel, you get a single blank worksheet. To add more sheets, you need to click the **"New sheet"** button, which is a **small plus-in-a-circle** icon that appears immediately to the right of your last worksheet tab.

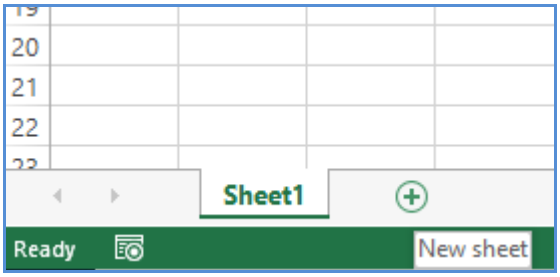


Figure 2.8.1 New Sheet

If you continue adding worksheets, you will eventually find that all the worksheet tabs will not fit at the bottom of your workbook window. If you have more worksheets than fit into the tab list, you will need to use the scroll buttons, which are immediate to the left of the worksheet tabs) to review the list of worksheets. The figure shows the scroll buttons. You can scroll across the sheet through these buttons.

Tip: **If you want to see all sheets in one go, just Right-click on the right scroll Button**

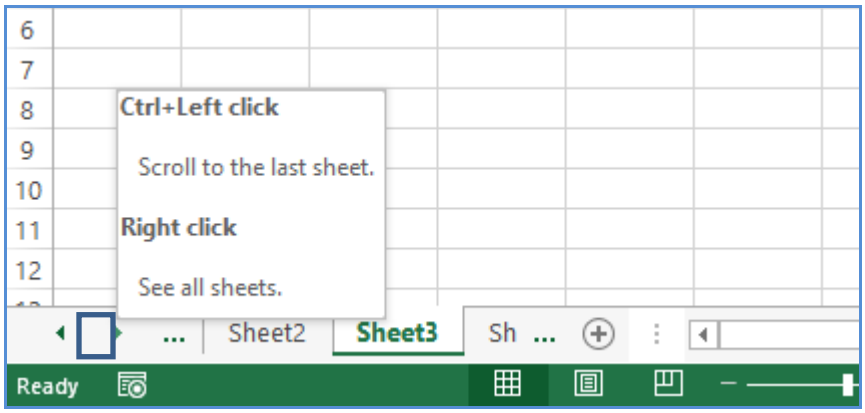


Figure 2.8.2 Right Click to see sheet

You want to Remove a worksheet, simply move to the sheet you want to remove, and then choose **Home>Cells>Delete>Delete Sheet** (you can also right-click a tab, and then choose Delete).

2.8.2 MOVING BETWEEN SHEETS

To move from one worksheet to another, you have different Methods:

- Click the worksheet tabs at the bottom of Excel's grid window (just above the status bar).
- Use Shortcut **Ctrl+Page Down** to move to the **next worksheet**.
- Use Shortcut **Ctrl+Page Up** to move to the **previous worksheet.**

2.8.3 HIDING WORKSHEETS

You can also hide a worksheet temporarily. When you hide a worksheet, its tab disappears, but the worksheet itself remains part of your workbook file, available whenever you choose to unhide it. You cannot print a hidden worksheet, either.

- To hide a worksheet**, right-click** the worksheet tab, and then **choose Hide**.
- To redisplay a hidden worksheet, right-click any worksheet tab, and then choose Unhide.
- The Unhide window appears along with a list of all hidden sheets, as shown in Figure.
- Select a sheet from the list,
- click OK to unhide it.

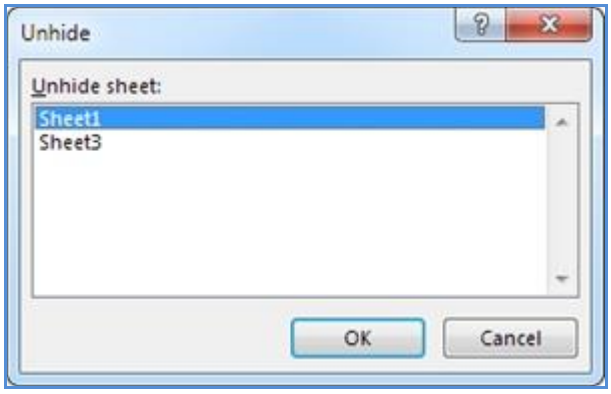


Figure 2.8.3 Hidden Sheets

Case Study: In an ongoing Forensic audit your principal CA Chandiwala gives you a workbook containing a Bank Reconciliation which is reconciled but when he crosschecks Cheques issued not presented at" 'ITGI Ann1'!D14" and goes to the sheet where the details of pending cheques is there, he finds totals are not matching but the reference is correct, as per attached Figure

He is perplexed and gives you the workbook to find if there is any fraud involved. How will you check the authenticity of the sheets and integrity of the workbook?

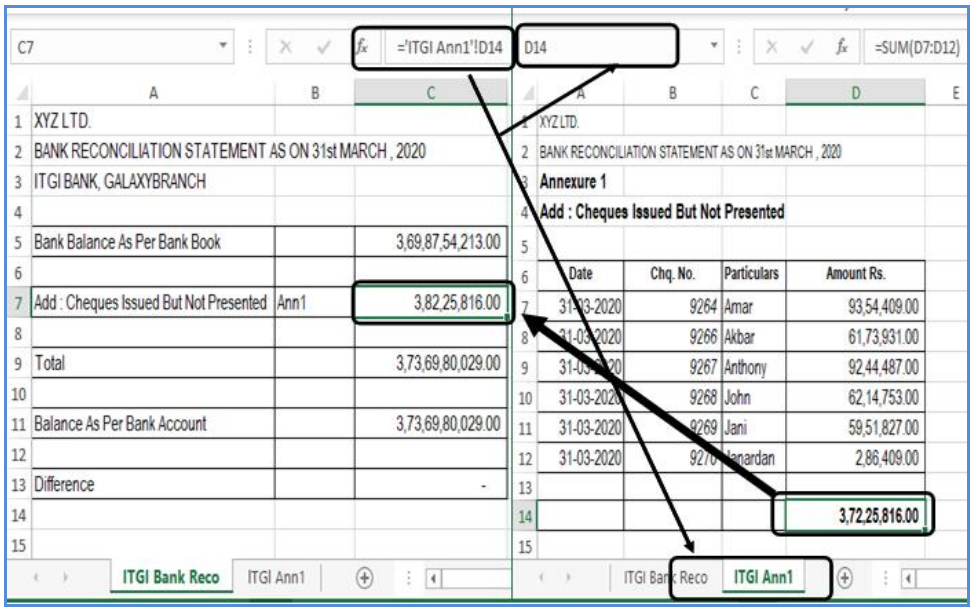


Figure 2.8.4 Excel Reconciliation Sheet

Strategy:

This is a case where everything seems to be correct, Cell reference is correct, but still, figures are not matching. This is how fraudster creates an illusion. If you are not cautious, you are bound to be misled.

As a part of checking the integrity of workbooks, it is always a good control to **check for hidden sheets**.

In this case when you check for Hidden sheets

1. **Right-click** on any tab.
2. Select **Unhide**
3. You find there is another hidden sheet with a matching name

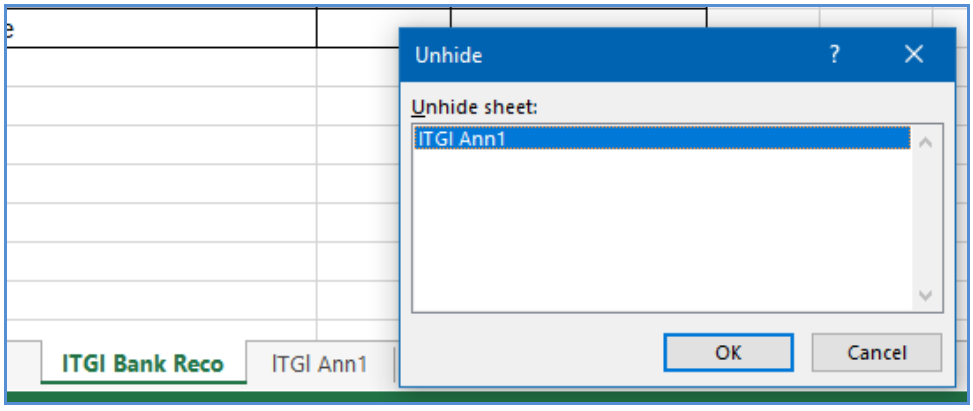


Figure 2.8.5 Unhide sheets

4. Click Unhide
5. The sheet is unhidden

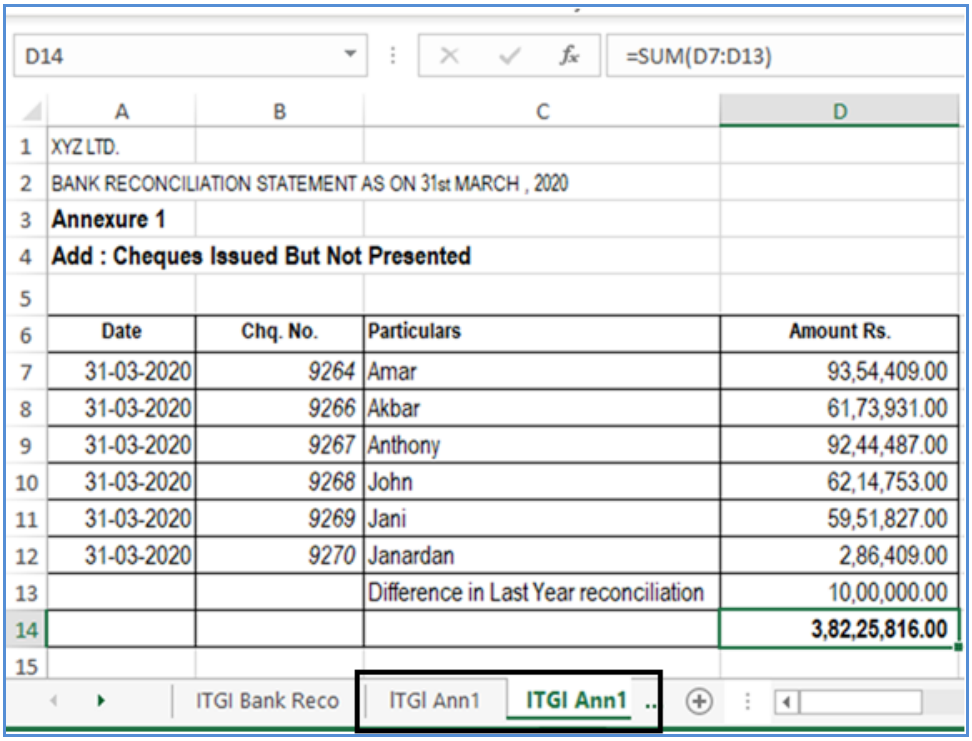


Figure 2.8.6 Similar Names

6. Now you have discovered the sheet from where the value is being referenced.
7. But there is another challenge you now have two sheets with the same name ITGI Ann1
8. In Excel, two sheets cannot have the same name.
9. If you observe closely, in one it is “L” in lower case and the other one consists of “I” in upper case visually they look similar.
10. You discover how the fraudster has created the illusion.

Gist: You have understood how to hide and unhide sheets, and as a control to check the integrity of the workbook.

Commands learnt: Home > Cells>Format>Hide/Unhide

2.8.4 REARRANGING WORKSHEETS

You can easily rearrange any of your sheets just by dragging their tabs from one place to another, as shown in Figure 2.8.7

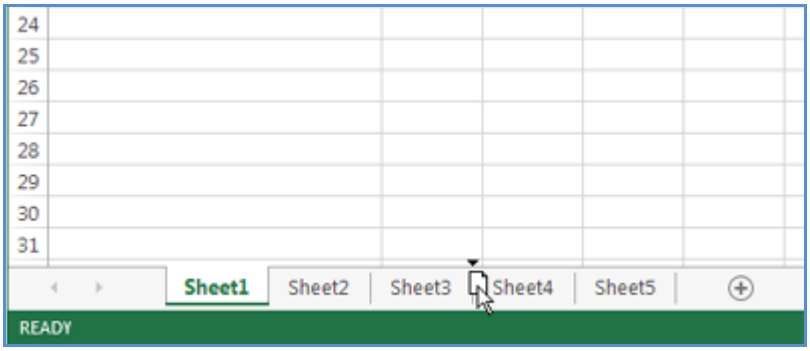


Figure 2.8.7 Rearrange Sheets

Tip: You can use a similar technique to create copies of a worksheet. Click the worksheet tab and begin dragging, just as you would to move the worksheet. Before you release the mouse button, press the Ctrl key (you will see a plus sign [+] appear). Keep holding the Ctrl key until you release the mouse button, at which point Excel creates a copy of the worksheet in the new location.

2.8.5 MOVING AND COPYING WORKSHEETS FROM ONE WORKBOOK TO ANOTHER

You may want to move (or copy) a worksheet from one Excel file to another, as follows:

Open both spreadsheet files.

- The file that contains the worksheet you want to **move or copy** is called the source file;
- The other file where you want to place the worksheet copy is the destination file.

Go to the source workbook.

- **Right-click** the worksheet you want to transfer,
- From the shortcut menu that appears,
- choose Move or Copy.
    - To transfer multiple worksheets at once, hold down the Ctrl key, > select all the worksheets you want to move or copy.
    - Excel highlights all the worksheets you select (and groups them together).
    - Right-click the selection, and then choose Move or Copy.

    Choose the destination file from the **"To book"** drop-down list.
- The "To book" menu shows all the currently open workbooks (including the source workbook).

Specify where you want to insert the worksheet.

If you want to copy the worksheet, turn on the "**Create a copy**" checkbox at the bottom of the window.

Click OK

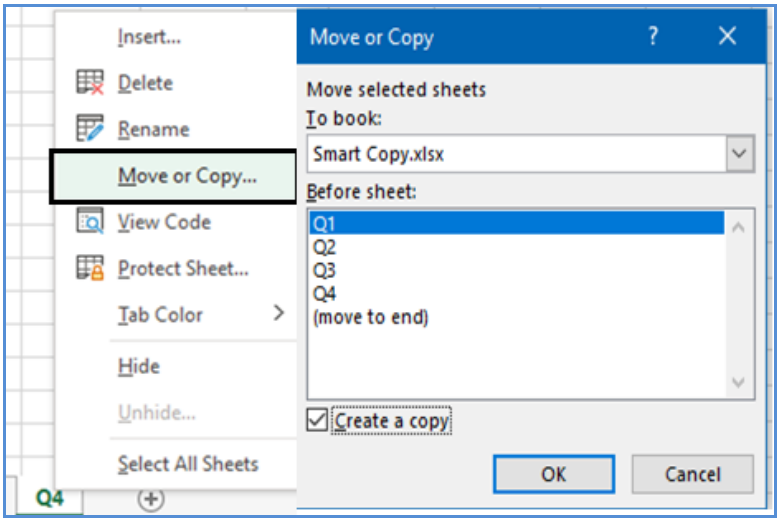


Figure 2.8.9 Moving/ Copying Sheets

Tip: You can also move worksheets to a new workbook, which it automatically creates for you. To do so, choose "(new book)" in the "To book" list. The resulting workbook has only the worksheets you transferred to it.

2.9 MANAGING ROWS AND COLUMNS

Every worksheet has exactly 1,048,576 rows and 16,384 columns.

2.9.1 INSERTING ROWS AND COLUMNS

The number of rows and columns in a worksheet is fixed, you can still insert and delete rows and columns if you need to make room for additional information. These operations do not change the number of rows or columns. Instead, inserting a new row moves down the other rows to accommodate the new row. The last row is simply removed from the worksheet if it is empty. Inserting new column shifts the columns to the right, and the last column is removed if it is empty.

There are few methods to insert rows or columns

Method 1

Select an entire row or multiple rows by clicking the row numbers in the worksheet border. Right-click and choose Insert from the shortcut menu.

Method 2

Move the cell pointer to the row that you want to insert, and then choose **Home > Cells > Insert > Insert Sheet Rows**. If you select multiple cells in the column, additional rows that correspond to the number of cells selected in the column are inserted and moves the rows below the insertion down.

Columns can be inserted similarly by clicking on Column where you want to insert.

2.9.2 DELETING ROWS AND COLUMNS

You can also delete rows or columns.

To delete a row or multiple rows

- Select the rows, Right-click, and choose **Delete** from the shortcut menu.
- Another method is through ribbon option **Home >Cells > Delete Sheet Rows**.

The same process can be used for deleting columns, Select the column(s) and choose **Delete** from the shortcut menu.

2.9.3 HIDING ROWS AND COLUMNS

You can hide rows or columns. Hiding rows and columns may be useful if you do not want users to see some particular information, or you do not want to print it.

To hide rows in your worksheet

- Select the row or rows that you want to hide by clicking in the row header on the left.
- Then right-click and choose **Hide** from the shortcut menu.
- Another method using Ribbon commands on the **Home > Cells >Format > Hide & Unhide** drop-down list.

2.10 SUMMARY

Excel is a versatile tool for data manipulation, analysis, report building, visualisation.

In this chapter, you learned the basics of Excel and a basic overview of the screen including an understanding of Ribbons. You have learned that you can increase efficiency by using Quick Access Toolbar.

You have also learnt about various Data types and the new rich data types like Stocks and Geography. Excel offers a host of features for customisation options. You learned how Go to Select can assist you in Looking for consistencies, finding irregularities, and clean data.

Finally, you learnt how to smart copy and paste including various options in paste special

In addition, we learned to manage worksheets, rows, and columns.

References

[1] Michael Alexander, Richard Kusleika, John Walkenbach, 'Excel 2019 Bible', Wiley,2018

[2] Greg Harvey, 'Excel 2019 All-in-One for Dummies, John Wiley & Sons,2018

[3] www.chandoo.org

CHAPTER

3 EXCEL – MAKING DATA CONSISTENT

LEARNING OBJECTIVES

- Understanding Naming of cells, ranges
- Understanding insertion of Tables in Excel
- Understanding Data Validation features of Excel
- Understanding creation of Dependent Drop-Down list
- Understanding INDIRECT Function

3.1 INTRODUCTION

Consistency is an important feature in Excel and makes the Excel sheets more meaningful and understandable. In this chapter you will learn how Naming a Cell or a range makes the formulas more understandable for any user of the sheet. Tables feature in Excel brings consistency in format, formulas, and makes the tabular data more manageable. Data Validation is an outstanding feature in Excel which aids in defining restrictions on what data can or should be entered in a cell.

3.2 NAMING IN EXCEL

In real world you are more used to names whether for people, objects or location. You understand Delhi more easily rather than some Latitude & Longitude reference, same way dealing with cryptic cell and range addresses can sometimes be confusing.

Excel allows you to assign human understandable descriptive names to cells and ranges. Rather than saying A2:A50 it would be more meaningful to call it Sales and while referring in formulas, you can have a formula =SUM(Sales) rather than =SUM(A2:A50).

Advantages

- Navigation becomes easy with names to areas of your worksheet either by using the Name box, or pressing F5 Key and specifying the range name.

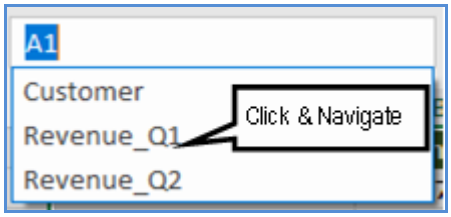


Figure 3.2.1 Name Box

- Chances of error reduce when writing a name rather than cell addresses.
- Writing formulas become much convenient. You can paste a cell or range name into a formula by using F3.
- Names make your formulas more meaningful for other users of the worksheets. A formula such as =Income—Expenditure is more meaningful than =B10-B50.

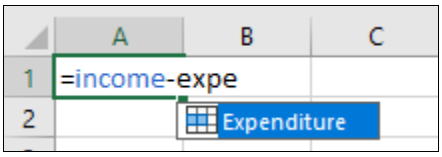


Figure 3.2.2 Formula with Name

Rules for Creating Names

There are some rules for naming in Excel which need to be followed in order to name a range.

- The first character of a name must be one of the following characters:
  - letter
  - underscore (_)
  - backslash (\).
- Remaining characters in the name can be
  - letters
  - numbers
  - periods
  - underscore characters
- The following are not allowed:
  - Space characters are not allowed as part of a name.
  - Names can't look like cell addresses, such as C$58 or B1C2.
  - Names are not case sensitive. For example, East and EAST are treated as the same name.
- Names cannot be more than 255 characters long.

3.2.1 HOW TO CREATE NAMES

There are broadly three methods for creating names:

Method 1

Using the Name box

The fastest way to create a named range is to use the Name box.

- Select a cell or a range of cells that you want to name

- Type a name into the Name Box
- Press the Enter key

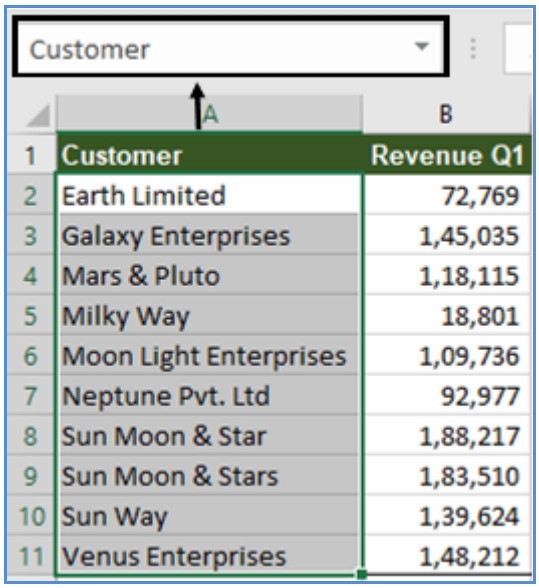


Figure 3.2.3 Name using Name Box

The Name box is a drop-down list and shows all names in the workbook.

Method 2

New Name dialog box

For more control over naming cells and ranges, use the New Name dialog box.

- Select the cell(s).
- On the **Formulas** tab, in the **Define Names** group, click the '**Define Name'** button.
- In the **New Name dialog box**, specify three things:
  - In the **Name** box, type the range name.
  - In the **Scope** dropdown, set the name scope by default it will be workbook.
  - In the **Refers** to box, your selected range will appear.
- Click OK to save the changes and close the dialog box.

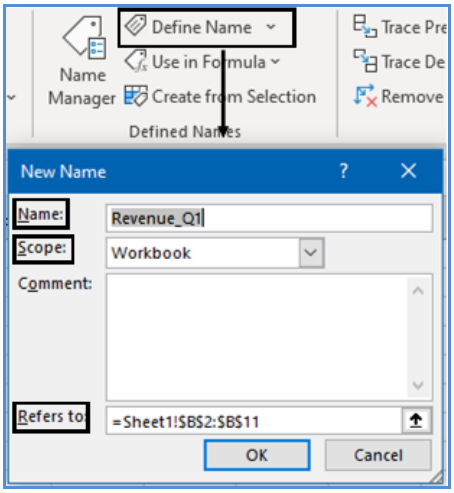


Figure 3.2.4 Name using Define name

Method 3

Using Create from Selection Option

You may have a worksheet that contains the text that you want to use for names for adjacent cells or ranges, as an example when you have data in tabular form, and you want to create named range for each column/row.

For example, in the dataset below, if you want to quickly create three named ranges (Customer, Revenue_Q1, Revenue_Q2), then you can use the method shown below.

- Select the entire table including the column headers.
- Go to the **Formulas tab > Define Names group** and click the **Create from Selection** button.
- You can also press the keyboard shortcut **Ctrl + Shift + F3**.
- **Create Names from Selection** dialogue box will open.
- Select the column or row with headers, or both, and click OK as shown in Figure.

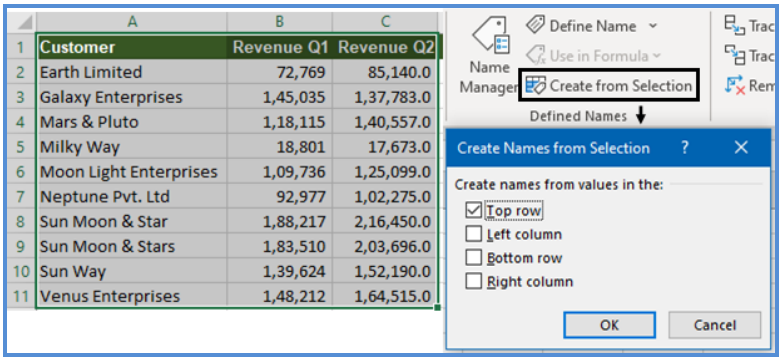


Figure 3.2.5 Name using create from selection

As a result, Excel will create 3 named ranges, picking up names from the Column headers automatically as you can see in name manager. Excel automatically replaces any spaces between words in the header labels, with underscores (_).

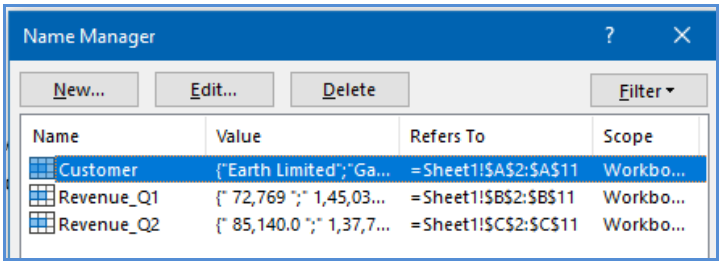


Figure 3.2.6 Name using create from selection

3.2.2 EXCEL NAME MANAGER

A workbook can have any number of named cells and ranges. In the **Formula** tab Excel **Name Manager** is specially designed to manage names: change, filter, or delete existing names as well as create new ones.

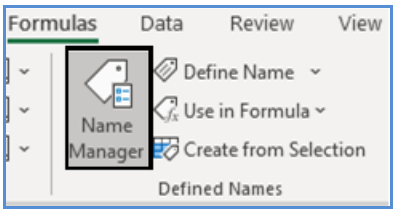


Figure 3.2.7 Name Manager

When you click on **Name Manager** or press **CTRL+F3,** Name manager dialog box appears where you can see all names in the current workbook immediately. You can select the name you want to work with and click one of the 3 buttons at the top of the window edit, delete, or filter. As an example you want to edit the name of a range as shown in figure 3.8, you click on edit, it will open the **Edi**t Name dialog box where you can change the name and reference. Scope however cannot be changed.

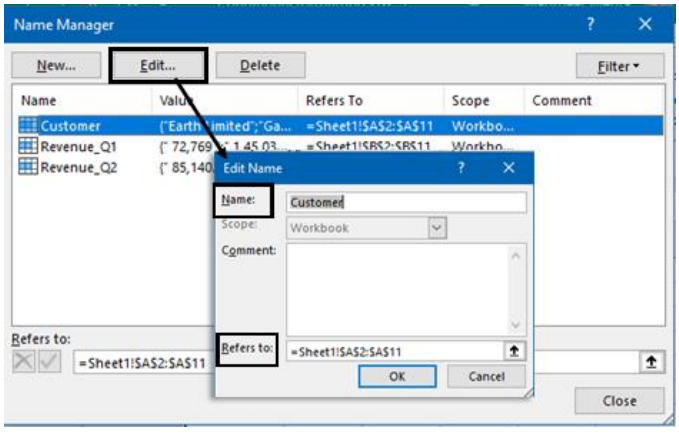


Figure 3.2.8 Edit name

In case you want to delete a named range, select it in the Name Manager and click the **Delete** button at the top.

3.2.3 USING NAMED RANGES

You can use names in formulas. For example, you could have a group of cells with Revenue figures for Q1. Name those cells **Revenue_Q1**, then use this formula to calculate the total amount **=SUM,** and rather than giving range, you can simply type **Revenue_Q1** or you can press **F3,** it will open **Paste name** dialog box and you can select the name you want to refer as shown in Figure 3.9. This formula will be more meaningful for any user of the spreadsheet.

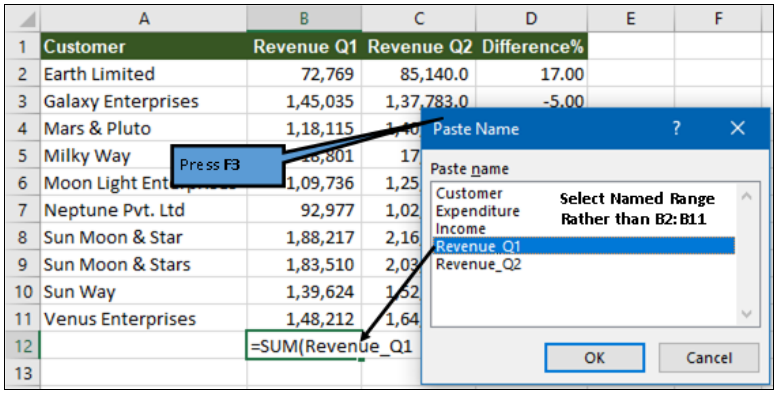


Figure 3.2.9 Using named ranges in a formula

***CASE STUDY**: In a Forensic investigation your principal CA Chandiwala gives you an Excel sheet as shown in Figure 3.2.10 with calculations and cross-verification of GST amount which he wants to produce as evidence in court. He wants you to make and structure the Excel sheet in a way so that any user should be able to understand the calculation. How will you do it?*

	A	B	C	D	E	F
1	**Customer**	**Revenue Q1**	**GST 18%**		**GST**	18%
2	Earth Limited	72,769	13,098			
3	Galaxy Enterprises	1,45,035	26,106			
4	Mars & Pluto	1,18,115	11,261		GST Aggregate	
5	Milky Way	18,801	3,384		Actual	
6	Moon Light Enterprises	1,09,736	19,752		Difference	
7	Neptune Pvt. Ltd	92,977	16,736			
8	Sun Moon & Star	1,88,217	33,879			
9	Sun Moon & Stars	1,83,510	33,032			
10	Sun Way	1,39,624	25,132			
11	Venus Enterprises	1,48,212	26,678			
12	**Actual Amount**					

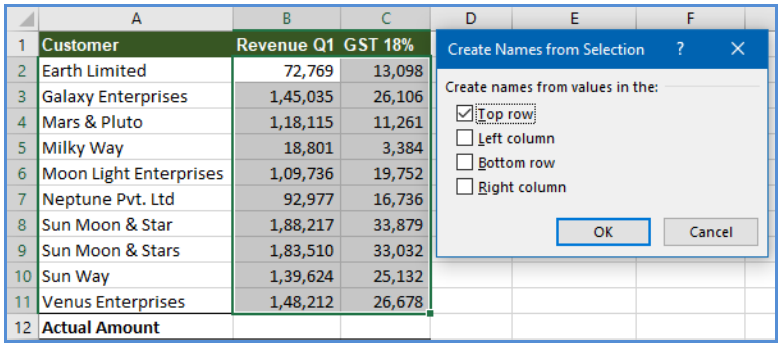
Figure 3.2.10 Excel sheet with calculations

Strategy:

You can use the feature of **Named Ranges** to make the formulas meaningful.

1. Select the data B1: C11
2. Go to the **Formulas tab > Define Names** group and click the Create from Selection button

3. Create Names from Selection dialogue box will open
4. Select the Top Row as shown in Figure
5. It will name the B2 to B11 as Revenue_Q1 and C2 to C11 as GST_18%



	A	B	C
1	Customer	Revenue Q1	GST 18%
2	Earth Limited	72,769	13,098
3	Galaxy Enterprises	1,45,035	26,106
4	Mars & Pluto	1,18,115	11,261
5	Milky Way	18,801	3,384
6	Moon Light Enterprises	1,09,736	19,752
7	Neptune Pvt. Ltd	92,977	16,736
8	Sun Moon & Star	1,88,217	33,879
9	Sun Moon & Stars	1,83,510	33,032
10	Sun Way	1,39,624	25,132
11	Venus Enterprises	1,48,212	26,678
12	Actual Amount		

Figure 3.2.11 Naming the ranges

6. Next Go to Cell **F1** which contains the GST rate and
7. Type a name into the Name Box "**GST_Rate**"
8. Press the **Enter** key as shown in figure
9. Type a name into the Name Box

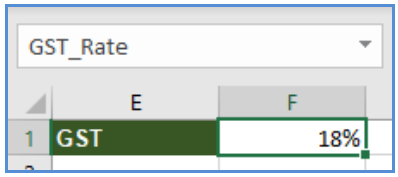


Figure 3.2.12 Naming a Cell

10. Next Go to cell **B12** and enter formula =*SUM(Revenue_Q1)* Similarly in **C12** enter Formula =*SUM(GST_18)*.

B	C
Revenue Q1	GST 18%
72769	13098
145035	26106
118115	11261
18801	3384
109736	19752
92977	16736
188217	33879
183510	33032
139624	25132
148212	26678
=SUM(Revenue_Q1)	=SUM(GST_18)

Figure 3.2.13 Formula with a named range

11. You will now go to cell F4 and type the formula =SUM(Revenue_Q1)*GST_Rate, which is much more meaningful and understandable by any user of Excel rather than cryptic =SUM(B2:B11)*F1.

	A	B	C	D	E	F	G
1	**Customer**	**Revenue Q1**	**GST 18%**		**GST**	18%	
2	Earth Limited	72,769	13,098				
3	Galaxy Enterprises	1,45,035	26,106				
4	Mars & Pluto	1,18,115	11,261		GST Aggregate	=SUM(Revenue_Q1)*GST_Rate	
5	Milky Way	18,801	3,384		Actual		
6	Moon Light Enterprises	1,09,736	19,752		Difference		
7	Neptune Pvt. Ltd	92,977	16,736				
8	Sun Moon & Star	1,88,217	33,879				
9	Sun Moon & Stars	1,83,510	33,032				
10	Sun Way	1,39,624	25,132				
11	Venus Enterprises	1,48,212	26,678				
12	**Actual Amount**	12,16,996	2,09,058				

Figure 3.2.14 Formula with names

12. You can name this cell **F4** as **Calculated_GST**
13. You will then copy the actual GST from Cell **C12** to F5 and name it **Actual_GST,**
14. In Cell **F6** you will then type the formula by pressing **Key F3** and in Paste Name selecting **Calculated_GST-Actual_GST.**

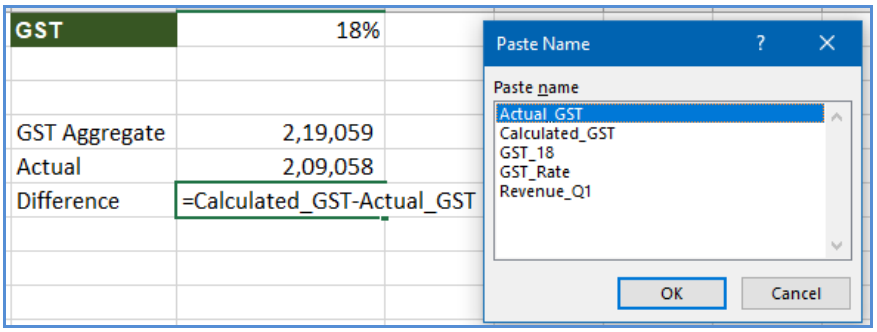


Figure 3.2.15 Paste Name with F3

15. All these formulas are transparent and can easily be understood by any user of the Excel sheet.

Gist: You have understood the use of the Name Manager to make the formulas meaningful and understandable by any user of Excel sheet.

Commands learnt: **Formulas tab > Define Names> Create from Selection, and using Name Box.**

3.3 TABLES

A table is a rectangular range of data that usually has a row of text headings to describe the contents of each column. Any data which is organized in rows and columns is not a table. The magic happens when you convert that range of cells into a Table.

Excel table is a special feature that allows you to manage the table's contents independently from the rest of the worksheet data. Excel's table feature makes common tasks much easier — and a lot better looking and Excel works very intelligently in a Table

You can create and use an Excel table to manage and analyze data easily. Further, with Excel Tables, you get built-in Filtering, Sorting, and Row Banding that ease your reporting activities.

Table Vs Range

A table is a more structured way of working with data than a range.

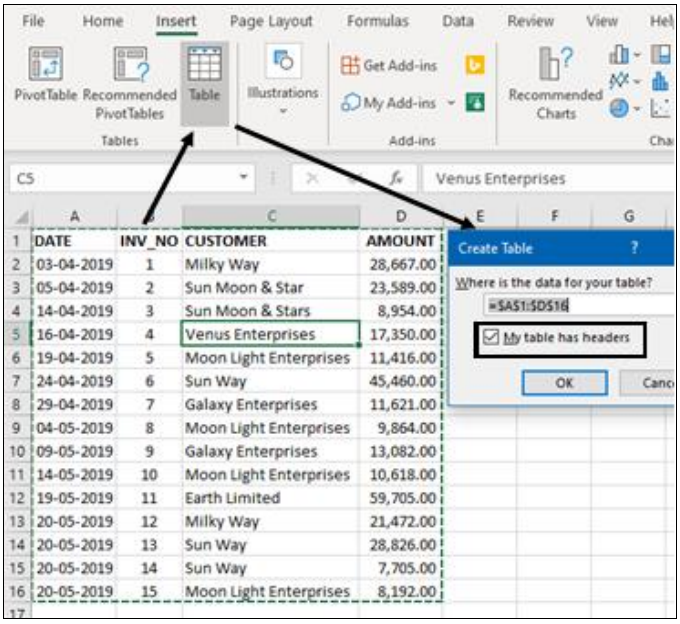
You can convert a range into a table and Excel automatically provides −

- Table Name
- Column Header Names
- Formatting to the Data (Cell Color and Font Color) for better Visualization

3.3.1 INSERT AND FORMATTING A TABLE

Wherever the data is organized in rows and columns, you can easily covert this range of cells into a table:

- Select any cell within your data set
- On the **Insert** tab, in the Tables group, click the **Table** button
- Press the Ctrl + T shortcut
- The **Create Table** dialog box appears with all the data selected for you automatically; you can adjust the range if needed
- Select **My table has headers** box to convert your first row as headers



	A	B	C	D
1	DATE	INV_NO	CUSTOMER	AMOUNT
2	03-04-2019	1	Milky Way	28,667.00
3	05-04-2019	2	Sun Moon & Star	23,589.00
4	14-04-2019	3	Sun Moon & Stars	8,954.00
5	16-04-2019	4	Venus Enterprises	17,350.00
6	19-04-2019	5	Moon Light Enterprises	11,416.00
7	24-04-2019	6	Sun Way	45,460.00
8	29-04-2019	7	Galaxy Enterprises	11,621.00
9	04-05-2019	8	Moon Light Enterprises	9,864.00
10	09-05-2019	9	Galaxy Enterprises	13,082.00
11	14-05-2019	10	Moon Light Enterprises	10,618.00
12	19-05-2019	11	Earth Limited	59,705.00
13	20-05-2019	12	Milky Way	21,472.00
14	20-05-2019	13	Sun Way	28,826.00
15	20-05-2019	14	Sun Way	7,705.00
16	20-05-2019	15	Moon Light Enterprises	8,192.00

Figure 3.3.1 Create Table

- Click OK
- The result will be, Excel converts your range into a Table

	A	B	C	D
1	DATE	INV_N	CUSTOMER	AMOUNT
2	03-04-2019	1	Milky Way	28,667.00
3	05-04-2019	2	Sun Moon & Star	23,589.00
4	14-04-2019	3	Sun Moon & Stars	8,954.00
5	16-04-2019	4	Venus Enterprises	17,350.00
6	19-04-2019	5	Moon Light Enterprises	11,416.00
7	24-04-2019	6	Sun Way	45,460.00
8	29-04-2019	7	Galaxy Enterprises	11,621.00
9	04-05-2019	8	Moon Light Enterprises	9,864.00
10	09-05-2019	9	Galaxy Enterprises	13,082.00
11	14-05-2019	10	Moon Light Enterprises	10,618.00
12	19-05-2019	11	Earth Limited	59,705.00
13	20-05-2019	12	Milky Way	21,472.00
14	20-05-2019	13	Sun Way	28,826.00
15	20-05-2019	14	Sun Way	7,705.00
16	20-05-2019	15	Moon Light Enterprises	8,192.00

Figure 3.3.2 Converted to Table

Formatting a Table

- You can choose from several table styles in Table Design. These styles can be used depending on what colour and pattern you want to display your data in the table. As an example, you can select a green colour style from Quick Styles as shown in Figure

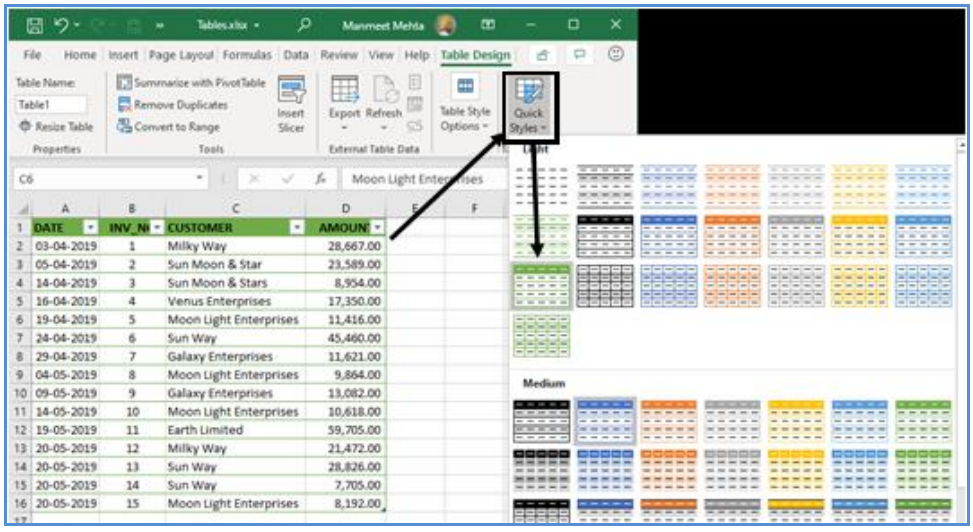


Figure 3.3.3 Formatting a Table

Naming a Table

Excel assigns a name to every table that is created.

To look at the name of the table you just created

- Click in the table, as soon as you click, a Contextual **Table Design** tab will appear on the Ribbon.
- In the Properties group, in the Table Name box, your Table Name will be displayed.
- You can edit this Table Name to make it more meaningful to your data.
- Click the Table Name box. Clear the Name and type **Tbl_Sales.**

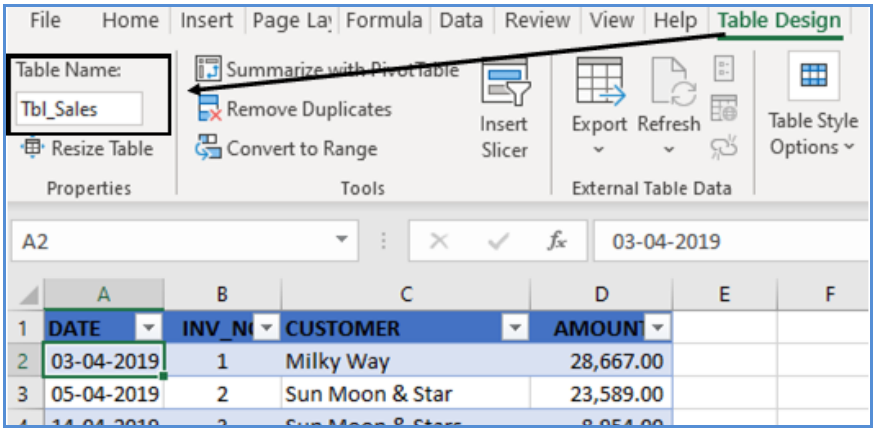


Figure 3.3.4 Naming a Table

3.3.2 BENEFITS OF A TABLE

- Excel provides Contextual **Table Design** in the ribbon ranging from properties to styles.
- If you have multiple rows in a table, and you scroll down the sheet so that the header row disappears, the **column letters in the worksheet are replaced by the table headers**.
- Excel automatically provides a **Sort and Filter** button in each column header to sort the data or filter the table.
- When you write a formula in any cell in a column of the table, it gets **copied consistently** to all the cells in that column.
- You can use **table name** and column header names in the formulas, without having to use cell references or creating range names.
- You can **Auto-extend** the table size by adding more rows or more columns by just typing in the first cell of a column or row.
- You can create and use **slicers** for a table for filtering data

3.3.3 ALWAYS VISIBLE COLUMN HEADINGS WHILE SCROLLING

When you are working with a large table that does not fit on a screen, the header row always remains visible when you scroll down. Rather than Column A, B, etc. you will find proper column names as headers.

	DATE	INV_NO	CUSTOMER	AMOUNT
4	14-04-2019	3	Sun Moon & Stars	8,954.00
5	16-04-2019	4	Venus Enterprises	17,350.00
6	19-04-2019	5	Moon Light Enterprises	11,416.00
7	24-04-2019	6	Sun Way	45,460.00
8	29-04-2019	7	Galaxy Enterprises	11,621.00
9	04-05-2019	8	Moon Light Enterprises	9,864.00

Figure 3.3.5 Always Visible Column Headings

FILTER A TABLE IN EXCEL

When you create a Table auto-filter is activated by default. To filter the table's data:

- Click the drop-down arrow in the column header.
- Uncheck the boxes next to the data you want to filter out or
- Uncheck the Select All box to deselect all the data, and then check the boxes next to the data you want to show.
- You can also use the Filter by Color and various options where appropriate.
- Click OK.

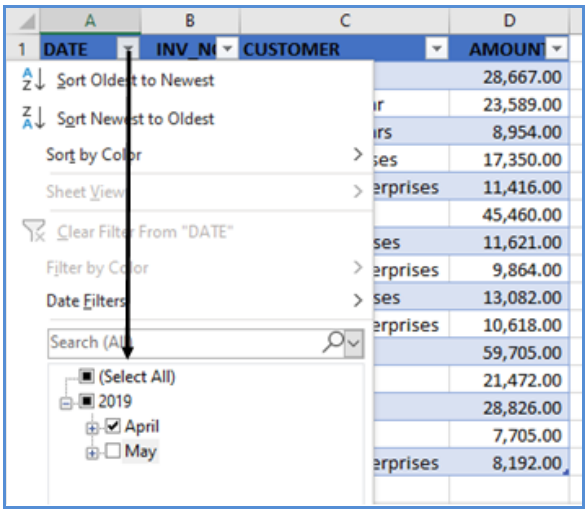


Figure 3.3.6 Filter in Table

- You can, however, Turn off the filter from Table Design Option

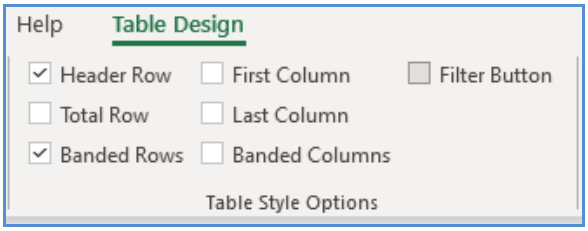


Figure 3.3.7 Turn Off Filter

SUM TABLE COLUMNS

Excel table has a wonderful feature called Total Row which can summarize data without formulas.

To sum a table's data.

- Select any cell in the table
- On the Table Design tab, in the Table Style Options group, put a tick mark in the Total Row box
- It will Add the Total row to a table
- The Total row is inserted at the bottom of the table and shows the total in the last column
- The data in the last column is summed
- It also has the capability to calculate data in a different way, e.g. count or average, just select the corresponding function

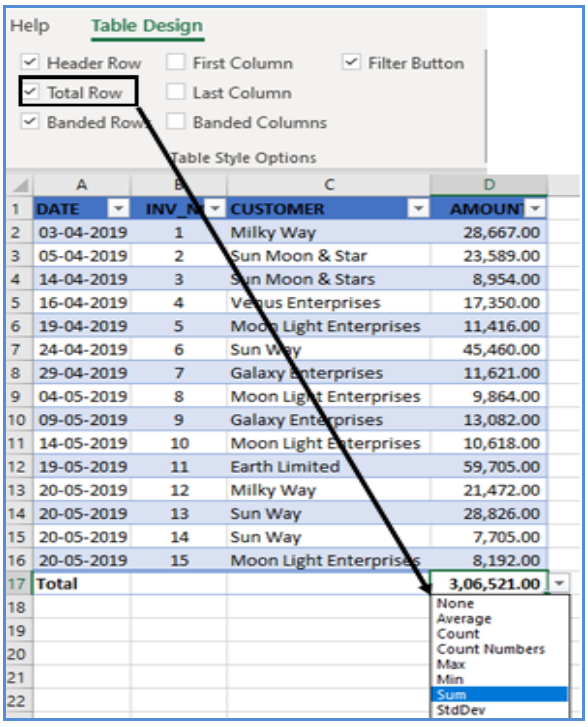


Figure 3.3.8 Total in Table

AUTO-EXPANSION OF TABLE TO INCLUDE NEW DATA

In an Excel worksheet adding more rows or columns means more formatting and reformatting. When you type anything next to a table, Excel assumes you want to add a new entry to it and expands the table to include that entry.

	DATE	INV_NO	CUSTOMER	AMOUNT
10	09-05-2019	9	Galaxy Enterprises	13,082.00
11	14-05-2019	10	Moon Light Enterprises	10,618.00
12	19-05-2019	11	Earth Limited	59,705.00
13	20-05-2019	12	Milky Way	21,472.00
14	20-05-2019	13	Sun Way	28,826.00
15	20-05-2019	14	Sun Way	7,705.00
16	20-05-2019	15	Moon Light Enterprises	8,192.00
17	20-05-2019			
18				

Figure 3.3.9 Auto Expansion of rows in the table

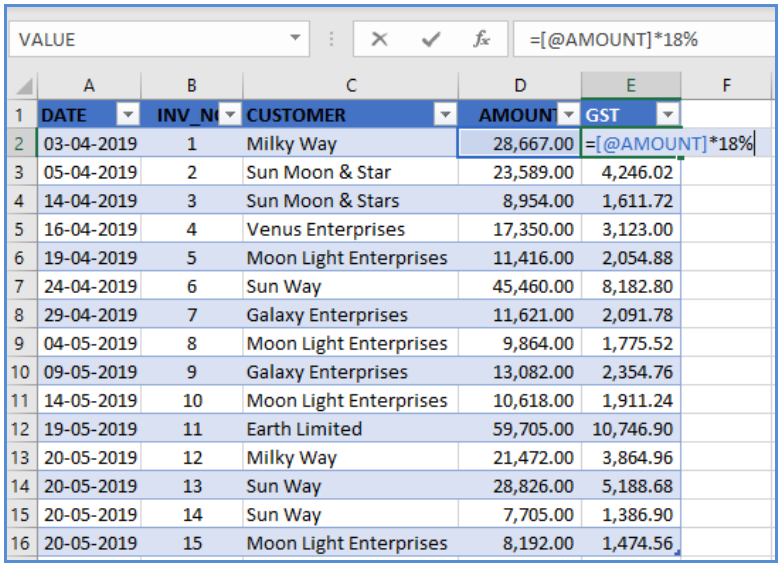
3.3.4 EXCEL TABLE FORMULAS

In Table data, Excel uses a special formula syntax called **structured references**. Compared to regular formulas, they have several advantages:

- **Easy-to-create**: When you select the table's data while making a formula, Excel will build a structured reference automatically.
- **Easy-to-read**: Structured references refer to the table parts by name, which makes formulas easier to understand.
- **Auto-filled:** When you enter a formula in any single cell, and it will be immediately copied throughout the column.
- **Changed automatically**: When you modify a formula anywhere in a column, the other formulas in the same column will change accordingly.
- **Updated automatically**: Every time the table is resized or the columns renamed, structured references update dynamically.

As an example, suppose you want to include a column for GST @ 18%.

You will just write GST in Cell E1 Table will Insert a column to the right of the column Amount. In any of the Cells in that empty column, type the Formula, =[@AMOUNT]*18%, and Press Enter. You will find that the formula propagates automatically to the other cells as shown in Figure



	A	B	C	D	E	F
1	DATE	INV_NO	CUSTOMER	AMOUNT	GST	
2	03-04-2019	1	Milky Way	28,667.00	=[@AMOUNT]*18%	
3	05-04-2019	2	Sun Moon & Star	23,589.00	4,246.02	
4	14-04-2019	3	Sun Moon & Stars	8,954.00	1,611.72	
5	16-04-2019	4	Venus Enterprises	17,350.00	3,123.00	
6	19-04-2019	5	Moon Light Enterprises	11,416.00	2,054.88	
7	24-04-2019	6	Sun Way	45,460.00	8,182.80	
8	29-04-2019	7	Galaxy Enterprises	11,621.00	2,091.78	
9	04-05-2019	8	Moon Light Enterprises	9,864.00	1,775.52	
10	09-05-2019	9	Galaxy Enterprises	13,082.00	2,354.76	
11	14-05-2019	10	Moon Light Enterprises	10,618.00	1,911.24	
12	19-05-2019	11	Earth Limited	59,705.00	10,746.90	
13	20-05-2019	12	Milky Way	21,472.00	3,864.96	
14	20-05-2019	13	Sun Way	28,826.00	5,188.68	
15	20-05-2019	14	Sun Way	7,705.00	1,386.90	
16	20-05-2019	15	Moon Light Enterprises	8,192.00	1,474.56	

Figure 3.3.10 Auto Propagation of Formula

3.3.5 REMOVE DUPLICATES

When you gather data from different sources, you may have duplicate values. If you need to remove the duplicate values before going further with analysis. There is a built-in feature in Tables to Remove Duplicates

- Click the table.
- On the Table DESIGN tab, click Remove Duplicates in the Tools group on the Ribbon.
- Remove Duplicates dialog box appears
- The column headers appear under columns in the Remove Duplicates dialog box.
- Check the column headers depending on which column you want to remove the duplicates and click OK.
- You will get a message on how many rows with duplicate values are removed and how many unique values remain. The cleaned data will be displayed in the table.

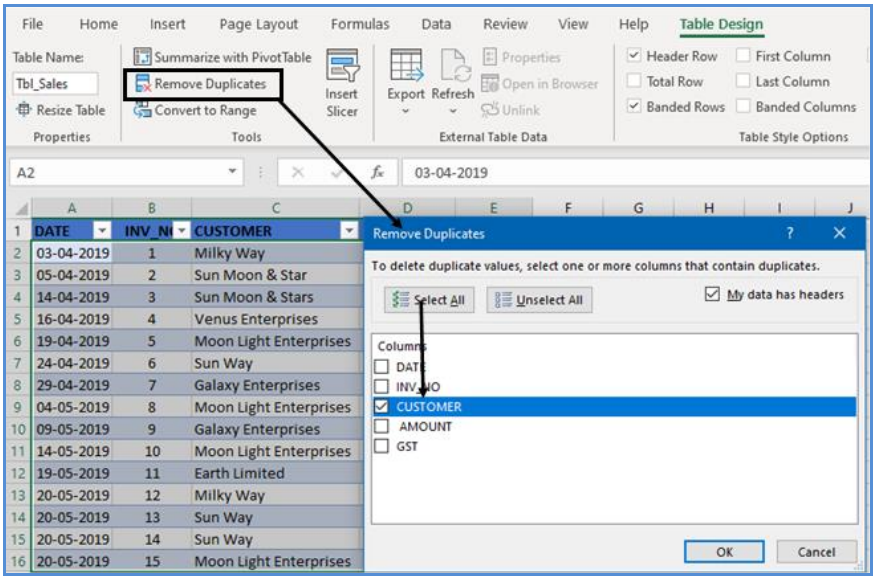


Figure 3.3.11 Remove Duplicates

3.3.6 CONVERTING A TABLE TO RANGE

To convert a table back to a range, just do the following:

- Right-click any cell in your table, and then click **Table> Convert to Range**, Or
- Click the **Convert to Range** button on the Table Design tab, in the Tools group
- In the dialog box that appears, click Yes
- Remove a table in Excel
- This will remove the table but retain all the data and formatting

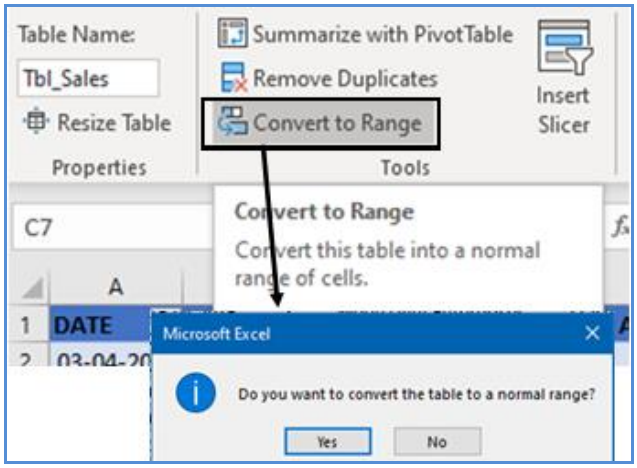
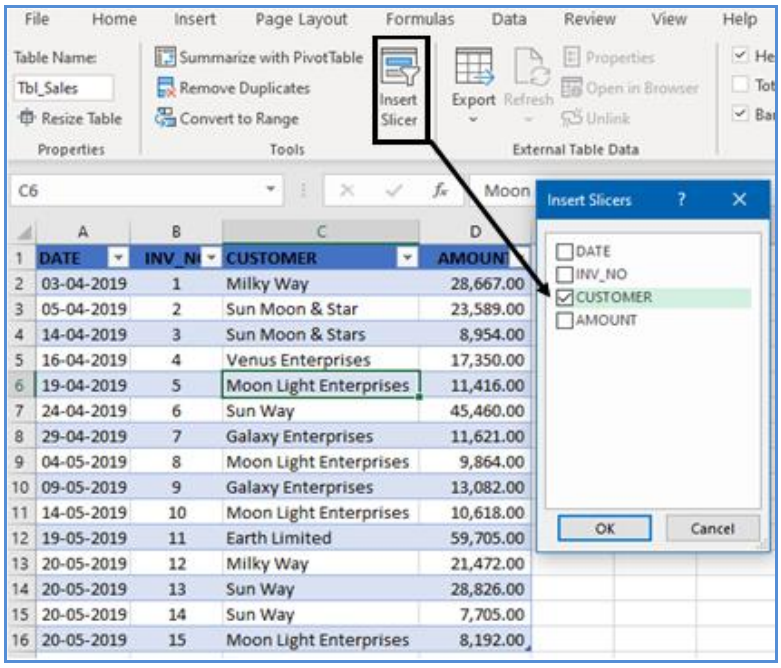


Figure 3.3.12 Convert to Range

3.3.7 SLICERS IN A TABLE

Slicers are a new feature introduced in Tables can also be used for filtering table data.

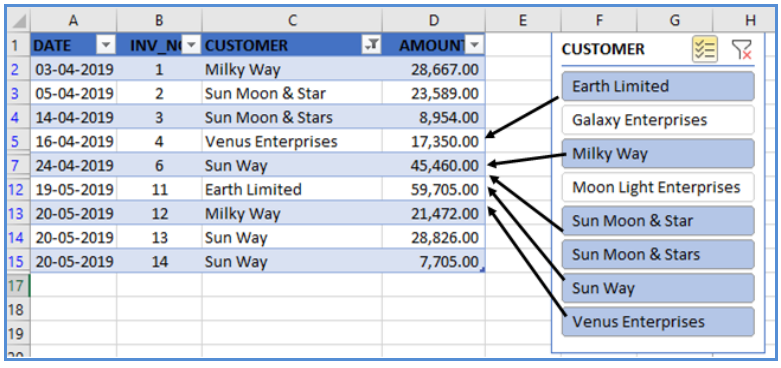
- You can add a slicer for your Excel table
- Go to the Design tab > Tools group, and click the **Insert Slicer** button.
- In the **Insert Slicers dialog box**, check the boxes for the columns that you want to create slicers for
- Click OK



	A	B	C	D
1	DATE	INV_NO	CUSTOMER	AMOUNT
2	03-04-2019	1	Milky Way	28,667.00
3	05-04-2019	2	Sun Moon & Star	23,589.00
4	14-04-2019	3	Sun Moon & Stars	8,954.00
5	16-04-2019	4	Venus Enterprises	17,350.00
6	19-04-2019	5	Moon Light Enterprises	11,416.00
7	24-04-2019	6	Sun Way	45,460.00
8	29-04-2019	7	Galaxy Enterprises	11,621.00
9	04-05-2019	8	Moon Light Enterprises	9,864.00
10	09-05-2019	9	Galaxy Enterprises	13,082.00
11	14-05-2019	10	Moon Light Enterprises	10,618.00
12	19-05-2019	11	Earth Limited	59,705.00
13	20-05-2019	12	Milky Way	21,472.00
14	20-05-2019	13	Sun Way	28,826.00
15	20-05-2019	14	Sun Way	7,705.00
16	20-05-2019	15	Moon Light Enterprises	8,192.00

Figure 3.3.13 Insert Slicers

- As a result, one or more slicers will appear in your worksheet.
- You simply click the items you want to show in your table.



	A	B	C	D
1	DATE	INV_NO	CUSTOMER	AMOUNT
2	03-04-2019	1	Milky Way	28,667.00
3	05-04-2019	2	Sun Moon & Star	23,589.00
4	14-04-2019	3	Sun Moon & Stars	8,954.00
5	16-04-2019	4	Venus Enterprises	17,350.00
7	24-04-2019	6	Sun Way	45,460.00
12	19-05-2019	11	Earth Limited	59,705.00
13	20-05-2019	12	Milky Way	21,472.00
14	20-05-2019	13	Sun Way	28,826.00
15	20-05-2019	14	Sun Way	7,705.00
17				
18				
19				

Figure 3.3.14 Slicers in a Table

CASE STUDY: ***In an Investigation your principal CA Chandiwala gives you an Excel sheet as shown in Figure3.3.15 where data Balance confirmations are being added periodically by various field assistants and the formatting is a mess, moreover he wants to maintain some consistency in formulas. He tells you to find a way in which formatting remains consistent and presentable even when data is added by different field assistants and formula could be added to consistently find the difference between Book balances and confirmations. You are a smart professional, is there a smart way to format this sheet?***

	A	B	C	D	E
1	S No.	DATE OF RESPONSE	CUSTOMER	Balance Books of Acc	Balance Confirmed
2	1	13-04-2020	ABC Inc.	8,80,797.00	7,70,797.00
3	2	05-04-2020	DEF Ltd.	7,07,965.00	7,07,965.00
4	3	14-04-2020	GHI LLP	55,101.00	55,101.00
5	4	16-04-2020	KLM Enterprises	6,09,739.00	6,09,739.00
6	5	19-04-2020	NOP Limited	9,10,153.00	7,10,153.00
7	6	24-04-2020	QRS Machines	82,041.00	82,041.00
8	7	29-04-2020	TUV P. Ltd.	8,78,148.00	9,28,148.00
9	8	04-05-2020	WXY Associates	8,53,976.00	8,53,976.00
10	9	09-05-2020	CAD Way	1,45,522.00	1,45,522.00
11	10	14-05-2020	HIK Associates	7,13,220.00	7,13,220.00

Figure 3.3.15 Excel sheet with poor formatting

Strategy:

You can use the feature of **Tables** in excel to make the data, formatting and consistent.

1. Select anywhere in the data.
2. Got to **Insert > Tables** and click or use shortcut CTRL+T
3. A Dialog box '**Create Table'** will appear, select '**My table has headers'**

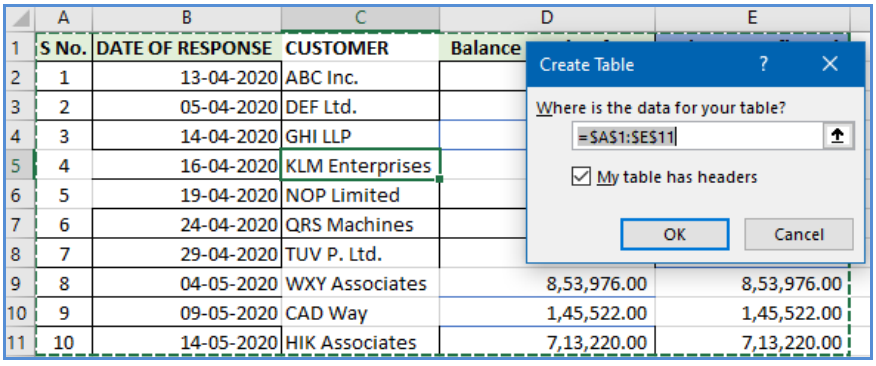


Figure 3.3.16 Create Table

4. Click OK
5. The data is converted to Table

S No	DATE OF RESPONSE	CUSTOMER	Balance Books of Ac	Balance Confirmed
1	13-04-2020	ABC Inc.	8,80,797.00	7,70,797.00
2	05-04-2020	DEF Ltd.	7,07,965.00	7,07,965.00
3	14-04-2020	GHI LLP	55,101.00	55,101.00
4	16-04-2020	KLM Enterprises	6,09,739.00	6,09,739.00
5	19-04-2020	NOP Limited	9,10,153.00	7,10,153.00
6	24-04-2020	QRS Machines	82,041.00	82,041.00
7	29-04-2020	TUV P. Ltd.	8,78,148.00	9,28,148.00
8	04-05-2020	WXY Associates	8,53,976.00	8,53,976.00
9	09-05-2020	CAD Way	1,45,522.00	1,45,522.00
10	14-05-2020	HIK Associates	7,13,220.00	7,13,220.00

Figure 3.3.17 Table created

6. But the formatting is not Good
7. Go to **Table design > Quick Styles,** select the desired style and right-click,

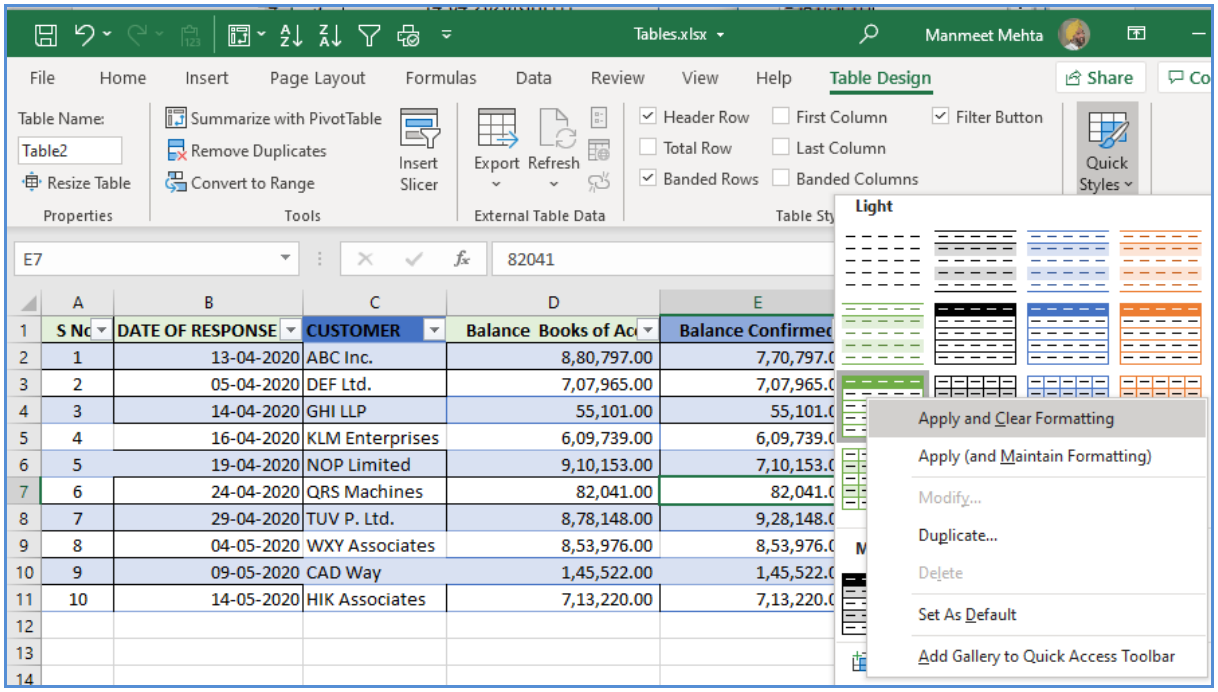


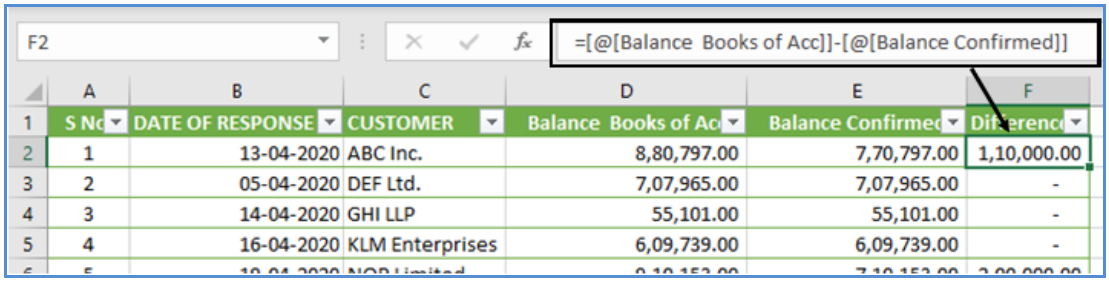
Figure 3.3.18 Quick Style

8. Select the option '**Apply and clear formatting**'
9. It will delete all previous formatting and apply new formatting consistently

	A	B	C	D	E
1	S No	DATE OF RESPONSE	CUSTOMER	Balance Books of Ac	Balance Confirmed
2	1	13-04-2020	ABC Inc.	8,80,797.00	7,70,797.00
3	2	05-04-2020	DEF Ltd.	7,07,965.00	7,07,965.00
4	3	14-04-2020	GHI LLP	55,101.00	55,101.00
5	4	16-04-2020	KLM Enterprises	6,09,739.00	6,09,739.00
6	5	19-04-2020	NOP Limited	9,10,153.00	7,10,153.00
7	6	24-04-2020	QRS Machines	82,041.00	82,041.00
8	7	29-04-2020	TUV P. Ltd.	8,78,148.00	9,28,148.00
9	8	04-05-2020	WXY Associates	8,53,976.00	8,53,976.00
10	9	09-05-2020	CAD Way	1,45,522.00	1,45,522.00
11	10	14-05-2020	HIK Associates	7,13,220.00	7,13,220.00

Figure 3.3.19 Clean Style

10. Name the Table as Tbl_DrBal
11. In Cell F1 type Difference and click a new column with the same formatting will be inserted.
12. In cell F2 you can type formula **=[@[Balance Books of Acc]]-[@[Balance Confirmed]].**
13. The formula will be consistently propagated to all the cells in the '**difference**' column.
14. Now if anyone starts typing in the next column Consistent formatting and formulas are maintained.
15. The data also becomes more presentable.



	A	B	C	D	E	F
1	S No	DATE OF RESPONSE	CUSTOMER	Balance Books of Ac	Balance Confirmed	Difference
2	1	13-04-2020	ABC Inc.	8,80,797.00	7,70,797.00	1,10,000.00
3	2	05-04-2020	DEF Ltd.	7,07,965.00	7,07,965.00	-
4	3	14-04-2020	GHI LLP	55,101.00	55,101.00	-
5	4	16-04-2020	KLM Enterprises	6,09,739.00	6,09,739.00	-

Figure 3.3.20 Consistent formula

Gist: You have understood the use of **Tables** to make the Formatting and formulas Consistent

Commands learnt: **Insert > Tables.**

3.4 DATA VALIDATION

When we are the only one populating an Excel worksheet with data, it's not too likely that the wrong data will be entered. But things are not the same if many people are populating an Excel worksheet. Fortunately, Excel has data Validation.

The Excel data validation feature allows you to set up rules that restrict or validate what can be entered into a cell. Technically, you create a validation rule that controls what kind of data can be entered into a certain cell and we can

lock out certain types of errors and guide the people using our workbook to make sure they fill in the right information.

Among other things, we can use **data validation** to do the following:

- Restrict data to predefined items in a list — eg Only Specified Departments
- Restrict numbers outside a specified range — eg between 100 and 600
- Restrict dates outside a certain time frame — eg not beyond some end date
- Restrict times outside a certain time frame — eg only up to 05:00 PM
- Limit the number of text characters — eg 10 digit PAN
- Validate data based on formulas or values in other cells — eg. Salary not exceeding Rs.4,00,000

Data validation options are located on the Data tab, in the Data Tools group. We can choose to show an input message when the user selects the cell. Input messages are generally used to offer users guidance about the type of data that we want entered in the cell

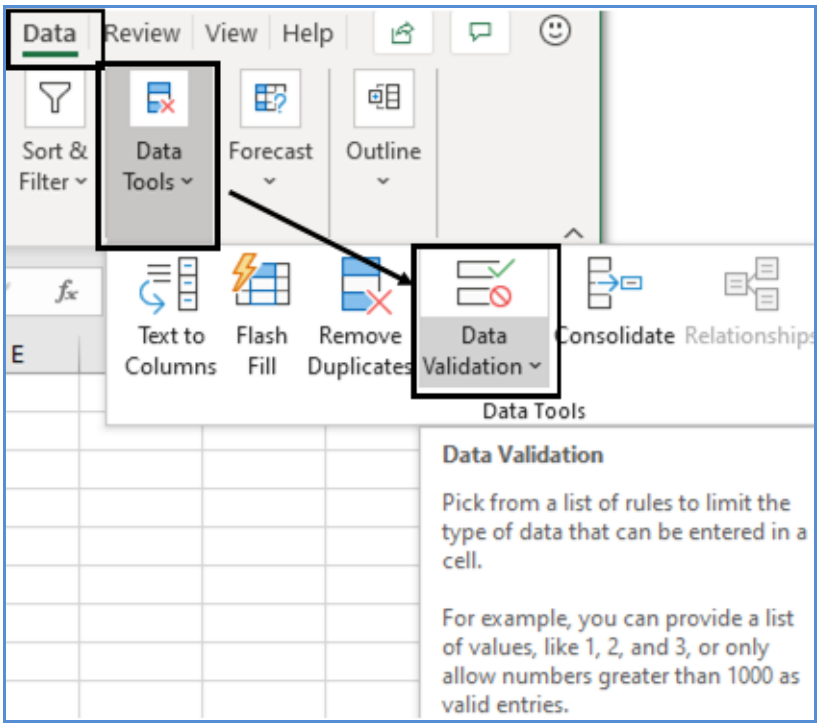


Figure 3.4.1 Data Validation

3.4.1 BASIC DATA VALIDATION

Data validation can be applied to cell to achieve the following objectives:

- Restrict data entry to values in a drop-down list
- Restrict data entry to a whole number within limits

- Restrict data entry to a decimal number within limits
- Restrict data entry to a date/ time within a time frame
- Restrict data entry to text of a specified length
- Use a formula to calculate what is allowed

Restrict data entry to values in a drop-down list

To make data entry easier in Excel, or to limit entries to certain items that we define "The List" choice is interesting because it doesn't just restrict invalid values, it also lets us add a handy drop-down list box that appears when anyone using the spreadsheet moves into that cell.

The person who is entering data can use the list to quickly insert an allowed value, without needing to type it in. We can also type values in by hand, but Excel assumes that if the value we enter does not match one of the entries in the list, our entry is invalid and will show error message.

Drop-down list of valid entries can be created out of range of cells elsewhere on the worksheet or maybe in another worksheet or workbook, it could be a named range also.

When a drop-down list is created for a cell, it displays an arrow next to that cell. To enter information in that cell, click the arrow, and then click the entry that we want.

- In the **Data Validation** dialog box, click the **Settings** tab.
- In the **Allow** box, select L**ist**.
- In the **Source** box, type the range of cells where your list resides e.g. D2:D6 as shown in Fig.

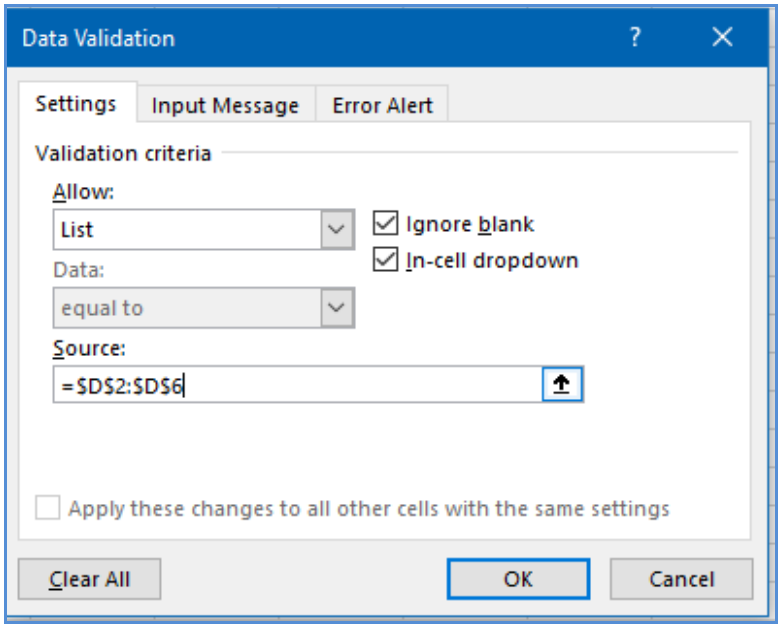


Figure 3.4.2 Data Validation through List

Restrict data entry to a whole number within limits

Eg force the person using the workbook to enter a whole number from 100 to 10000

- In the **Data Validation** dialog box, click the **Settings** tab.
- In the **Allow** box, select **Whole number**.
- In the **Data** box, select the type of restriction that we want Eg, to set upper and lower limits, select between.
- Enter the minimum, maximum, or specific value to allow. We can also enter a formula that returns a number value. For example we could only want values between 100 and 10000 as shown in Fig.

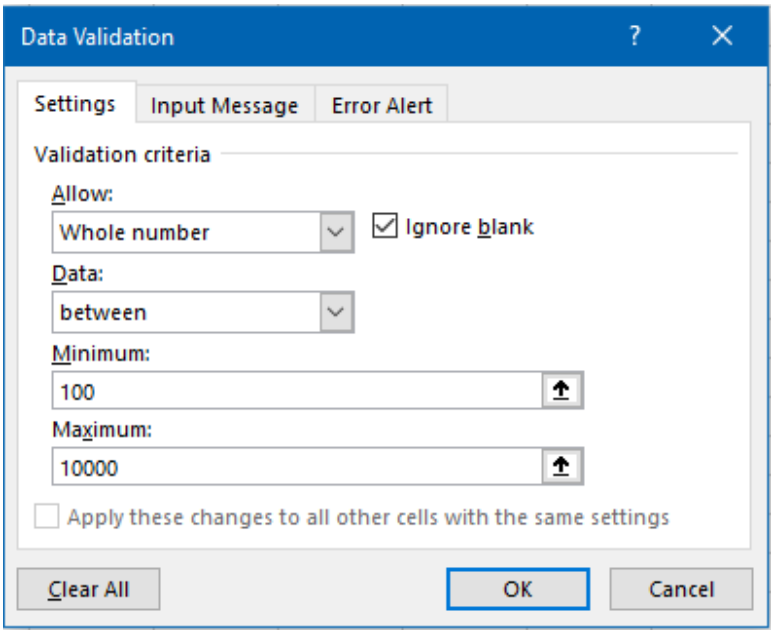


Figure 3.4.3 Data Validation Whole numbers

Restrict data entry to a decimal number within limits

- In the **Data Validation** dialog box, click the **Settings** tab.
- In the **Allow** box, select **Decimal**.
- In the **Data** box, select the type of **restriction** that we want. For example, to set upper and lower limits, select between.

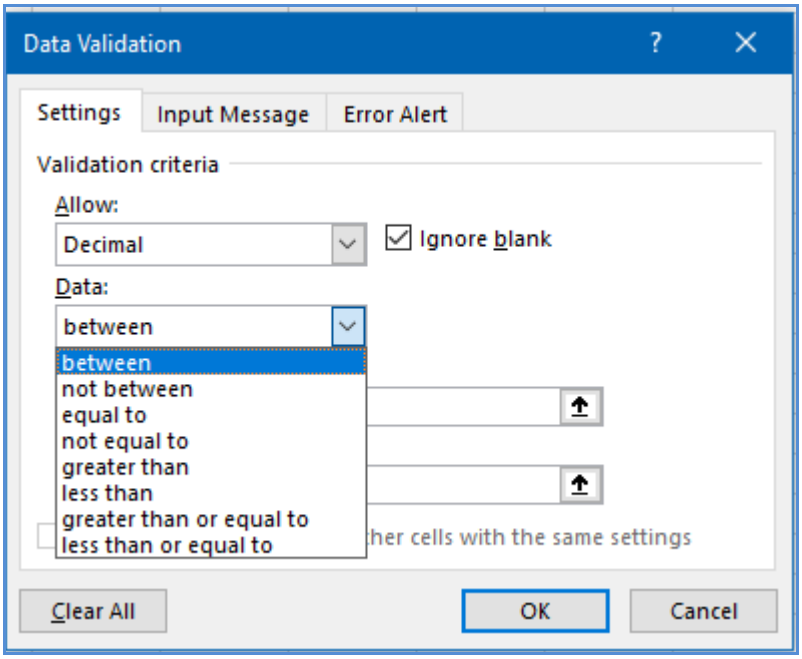


Figure 3.4.4 Data Validation decimal

Restrict data entry to a Date/Time within a time frame

For Example, we want to restrict only to dates in the current Financial year

- In the **Data Validation** dialog box, click the **Settings** tab.
- In the **Allow** box, select **Date /Time**.
- In the **Data** box, select the type of restriction that we want. For example, to allow dates after a certain day, select **greater than.**
- Enter the **start, end, or specific date/ time** to allow. We can also enter a formula that returns a date.

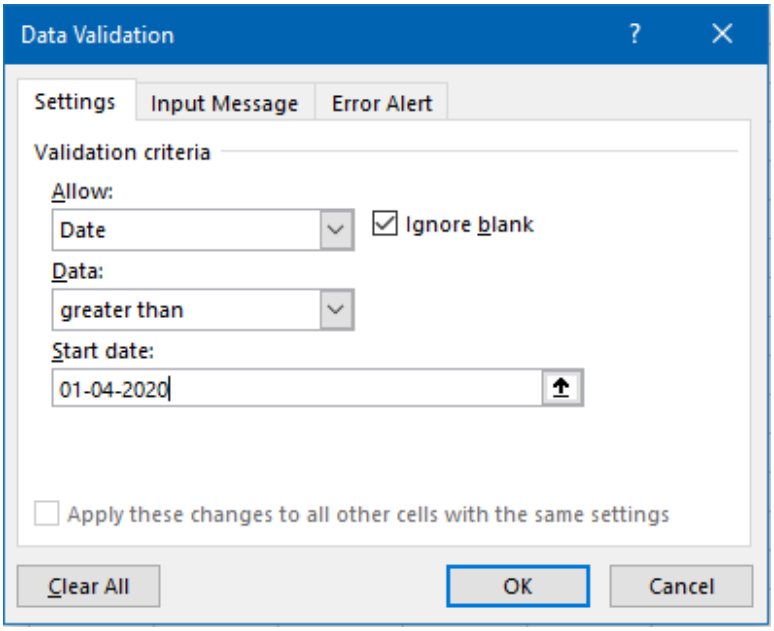


Figure 3.4.5 Data Validation Dates

Restrict data entry to text of a specified length

For Example, we want to enter PAN No which is 10 digit

- In the **Data Validation** dialog box, click the **Settings** tab.
- In the **Allow** box, select **Text** Length.
- In the **Data** box, select the type of restriction that we want. For example, in the above case **equal to**.
- In the **Length** box select *10 since PAN is 10 digit*

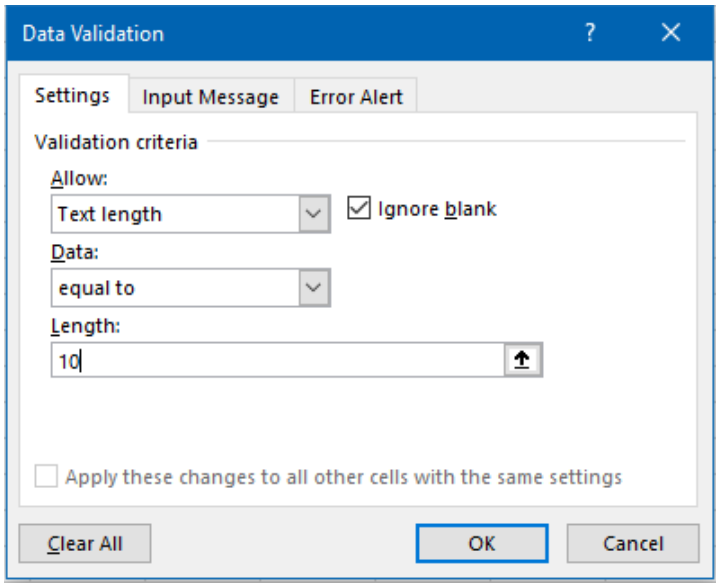


Figure 3.4.6 Data validation Text length

Customise validation through Formulas

- In the Data Validation dialog box, click the Settings tab.
- In the Allow box, select the Custom.
- In the Formula box, enter a formula that calculates a logical value (TRUE for valid or FALSE for invalid entries). For example: to allow only text in cell B2 we could take the formula =ISTEXT(B2)

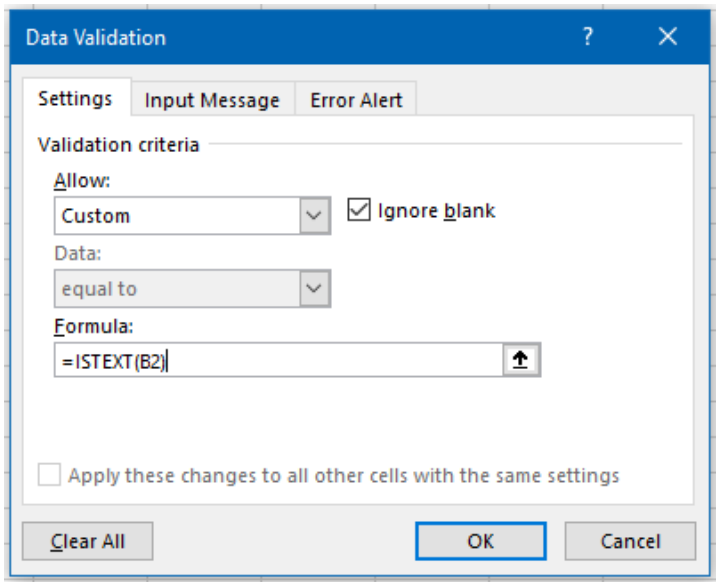


Figure 3.4.7 Data Validation based on formulas

3.4.2 CUSTOM INPUT MESSAGES & ERROR MESSAGES

If you want to display a message that explains to the user what data is allowed in each cell, open the Input Message tab, and do the following:

- Click the **Input Message** tab in the **Data Validation** dialog box.
- Check the box **Show input message** when cell is selected.
- Enter the title and text of your message into the corresponding fields.
- Click OK

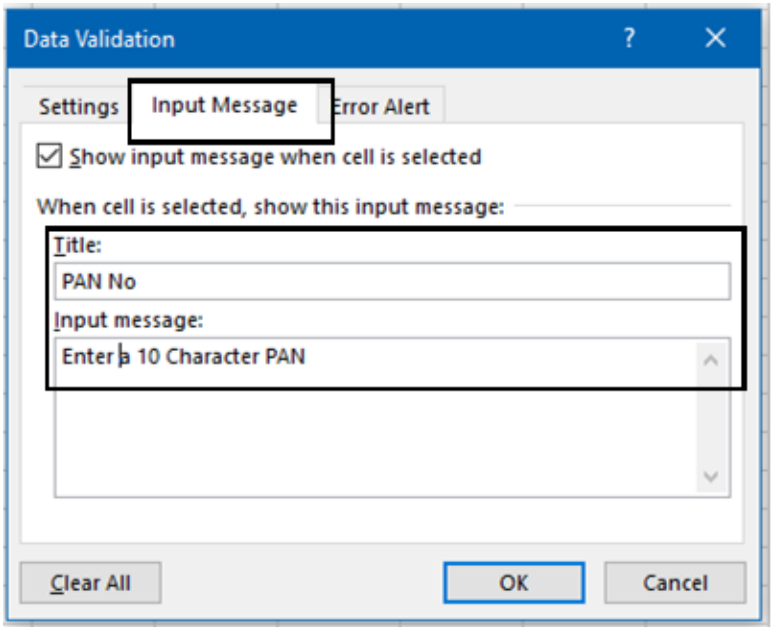


Figure 3.4.8 Input message

To configure a custom error message, go to the **Error Alert** tab

- Click the **Error Alert** tab in the **Data Validation** dialog box.
- Check the box **Show error alert after invalid data is entered.**
- You can select **Stop, Warning, Information** under **Style**: dropdown
- In the box under **Title**: type Invalid PAN:
- In the box under **Error message**: type “A valid PAN is 10 Characters”.
- Click OK.

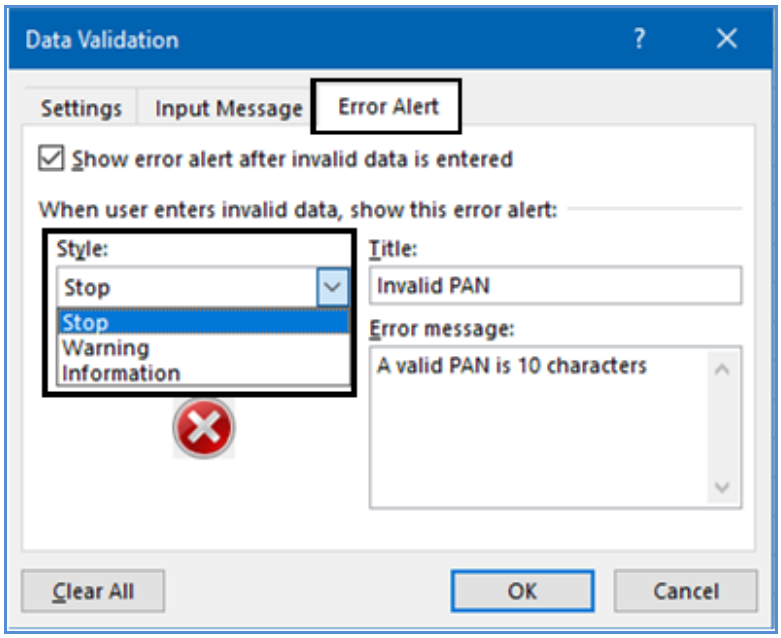


Figure 3.4.9 Error Alert

Food for thought: We need to know that Data validation has different options that have different effects on whether Excel tolerates invalid input. The choices are:

Stop. Excel displays the error message along with a Retry and Cancel button. The person using the workbook must click Cancel to reverse the change (which returns the cell to its last value) or Retry to put the cell back into edit mode and try to fix the problem. The Stop option is the only Style choice that completely prevents the person from using the workbook from entering invalid data.

Warning. In this case, the error message includes Yes and No buttons that let the person entering the data decide whether to go ahead with the input. Clicking Yes makes Excel accept the data entered into the cell, even if it breaks the validation rules.

Information. The error message comes with Cancel and OK buttons. Clicking OK enters the new (invalid) data in the cell; Cancel leaves the cell unchanged.

TIPS FOR WORKING WITH DATA VALIDATION

- If you are planning to share the workbook, share it only after having finished specifying data validation and protection settings. After sharing a workbook, you won't be able to change the validation settings unless you stop sharing. However, Excel will continue to validate the cells that you have designated while the workbook is being shared.
- You can apply data validation to cells that already have data entered in them. However, Excel does not automatically notify that the existing cells contain invalid data. In this scenario, you can highlight invalid data by instructing Excel to circle it on the worksheet. Once you have identified the invalid data, you can hide the circles again. If you correct an invalid entry, the circle disappears automatically.

- To quickly remove data validation for a cell, select it, and then open the Data Validation dialog box (Data tab, Data Tools group). On the Settings tab, click Clear All.
- To find the cells on the worksheet that have data validation, on the Home tab, in the Editing group, click Find & Select, and then click Data Validation. After you have found the cells that have data validation, you can change, copy, or remove validation settings.
- When creating a drop-down list, you can use the Define Name command (Formulas tab, Defined Names group) to define a name for the range that contains the list. After you create the list on another worksheet, you can hide the worksheet that contains the list and then protect the workbook so that users will not have access to the list.
- If data validation is not working, make sure that: Users are not copying or filling data -Data validation is designed to show messages and prevent invalid entries only when users type data directly in a cell. When data is copied or filled, the messages do not appear.
- Formulas are error free: Make sure that formulas in validated cells do not cause errors, such as #REF! or #DIV/0!. Excel ignores the data validation until you correct the error.
- Cells referenced in formulas are correct: If a referenced cell changes so that a formula in a validated cell calculates an invalid result, the validation message for the cell won't appear.

3.4.3 FIND CELLS THAT HAVE DATA VALIDATION

- You wish to **locate** the cells which contain **data validation**
- On the **Home** tab in the **Editing** group, click **Find & Select**, and then click **Data validation** to quickly find and select all cells that contain data Validation.

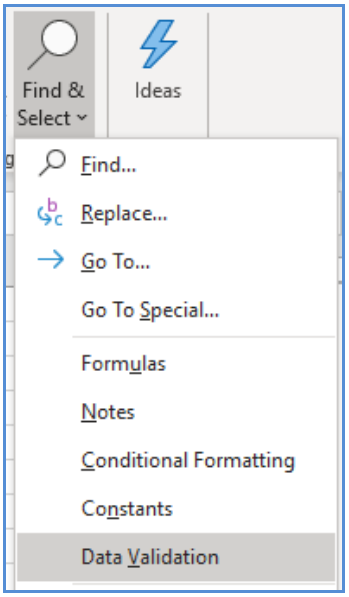


Figure 3.4.10 Find Data Validation

- Thus, we can find cells where data validation has been applied.

3.4.4 DISPLAY OR HIDE CIRCLE AROUND INVALID DATA

Data validation is designed to display messages and prevent invalid entries only when users type data directly in a cell. In the following cases, validation messages will not appear, and invalid data can be entered:

- When a formula in the cell calculates a result that is not valid.
- When a macro enters invalid data in the cell.

We can audit our worksheets to look for incorrect data that may cause inaccurate calculations or results. We can identify cells with data validation that contain invalid data by displaying a red circle around them so that we can easily find and correct any problems. Data validation begins policing a cell only after it has been applied. If we apply a validation rule and the cell already contains invalid information, Excel does not complain.

Circle invalid cells

- On the **Data** tab in **data tools** under **Data validation** go to **Circle Invalid Data**

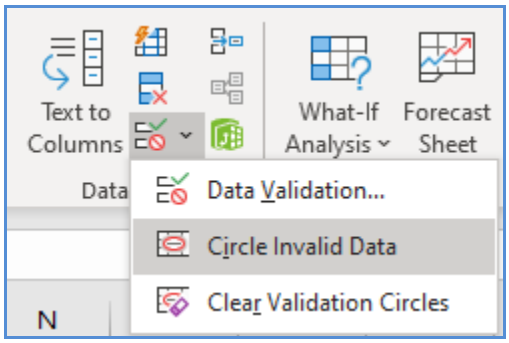


Figure 3.4.11 Circle Invalid data

- **Click** Circle Invalid Data
- Microsoft Office Excel displays a red circle around any cells that contain invalid data.

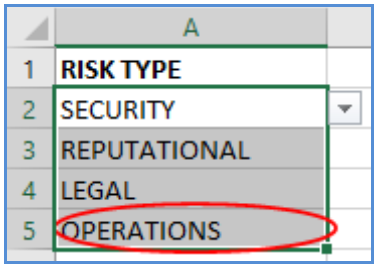


Figure 3.4.12 Invalid data

- All cells that do not meet their data validation criteria are circled, including values that were typed, copied, or filled in the cells, calculated by formulas, or entered by macros.

HIDE VALIDATION CIRCLES

We can do one of the following:

- To remove the circle from a single cell, enter valid data in the cell.
- To hide all circles, On the **Data** tab in **data tools** under **Data validation** go to **Clear Invalid Data**

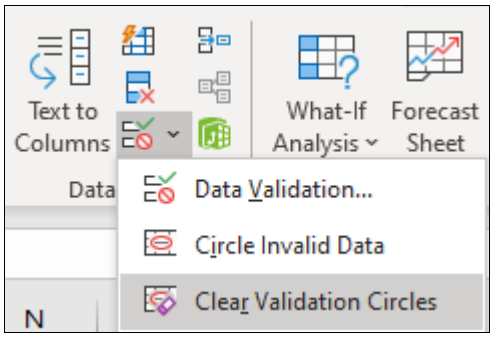


Figure 3.4.13 Hide Validation Circles

3.4.5 REMOVE DATA VALIDATION

To remove data Validation:

- On the **Data** menu, click **Validation**
- In the **Data Validation** dialog box, click the **Settings** tab, and then click **Clear All**.

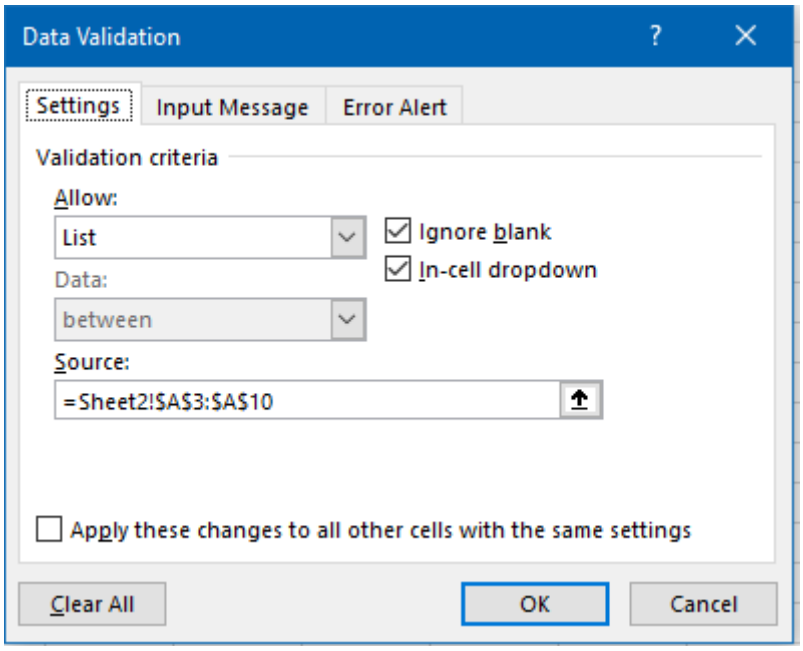


Figure 3.4.14 Remove validation

3.4.6 INDIRECT FUNCTION

The INDIRECT Function converts a text string into a cell reference, it means Excel INDIRECT can be used to indirectly reference cells, ranges, other sheets, or workbooks, that is why the name.

You can use this function to create a reference that will not change if row or columns are inserted in the worksheet. Or use it to create a reference from letters and numbers in other cells.

The INDIRECT function has two arguments: **INDIRECT (ref_text,a1)**

- **ref_text**: A cell reference or a text string (or both), that create a range reference to a cell, range of cells or named range.
- **a1**: (optional) Does the reference use A1 reference style?
  - If TRUE or omitted, ref_text is interpreted as an A1-style cell reference.
  - If FALSE, ref_text is treated as a R1C1 reference.

BASIC USE

You have number 500 in cell D1, and text D1 in cell A1. Now, put the formula =INDIRECT(A1) in any other cell and see what happens:

The INDIRECT function refers to the value in cell A1, which is D1.

The function is routed to cell D1 where it picks the value to return, which is number 500.

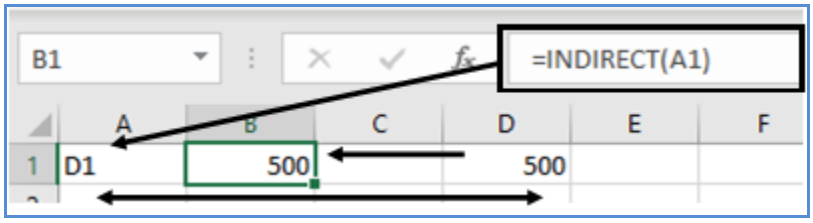


Figure 3.4.15 Indirect Function

INDIRECT FOR NAMED RANGES

You can refer to named ranges in an INDIRECT formula. As an example, the INDIRECT function is used to sum the selected named range.

In cells A1:D7, there are Customer data and B2:B7, C2:C7 & D2:D7 have been named as JAN, FEB & MAR.

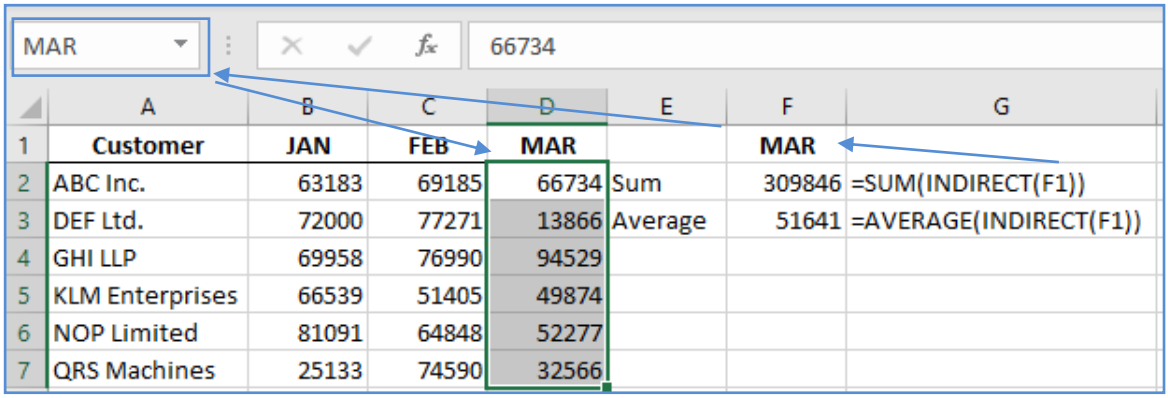
Now in cell F1, you type the name of one of the ranges, Eg. MAR

In cell F2, type the formula: =SUM(INDIRECT(F1))

Press the Enter key, and the formula returns the sum of numbers in the MAR range.

Change cell F1 to FEB, and the formula returns the sum of numbers in the FEB range.

Thus, INDIRECT indirectly refers to the **MAR** range.



	A	B	C	D	E	F	G
1	**Customer**	**JAN**	**FEB**	**MAR**		**MAR**	
2	ABC Inc.	63183	69185	66734	Sum	309846	=SUM(INDIRECT(F1))
3	DEF Ltd.	72000	77271	13866	Average	51641	=AVERAGE(INDIRECT(F1))
4	GHI LLP	69958	76990	94529			
5	KLM Enterprises	66539	51405	49874			
6	NOP Limited	81091	64848	52277			
7	QRS Machines	25133	74590	32566			

Figure 3.4.16 Indirect Named Ranges

3.4.7 DEPENDENT VALIDATION CREATION

We can limit the choices in an Excel data validation list, by using named ranges and the INDIRECT function, to create dependent data validation lists.

CASE STUDY: Your principal CA Chandiwala wants you to create two dropdown lists. The second list should be dependent on what is selected in the first one. He wants the first dropdown list with CA Partner-in-charge i.e. CA Chandiwala & CA Sonawala and independent List in 2nd cell, we want the dropdown list of Audit Managers from the respective lists.

Strategy:

We can use the INDIRECT function as the source of the second list.

1. On a blank sheet, set up a list of items for the first dropdown: Chandiwala & Sonawala Name the range Partner-in-Charge, as shown in Fig 3.4.17.

	A	B	C
1	**Partner-in charge**	**Chandiwala**	**Sonawala**
2	Chandiwala	Ram Mohan	Ramesh Chand
3	Sonawala	Shyam Mohan	Mool Chand
4		Chander Mohan	Prakash Chand

Figure 3.4.17 List

2. **Define** the first list.
3. In other columns, **set up a list of choices available for each Group** Eg. for **Chandiwala,** it could be Name of Audit managers as shown in Figure
4. Similarly, the second list **Sonawala** will also have Names of audit managers under him as shown in Figure.
5. Name these ranges as **Chandiwala** and **Sonawala**. The **range name** for this list must match the value in the original list.

6. To select the partner-in-charge from cell E2:E4, select **Data** > **Data Validation**. Change the **Allow** box to **List**; in the Source box, type **= Partner_in_Charge**

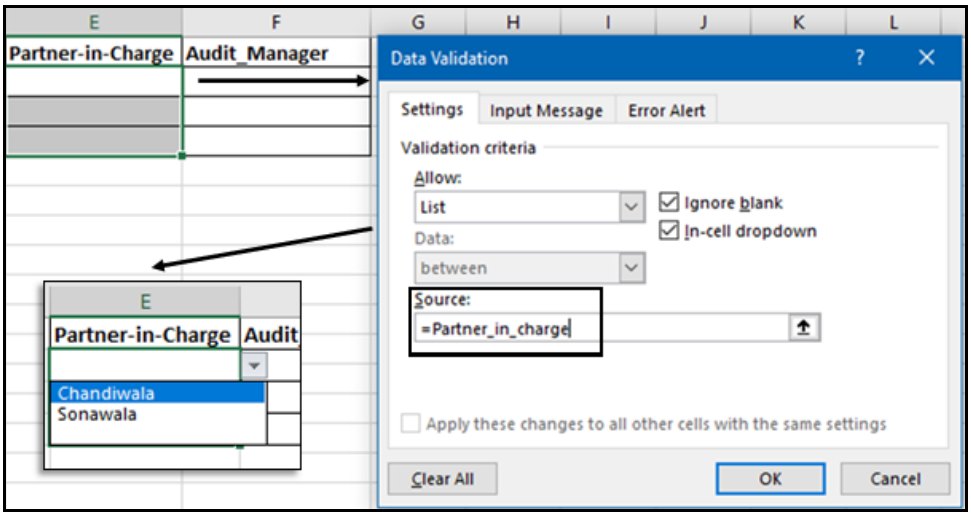


Figure 3.4.18 Data validation based on the list

7. Click **OK**. Cell D2 will have a dropdown list of Partner-In-charge.

8. To set up the second dropdown, select cell F2:F4 and then select **Data > Data Validation**. Change the Allow dropdown under Validation Criteria from Any Value to List. In the Source box, enter **=INDIRECT(E2),** as shown in Fig.

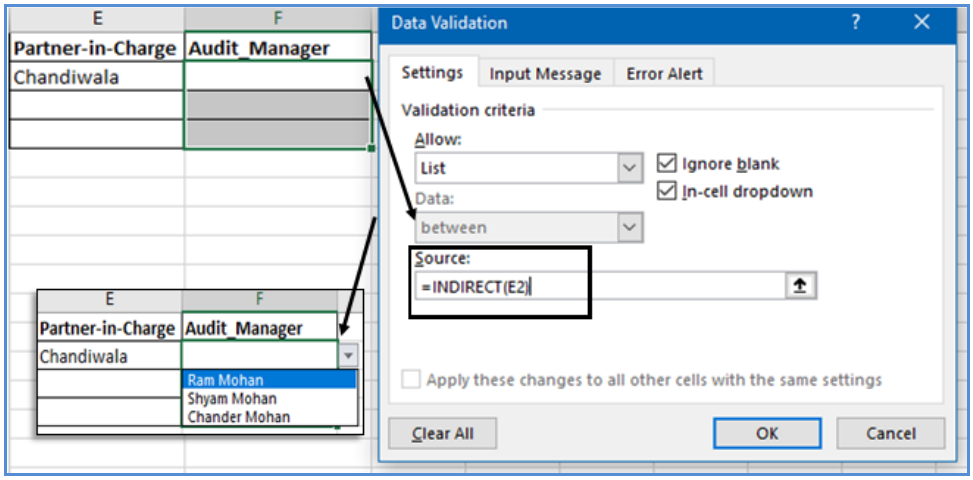


Figure 3.4.19 Data validation based on Indirect Formula

9. When we select a value in E2, the formula for the second dropdown list will automatically update, as shown in Fig. The INDIRECT function looks in E2 and hopes to find a formula there. When we select Chandiwala in E2, the validation formula becomes = Chandiwala. Because we cleverly set up a named range called Chandiwala, Excel can populate the list.

Gist: Using the **INDIRECT** function with **data validation** will allow us to set up a second validation list that is dependent on the choice in an earlier list.

Commands Learnt: **Data – Data Validation, INDIRECT ()**

3.4.8 ADVANCED USE OF DATA VALIDATION

Data validation is a wonderful tool to bring consistency of data and acts as an input check and is a kind of preventive control that can prevent any erroneous data to be entered. The below case study will make it clear how this versatile feature can be used for multiple types of checks.

CASE STUDY: Your principal CA Chandiwala wants you to prepare a risk register shown in Figure. At the same times to maintain data consistency he wants you to ensure the following:

	A	B	C	D	E	F	G	H	I	J	K
1	RISK REGISTER									RISK RESPONSE	
2	Risk ID	Risk Title	Risk Statement	Date Risk Identified	Risk Originator	Risk Category	Probability 1-5	Impact 1-5	Total Score	Risk Response Strategy	Risk Response Plan Description
3											

Figure 3.4.20 Risk register

Title	Data Validations
Risk ID	*Numeric*
Risk Title	*Text at least 20*
Date Risk Identified	*The day of entry*
Risk Category	*Only out of COMPLIANCE, INFOTECH, OPERATIONAL, REPUTATIONAL, SECURITY*
Probability	*Values 1 to 5*
Impact	*Values 1 to 5*
Risk Response Strategy	*Only out of ACCEPT, AVOID,MITIGATE,TRANSFER*

He asks you to find a way in which data should be restricted for entry as per above parameters You are a smart professional, is there a smart way to do it?

Strategy:

You can use various Data validations to achieve the data consistency.

To create these validations, you use the **Data Validation** Features under the **Data Tab**

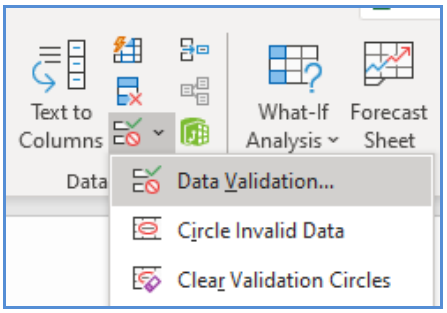


Figure 3.4.21 Create data validation

Risk ID: Numeric

1. You will select A3:A7
2. In the **Data Validation** dialog box, click the **Settings** tab. In the **Allow** box, select **Custom**.
3. In Formula Box write **=ISNUMBER(A3)** this will ensure that the data entered is numeric.

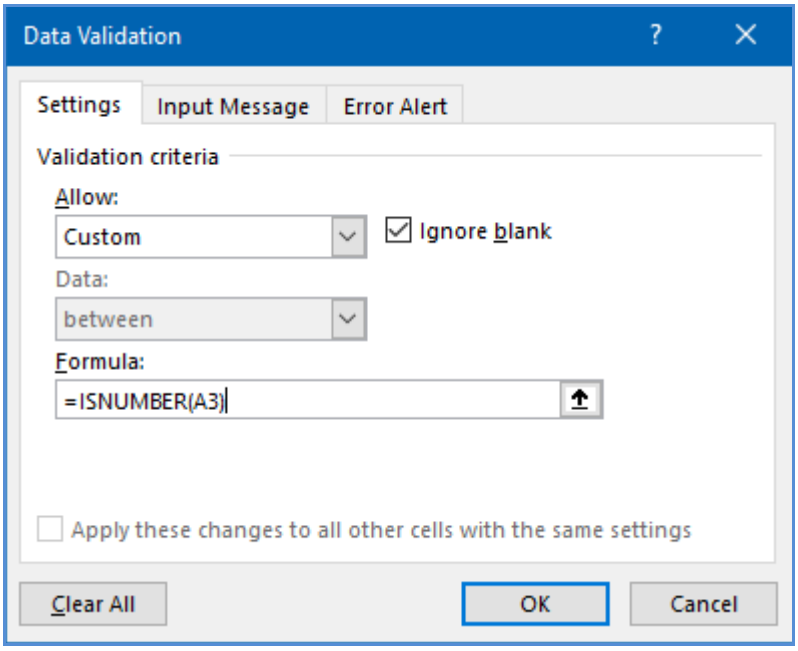


Figure 3.4.22 Data validation formula

4. You can also display an input message when the cell is clicked.

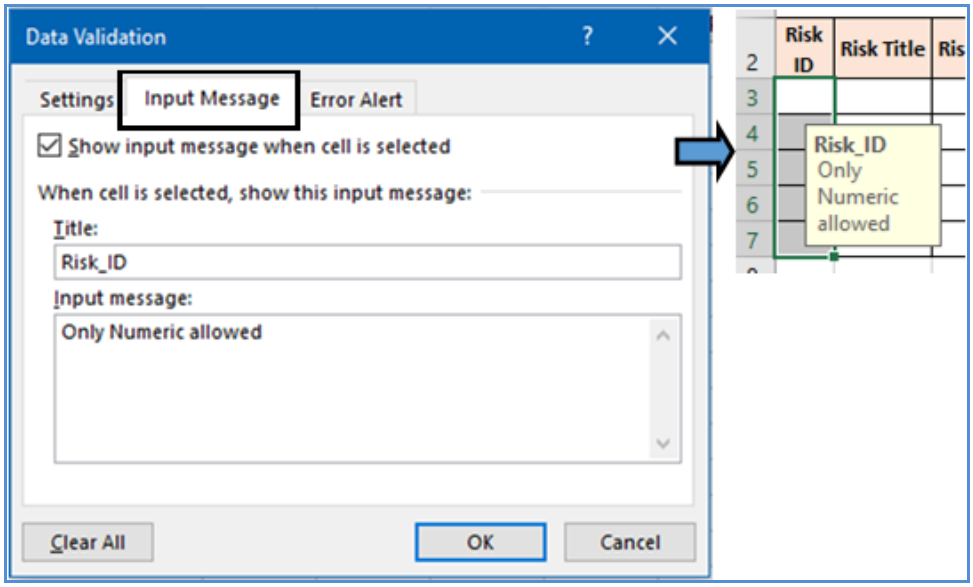


Figure 3.4.23 Input Message

5. You can also specify how Excel will respond when invalid data is entered.

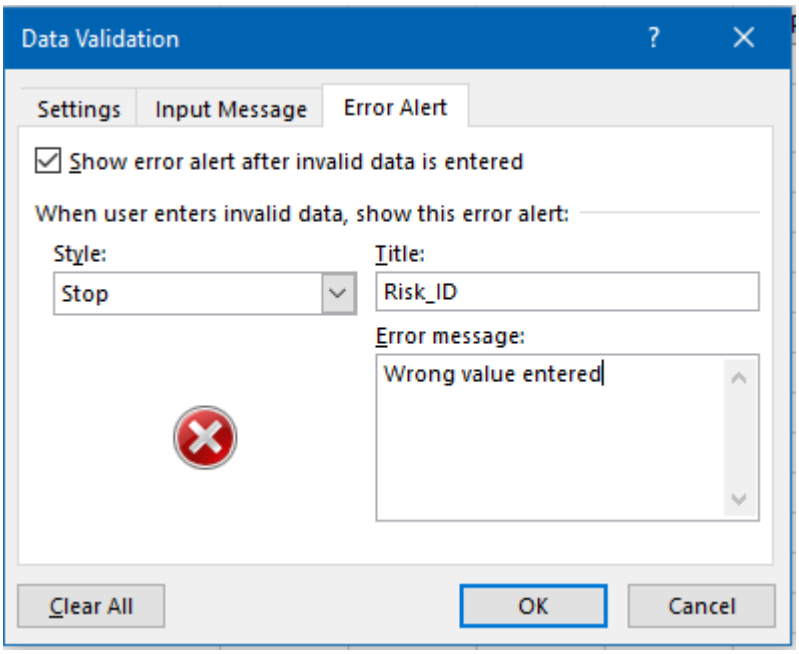


Figure 3.4.24 Error Alert

Risk Title: Text at least 20

6. You will select B3:B7
7. In the **Data Validation** dialog box, click the **Settings** tab. In the **Allow** box, select **Text**.
8. In Data select '**Greater than or equal to'**
9. In Minimum type **'20'** this will ensure that the data **Text at least 20 characters**

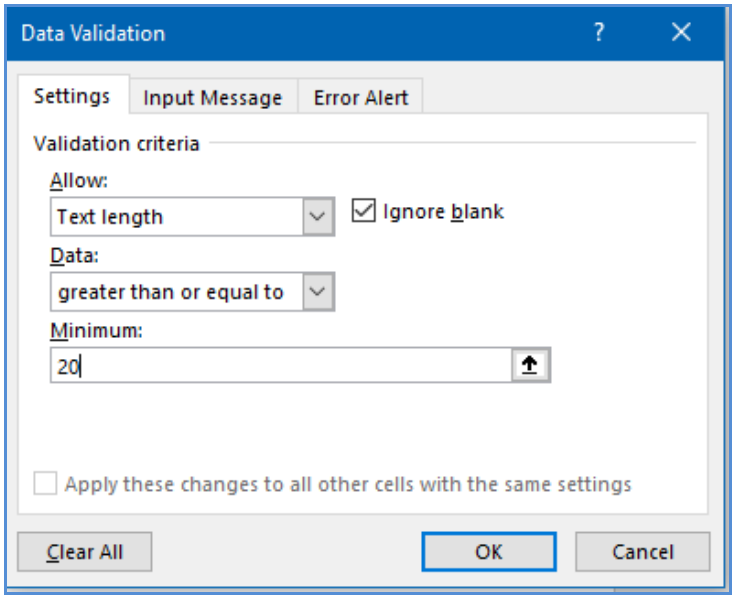


Figure 3.4.25 Data validation based on text length

Date Risk Identified: The day of entry

10. You will select D3:D7
11. In the **Data Validation** dialog box, click the **Settings** tab. In the **Allow** box, select **Date**.
12. In Data select '**equal to'**
13. In Date type **'=TODAY ()'** this will ensure that the data **will always be the date of entry**

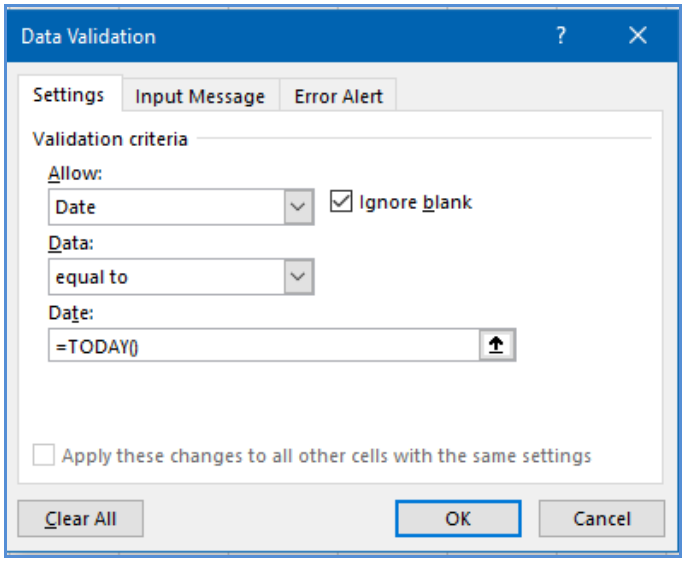


Figure 3.4.27 Data Validation based on Date

Risk Category: Only out of COMPLIANCE, INFOTECH, OPERATIONAL, REPUTATIONAL, SECURITY

Risk Response Strategy: Only out of ACCEPT, AVOID, MITIGATE, TRANSFER

14. To create a list of valid entries for the drop-down list, **type the entries** in a single column or row without blank cells.

15. We first need to sort the data in the order that we want it to appear in the drop-down list.

- Best approach is always **defining a name** for the list.
- The list could be in another worksheet or another workbook
- Different worksheet in the **same workbook** - Type the list on that worksheet, and then define a name for the list.
- Different worksheet in a **different workbook** -Type the list on that worksheet, and then define a name with an external reference to the list.

D	E
Risk Category	**Risk Response Strategy**
COMPLIANCE	ACCEPT
INFOTECH	AVOID
OPERATIONAL	MITIGATE
REPUTATIONAL	TRANSFER
SECURITY	

Figure 3.4.28 List for reference

16. **Select** the cell where you want the drop-down list.
17. On the **Data** Tab, click **Validation**, and then click the **Settings** tab.
18. In the **Allow** box, click **List**.
19. To specify the **location** of the list of valid entries, do one of the following:
    - If the list is in the current worksheet, enter a reference to our list in the **Source** box.
    - If the list is on a different worksheet in the same workbook or a different workbook, enter the name that we defined for our list in the Source box.
    - In both cases, make sure that the reference or name is preceded with an equal sign (=). For example, **enter =Risk_Category,** you can also use Key F3 to get **paste name** and select
20. Make sure that the **In-cell drop-down** check box is selected.
21. To specify whether the cell can be left blank, select, or clear the **Ignore blank** check box.

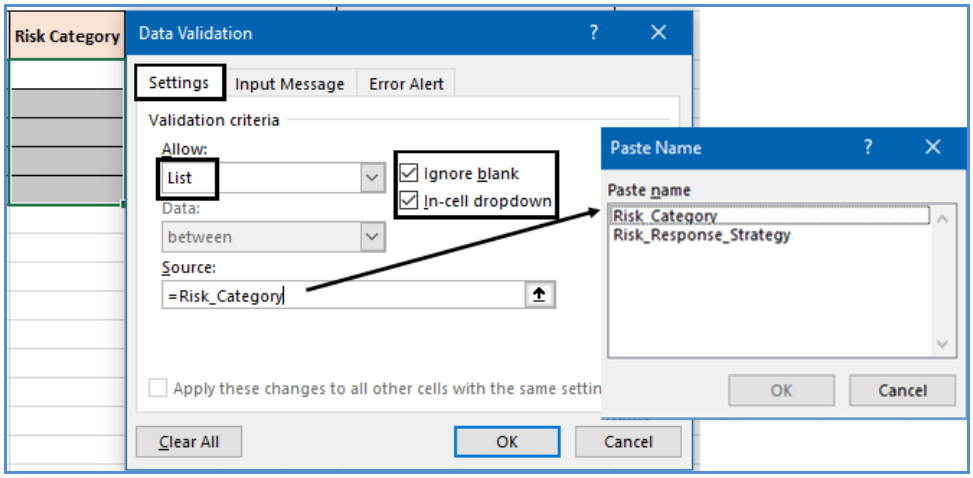


Figure 3.4.29 List based on Named Range

22. Now when you enter a dropdown will appear as under.

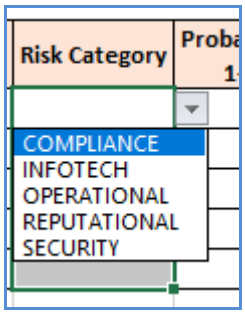


Figure 3.4.30 Dropdown based on Valid data

Probability: Values 1 to 5

Impact: Values 1 to 5

23. You will select G3:G7 similarly H3:H7
24. In the **Data Validation** dialog box, click the **Settings** tab. In the **Allow** box, select **Whole number**.
25. In Data select **''between'**
26. In Minimum type **'1'** and Maximum type '**5**' this will ensure that the data **will be values from 1 to 5**

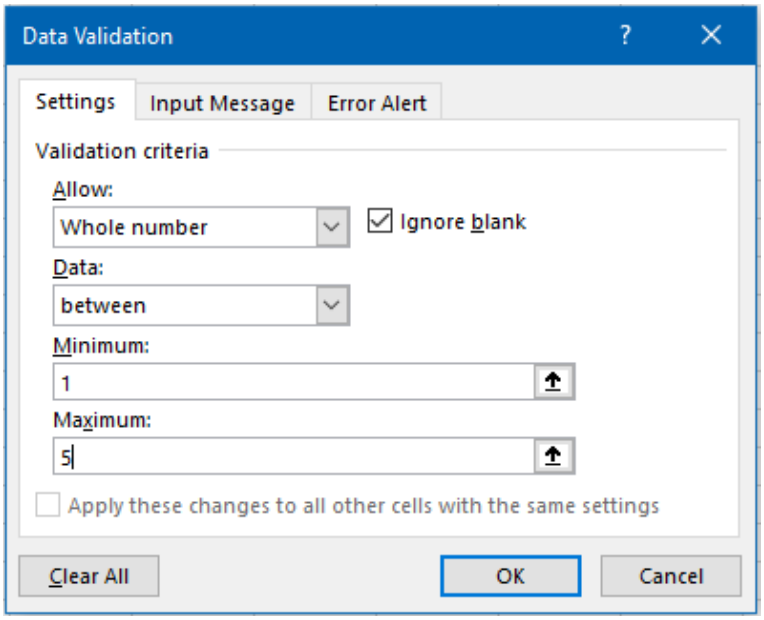


Figure 3.4.31 data validation based on Number

Gist: We have created Different data validation for the risk register.

Commands learnt: Data Validation

Food for thought: To delete a drop-down list, select the cell with the list. Click the Data menu, and then click Validation. In the Data Validation dialog box, click the Settings tab, and then click Clear All.

3.5 SUMMARY

Workbooks are meant to be shared but Consistency of Formulas and formatting as well as Data must be maintained anybody using our workbook should not accidentally or intentionally messes with the formatting & data.

In this chapter, we learned how to we can name our ranges to make formulas more meaningful and any user can understand the formulas. We have further learnt to insert Tables so that our data is consistent in formatting, rows, and columns added automatically thus making tables dynamic.

We also understood how you can apply the cell validations so that only values meeting our predefined criteria are entered or accepted by the cells. We have further learnt to copy, remove cell validations. Our need for creating dependent lists has been fulfilled using data validation along with function INDIRECT. Finding invalid data and finding cells with data validation has also been addressed in this chapter.

REFERENCES

[1] Michael Alexander, Richard Kusleika, John Walkenbach, 'Excel 2019 Bible', Wiley,2018

[2] Greg Harvey, 'Excel 2019 All-in-One for Dummies, John Wiley & Sons,2018

[3] www.chandoo.org

CHAPTER

4

ORGANISE DATA USING EXCEL

LEARNING OBJECTIVES

- Understanding Sorting with Excel
- Understanding the use and benefits of Filter
- Understand Advanced filtering options
- Understanding how to Group Data
- Understand the use of Subtotal in Excel

4.1 INTRODUCTION

Imagine having a long list of names or amounts or dates —or anything for that matter—and putting it in order. Putting a list in order, accomplishes several goals, including making your list easier to read and use, making it appear more orderly, and giving the content a perceived priority. Now that priority can of course change, as the same list can be put in order—or sorted—by any field within the list. Sorting gives some sense to data.

Often you have to extract some data based on certain criteria. You will be using features like Auto filter or even Advanced filter in Excel to extract information matching your requirement. Excel has the capability to Filter based on text, numbers, or even dates making Data analysis easy.

A lot of complex and detailed information becomes difficult to read and analyse. Excel offers you the facility to outline where you can group data manually or automatically to make it understandable. It even offers you a one-click solution to create subtotals.

Pro Tip:

Excel can work intelligently, you need to follow some rules to keep the data in a format which makes powerful data commands possible.

- Use only a single row of headings above the data. If you need to have a two-row heading, set it up as a single cell with two lines in the row.
- Never leave one heading cell blank. You might do this if you add a temporary column. If you forget to add a heading before you sort, this will affect the intelligence of Excel, and it will sort the headings down into the data.
- There should be no entirely blank rows or blank columns in the middle of Your data. You may have an occasional blank cell but never have entirely blank columns.
- If your heading row is not in row 1, be sure to have a blank row between your headings and any other filled cells.

- Formatting the heading cells in bold will help the Excel intelligenty understand that these are headings.

4.2 SORT

Excel allows us to sort on various criteria. You can sort data by:

- **text** (A to Z or Z to A),
- **numbers** (smallest to largest or largest to smallest),
- **dates** and **times** (oldest to newest and newest to oldest)
- **custom list** (such as Large, Medium, and Small)
- **cell color**,
- **font color**, or
- **cell icon**.

Excel further gives us the option of

- **Quick sorting** -you can quickly sort your data by using the A-Z  and Z-A Sort buttons on the Ribbon's Data tab.
- **Sort Dialog Box –** You can also sort on various criteria through Sort Dialog Box  on the Data tab.

4.2.1 SORTING-CUSTOM LIST

CASE STUDY*: *In an audit, your principal CA Chandiwala gives you a pending Invoice list which he asks you Sort according to Financial Year i.e. April to March, normally options available for sorting in Excel are either A to Z, Z to A or Custom list, In custom list also the options are according to Calendar Year i.e. January to December. Your challenge is to sort according to the financial year?

	A	B	C	D	E
1	CUST_NO	INVOICE_NO	INV_DATE	Due Month	NET SALES
2	10204	1001796	12-11-2019	December	2,00,965
3	10302	1001300	13-08-2019	September	1,04,106
4	10400	1001780	09-11-2019	December	70,502
5	10500	1002100	17-12-2019	January	49,418
6	10801	1001556	01-10-2019	November	44,146
7	10900	1001572	01-10-2019	November	63,913
8	11100	1001348	17-08-2019	September	5,271
9	11207	1001156	13-07-2019	August	4,15,107
10	11300	1001492	21-09-2019	October	20,97,938
11	11301	1000724	16-04-2019	May	2,28,124
12	11400	1002348	29-01-2020	February	26,356
13	11600	1000692	09-04-2019	May	1,93,325

Figure 4.2.1 Pending Invoice list

Strategy:

You can Sort this list Using the Custom sorting option as follows:

1. Choose Sort by dropdown- Due Month, Sort on- values, Since, you want to do a sort on Due month the sort can be in alphabetic order ie A to Z but that will lead to a funny situation where after April, August will come & then December and so on if you look at dropdown you also see an option Custom Lists From January to December,
2. But you are not interested in a calendar year you want the list sorted as per **Financial year ie April to March**. For the purpose, you can also add/create custom lists.
3. Go to the **Blank section** of Your excel list & create the desired custom list as shown in Fig4.2.2 and **Select** the range of cells.

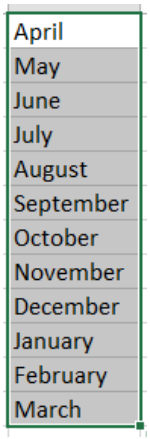


Figure 4.2.2 Custom list

4. Choose **File >Options >Advanced**. In the general category, Click the Edit Custom Lists Button as shown in Fig 4.2.3.

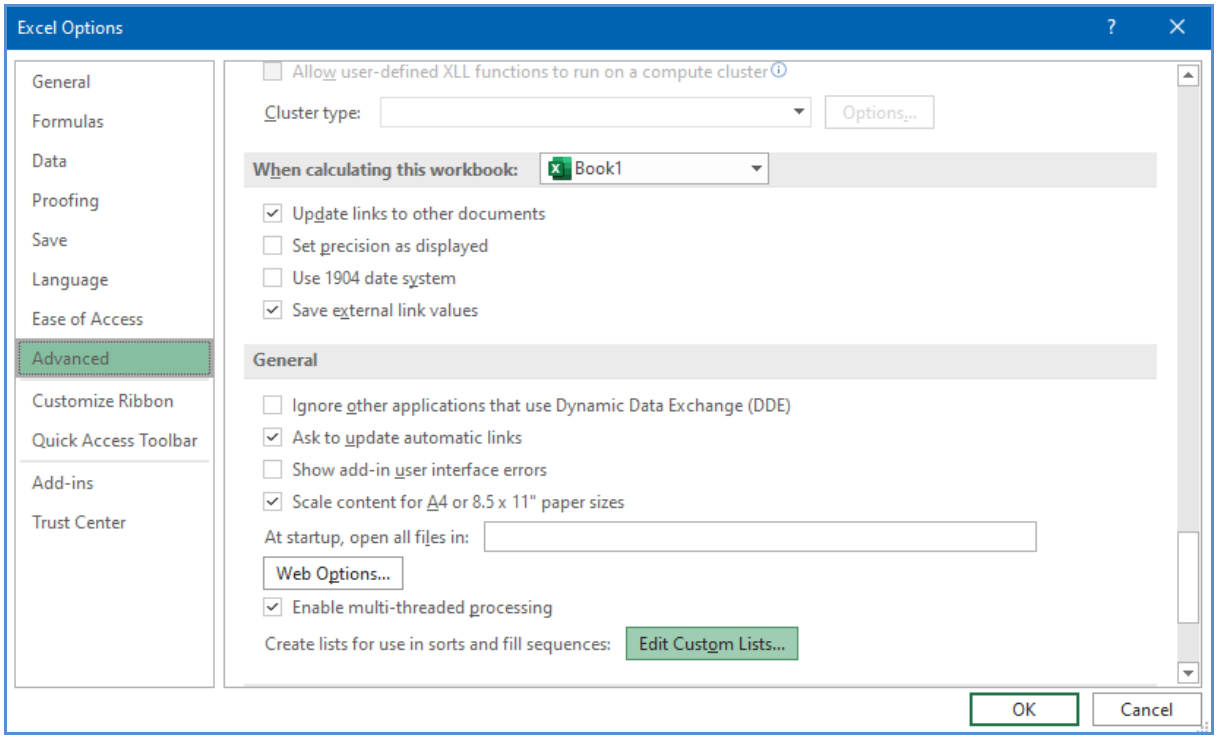


Figure 4.2.3 Edit Custom List

5. In the Custom List Dialog Box, Your range of values from the previous step is showing in the import text box as shown in Fig 4.2.4 Click the import button. Click OK

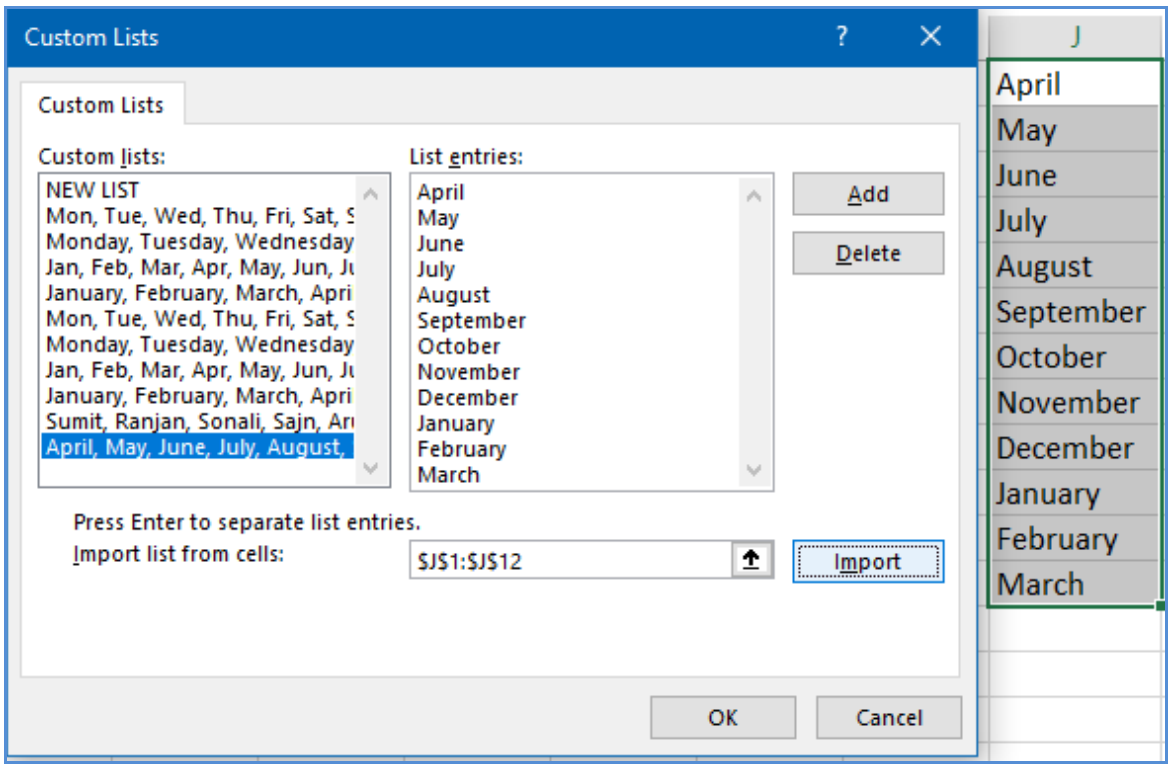


Figure 4.2.4 Custom list import

6. The custom list is added to the **custom list box** & you can now select this list to sort Your data as per Financial year as shown in Fig 4.2.5.

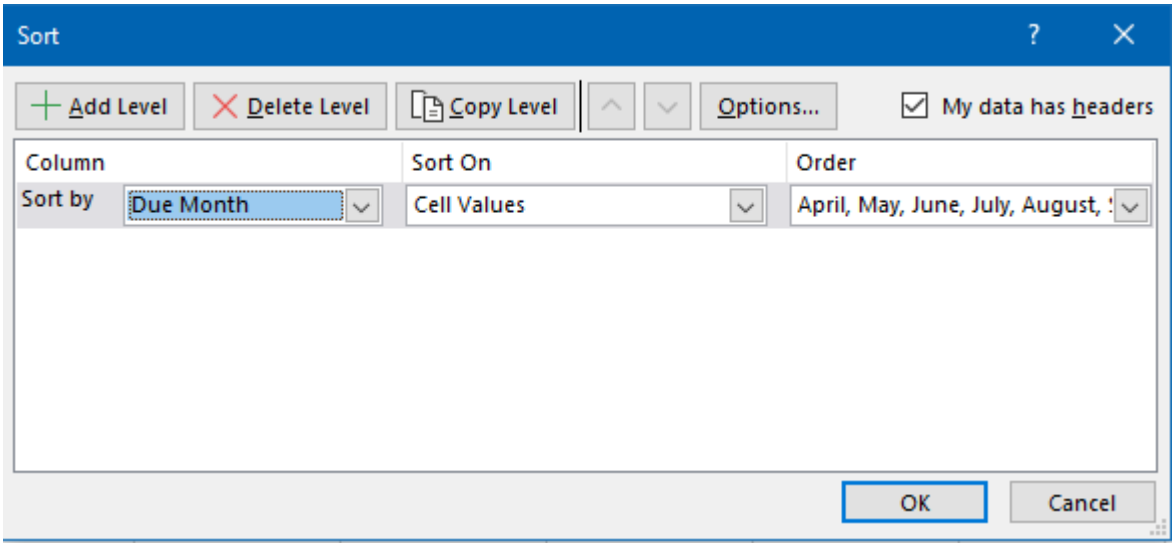


Figure 4.2.5 sort custom list

7. The resultant Invoice list sorted with due months in order of Financial Year as shown in Fig 4.2.6

	A	B	C	D	E
1	CUST_NO	INVOICE_NO	INV_DATE	Due Month	NET SALES
2	11704	1000580	06-03-2019	April	4,23,673
3	11704	1000628	13-03-2019	April	4,94,175
4	11805	1000596	13-03-2019	April	1,23,728
5	11805	1000660	27-03-2019	April	25,038
6	11806	1000548	02-03-2020	April	23,062
7	11806	1000612	13-03-2019	April	26,35,600
8	92326	1000564	05-03-2019	April	23,062
9	92326	1000644	20-03-2019	April	1,13,990
10	11301	1000724	16-04-2019	May	2,28,124
11	11600	1000692	09-04-2019	May	1,93,325
12	11704	1000756	23-04-2019	May	5,14,245
13	92326	1000676	05-04-2019	May	16,04,422

Figure 4.2.6 sorted list

Gist: You have analysed the outstanding Invoice data by sorting it according to the custom list.

Commands learnt: **Data > Sort**, Using & adding Custom Lists

4.2.2 COLUMN WISE SORTING

Excel can not only sort vertically i.e. top to bottom but also horizontally i.e. left to right

CASE STUDY: In an audit, your principal CA Chandiwala gives you a Customer wise, month-wise list and he wants to do a Monthly Trend Analysis, but the list has the monthly columns arranged alphabetically. Your principal tells you to cut and paste columns and arrange them in the Financial Year order. You are a smart auditor, is there a smart way to do it?

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	CUSTOMER	April	August	December	February	January	July	June	March	May	November	October	September
2	Earth Limited	-	13,354	1,68,654	1,35,192	69,899	1,38,801	3,10,176	19,19,390	1,05,908	26,020	1,33,971	19,12,887
3	Galaxy Enterprises	11,621	64,806	62,105	74,007	12,52,830	14,74,406	1,15,379	19,10,500	1,29,943	25,798	1,48,545	39,391
4	Mars & Pluto	-	10,352	8,192	-	23,567	-	88,306	30,397	59,531	22,952	23,840	13,844
5	Milky Way	28,667	30,602	76,210	61,621	34,742	7,472	12,72,140	2,06,687	41,643	11,050	88,276	1,16,321
6	Moon Light Enterprises	11,416	44,173	1,33,077	88,429	2,23,862	39,828	3,82,966	2,91,388	14,28,851	16,736	23,75,056	1,48,020
7	Neptune Pvt. Ltd	-	-	21,67,093	30,830	4,08,014	-	13,14,772	2,87,041	57,197	82,168	27,04,895	1,10,946
8	Sun Moon & Star	23,589	15,474	13,844	10,903	15,749	-	48,818	30,921	18,77,938	-	29,838	-
9	Sun Moon & Stars	8,954	48,097	56,101	13,166	46,725	-	86,420	41,288	14,17,918	13,950	57,786	13,107
10	Sun Way	45,460	15,050	1,65,694	84,338	1,91,423	30,586	4,19,054	2,14,397	1,68,519	87,978	2,09,329	1,45,727
11	Venus Enterprises	17,350	44,993	47,247	12,176	83,210	-	16,38,089	1,65,965	38,705	54,835	27,977	40,580
12	Grand Total	1,47,057	2,86,901	28,98,217	5,10,662	23,50,021	16,91,093	56,76,120	50,97,974	53,26,153	3,41,487	57,99,513	25,40,823

Figure 4.2.7 Monthwise list

Strategy:

You can Sort this list Using Column wise sorting option as follows:

1. Select the data B1:M12
2. Choose Sort from Data Tab.

3. You will options as shown in figure 4.2.8 and if you click

4. Select option of sorting **left to right**

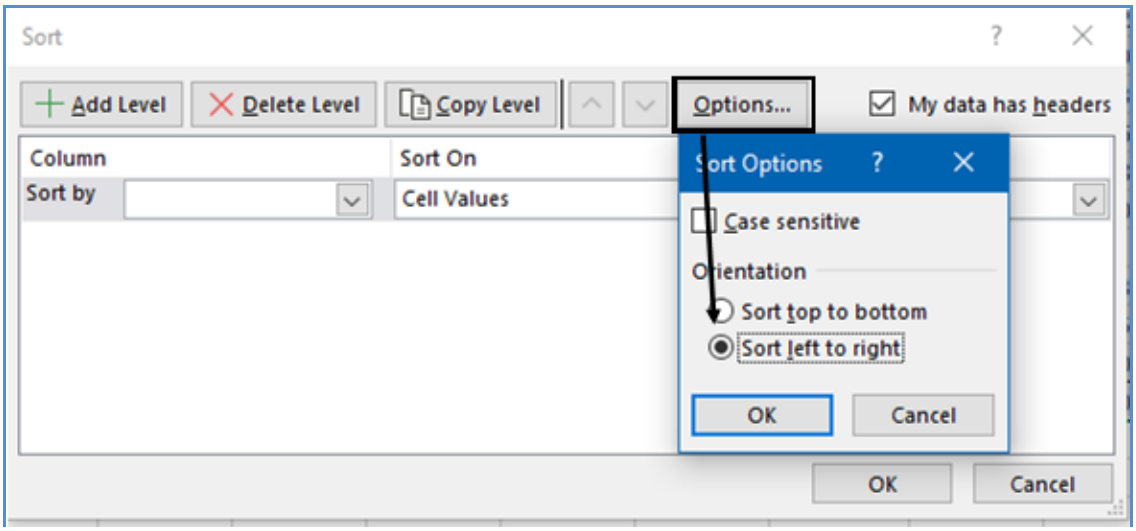


Figure 4.2.8 Sort Options

5. Sort by Row 1 and custom sort on April to March, the list you had created in the previous exercise

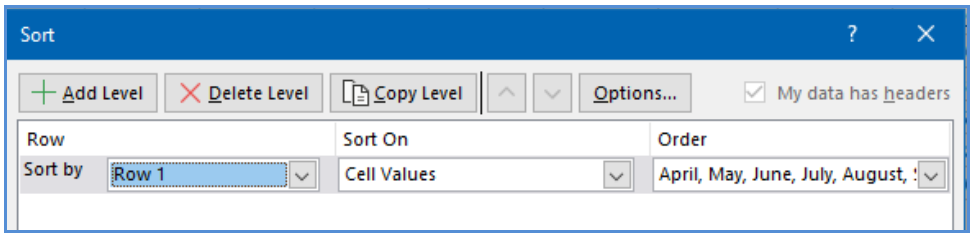


Figure 4.2.9 Custom Sort

6. Your Data is sorted Column wise as shown in Figure 4.2.10.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	CUSTOMER	April	May	June	July	August	September	October	November	December	January	February	March
2	Earth Limited	-	1,05,908	3,10,176	1,38,801	13,354	19,12,887	1,33,971	26,020	1,68,654	69,899	1,35,192	19,19,390
3	Galaxy Enterprises	11,621	1,29,943	1,15,379	14,74,406	64,806	39,391	1,48,545	25,798	62,105	12,52,830	74,007	19,10,500
4	Mars & Pluto	-	59,531	88,306	-	10,352	13,844	23,840	22,952	8,192	23,567	-	30,397
5	Milky Way	28,667	41,643	12,72,140	7,472	30,602	1,16,321	88,276	11,050	76,210	34,742	61,621	2,06,687
6	Moon Light Enterprises	11,416	14,28,851	3,82,966	39,828	44,173	1,48,020	23,75,056	16,736	1,33,077	2,23,862	88,429	2,91,388
7	Neptune Pvt. Ltd	-	57,197	13,14,772	-	-	1,10,946	27,04,895	82,168	21,67,093	4,08,014	30,830	2,87,041
8	Sun Moon & Star	23,589	18,77,938	48,818	-	15,474	-	29,838	-	13,844	15,749	10,903	30,921
9	Sun Moon & Stars	8,954	14,17,918	86,420	-	48,097	13,107	57,786	13,950	56,101	46,725	13,166	41,288
10	Sun Way	45,460	1,68,519	4,19,054	30,586	15,050	1,45,727	2,09,329	87,978	1,65,694	1,91,423	84,338	2,14,397
11	Venus Enterprises	17,350	38,705	16,38,089	-	44,993	40,580	27,977	54,835	47,247	83,210	12,176	1,65,965
12	Grand Total	1,47,057	53,26,153	56,76,120	16,91,093	2,86,901	25,40,823	57,99,513	3,41,487	28,98,217	23,50,021	5,10,662	50,97,974

Figure 4.2.10 Sorted Columns

Gist: You have analysed the data by sorting it according to the Columns.

Commands learnt: **Data > Sort**, Option: Left to right

4.3 FILTER

You often want to extract figures more than a certain limit from Your list, you again want to know sales made to Galaxy Ltd in a list of sales. Fortunately, Excel includes an easy-to-use Filter to show just what you want to see and hide the rest. Filtering doesn't change Your data in any way. As soon as you remove the filter, all Your data reappears, exactly the same as it was before.

Filtering is a way that you can use Excel to quickly extract certain data from Your spreadsheet. Unlike sorting, filtering doesn't just reorder the list. It hides the rows or columns containing data that do not meet the filter criteria you define. Excel has an AutoFilter feature that makes it very easy to extract data from Your spreadsheet. After filtering data, you can copy, find, edit, format, chart, and print the subset of filtered data without rearranging or moving it.

If you're using an Excel Table, it has built-in filtering features. If the data is not in an Excel Table, you can bring it up from **Filter** under **Sort & Filter** in **Data Tab.**

You can also filter by more than one column. Filters are additive, which means that each additional filter is based on the current filter and further reduces the subset of data.

4.3.1 ADDING A FILTER

The various methods of adding a filter are:

Method 1

In Ribbon go to **Data Tab > Sort & Filter Group > Filter** as shown in Figure 4.3.1

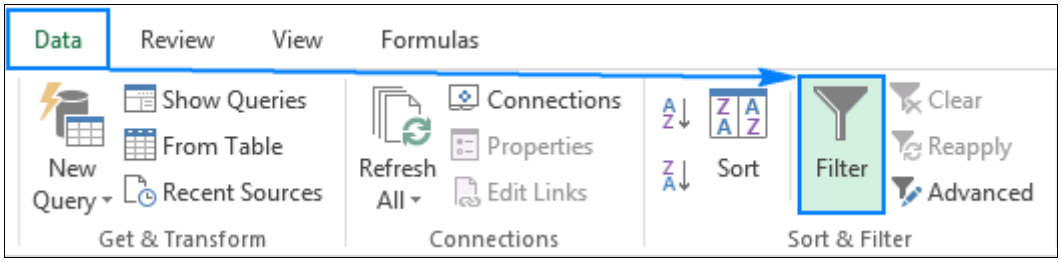


Figure 4.3.1 Filter in Data Tab

Method 2

In Ribbon go to **Home Tab > Editing Group> Sort & Filter > Filter** as shown in Figure 4.3.2

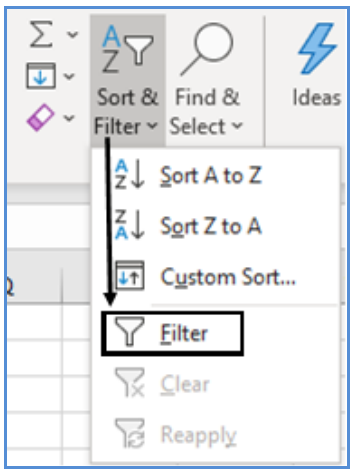


Figure 4.3.2 Filter on Home Tab

Method 3

Use a shortcut **ALT+A+T** or **CTRL+SHIFT+L**

When Filter is activated you see a drop-down arrow on the column heading, it indicates that filtering is added, but not applied yet. When you hover over the arrow, a screen tip displays (Showing All).

To filter data in Excel:

- Click the **drop-down arrow** for the column you want to filter.
- **Uncheck the Select All** box to quickly deselect all data.
- **Check the boxes** next to the data you want to display,
- Click **OK**.

Various types of filters are there:

- **Text filters** if the column you selected contains text
- **Date filters** if the column you selected contains dates
- **Number filters** if the column you selected contains numbers
- **Font color** if the column you selected contains font with color
- **Cell icon** if the column you selected contains cell icons
- **Advanced filter**

Excel also enables us to perform more intricate types of filtering. Two particularly useful types are the Top 10 filter and custom filtering.

- **Finding the Top (or Bottom) 10 in a column**
- **Using custom filters**

When filtering under Text or Number or date filters you find the option of **Custom Filtering**.

- In custom Filtering dialog box where you could enter two logical filtering conditions **"and", "or"**
- To filter the table column or selection so that **both criteria** must be true, select **And**.
- To filter the table column or selection so that **either or both criteria** can be true, select **Or.**

You can now enter two filtering requirements for the column of data which could be anything like

- equals
- does not equal
- is greater than
- is greater than or equal to
- is less than
- is less than or equal to
- begins with
- does not begin with
- ends with
- does not end with
- contains
- does not contain

value or text or date could also contain wildcard like ? or *.

You could even have **DYNAMIC FILTER** like in case of date filters you could have a filter for "**Today**", "**Yesterday**" or " **next month**", which would give a different filtered value every day.

Wherever you have a dynamic filter it is important to note that **AutoFilters doesn't refresh automatically**. To see the updated filtered results, you can **reapply** the filter from the sort & Filter group in Data Tab.

Filtering could even be done on the type of format; you can **Filter by Cell Color, Filter by Font Color, or Filter by Cell Icon.**

4.3.2 DATE AND TEXT FILTERS

Date Filters

Excel Date Filters have many choices that let you filter records for a certain time-period pretty fast.

By default, Excel Filter groups the dates in a given column by a **hierarchy of years, months, and days**. You can expand or collapse different levels by clicking the plus or minus signs next to a given group.

Date Filters also allow you to display or hide data for a particular **day, week, month, quarter, year**, before or after a specified date, or between two dates. Excel filter by date works in a single click. As an example, to filter rows containing records for the current quarter, you simply point to Date Filters and click This Quarter.

There is also an option of **Custom Autofilter** dialog box to select the Equals, Before, After, Between operator or Custom Filter.

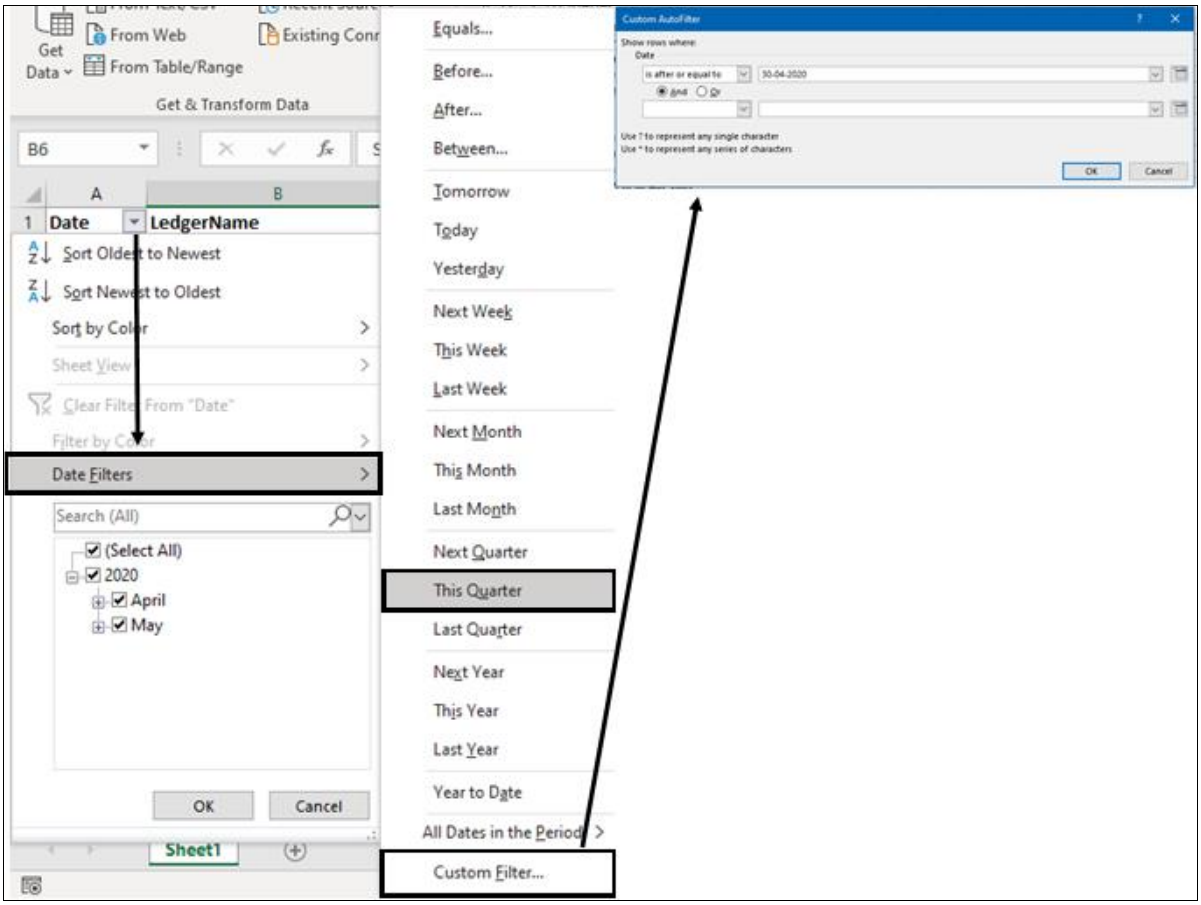


Figure 4.3.3 Date Filters

Text Filters

When you want to filter a text column for something very specific, there are several advanced options provided by Excel Text Filters such as:

- Filter cells that begin with or end with a specific character(s).
- Filter cells that contain or do not contain a given character or word anywhere in the text.
- Filter cells that are exactly equal or not equal to a specified character(s).

When you add a filter to a column containing text values, Text Filters will appear automatically in the AutoFilter.

As an example, to filter out rows containing “Stationery”:

- Click the **drop-down arrow** in the column heading, and point to Text Filters.
- In the drop-down menu, select the desired filter (Contains: Stationery).
- The Custom AutoFilter dialog box will show up.
- In the box to the right of the filter, type Stationery.
- Click OK.

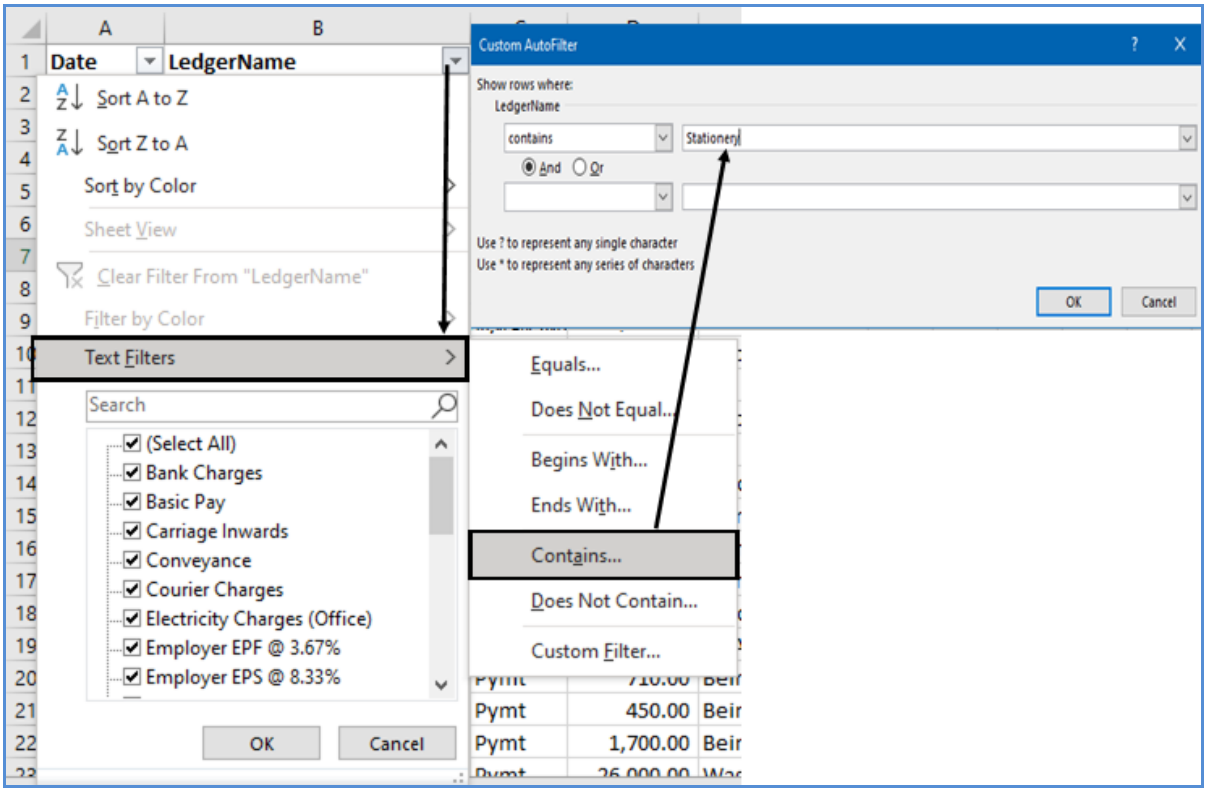


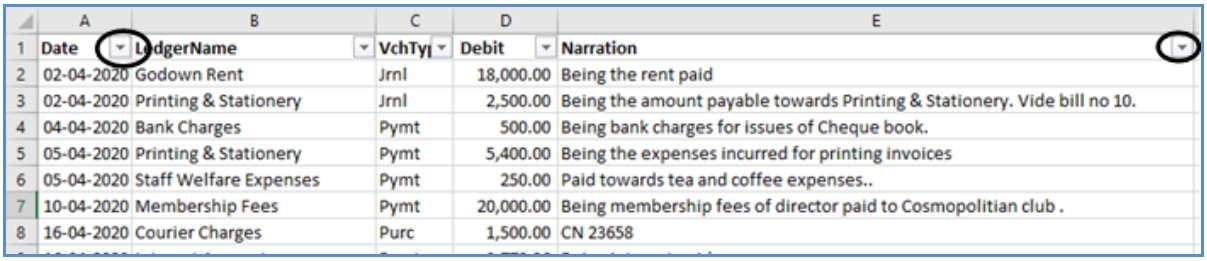
Figure 4.3.4 Text Filters

CASE STUDY: In an audit, your principal CA Chandiwala has given you a list of expenses in Excel drawn from the Accounting system of a company and asks you to find if any amount has been paid as a penalty in April so that the same can be reported in the Audit report. You are to find it using filters.

Strategy:

You can find records matching certain criteria by using **Filter** under **Data Tab** since you want to create a list of all expenses contain word "Penalty" in "April", you can use the Filter feature in Excel

1. Your data must **have Header Row.**
2. Select any one cell within data and select **Data > Filter**. Excel will add a dropdown to each heading as shown in Fig. 4.3.5



	A	B	C	D	E
1	Date	LedgerName	VchTyp	Debit	Narration
2	02-04-2020	Godown Rent	Jrnl	18,000.00	Being the rent paid
3	02-04-2020	Printing & Stationery	Jrnl	2,500.00	Being the amount payable towards Printing & Stationery. Vide bill no 10.
4	04-04-2020	Bank Charges	Pymt	500.00	Being bank charges for issues of Cheque book.
5	05-04-2020	Printing & Stationery	Pymt	5,400.00	Being the expenses incurred for printing invoices
6	05-04-2020	Staff Welfare Expenses	Pymt	250.00	Paid towards tea and coffee expenses..
7	10-04-2020	Membership Fees	Pymt	20,000.00	Being membership fees of director paid to Cosmopolitian club .
8	16-04-2020	Courier Charges	Purc	1,500.00	CN 23658

Figure 4.3.5 Excel will add a dropdown to each header

3. Select the **Date Dropdown**, Click the (Select All) checkbox to unselect all Names and then click the **April Box** as shown in Fig.4.3.6

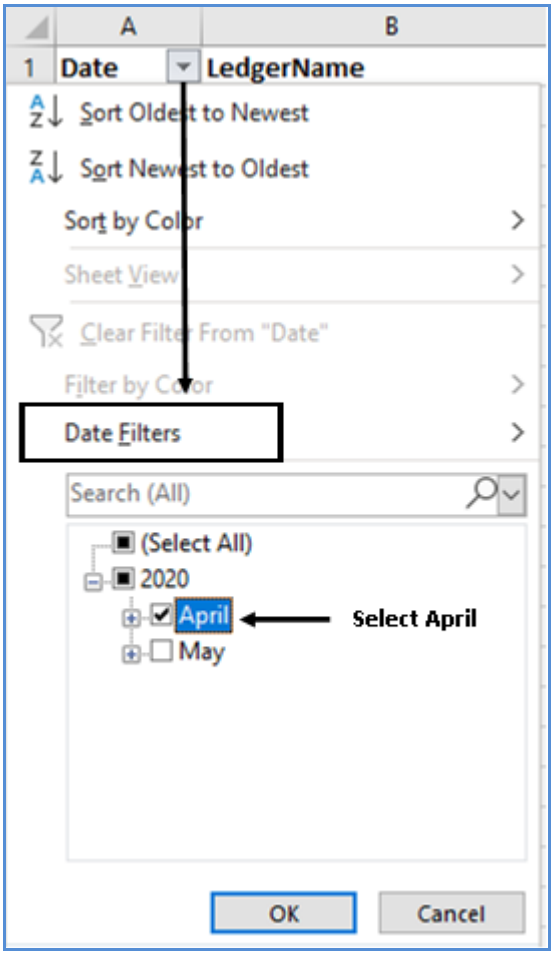


Figure 4.3.6 Date Filters

4. You will now see only rows where the dates are in **April** all other rows are now hidden.

5. Next, you want to know the amount paid for the penalty, you can search in narrations all the cases where narration contains the word "**Penalty**"

6. Select the **Narrations Dropdown**, Click **Text Filter** and go to **Contains**

7. **Custom Autofilter** dialog box will appear in right-hand box type “Penalty” as shown in Fig.4.3.7

8. Click OK

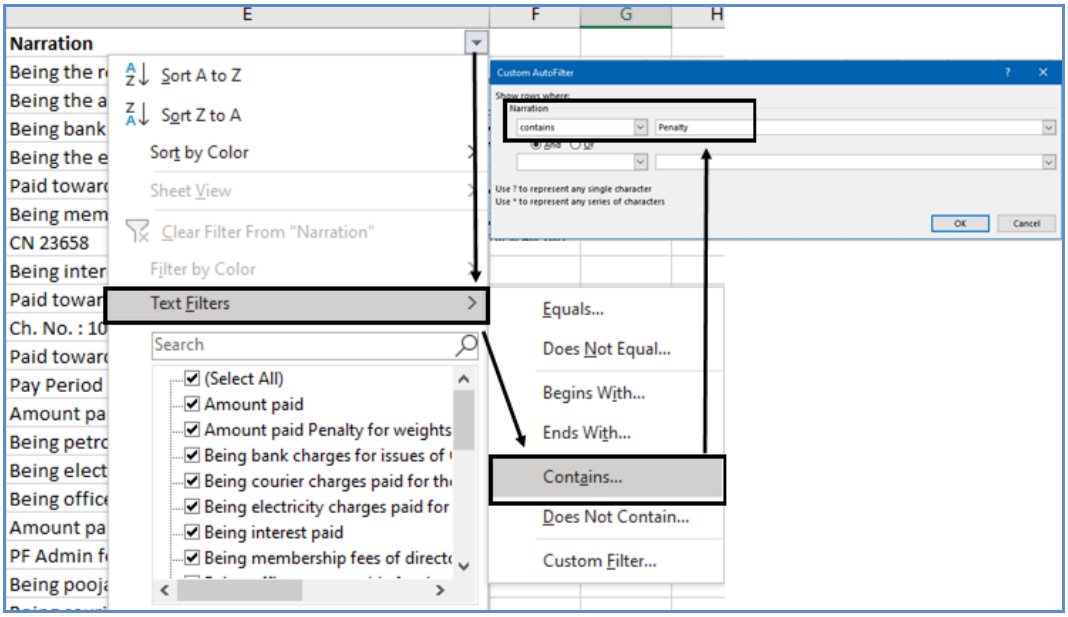


Figure 4.3.7 Use text filters

9. You will now see only rows where the narrations contain word Penalty all other rows are now hidden.

	A	B	C	D	E	F
1	Date	LedgerName	VchTyp	Debit	Narration	
18	30-04-2020	Fees & taxes	Pymt	12,000.00	Amount paid Penalty for weights & Measurement violation	

Figure 4.3.8 Filtered Data

Gist: You have created a list of Payments for a penalty in April using Filter Feature in excel

Commands learnt: Data > Filter (Date & Text)

4.3.3 ADVANCED FILTER

You can use Advance Filtering if you want to filter the data of more than one column.

You need to define your filtering criteria as a range.

In Excel Advanced Filter is helpful when you want to find data that meets two or more complex criteria such as extracting matches, filtering rows that match items in another list, finding exact matches including uppercase and lowercase characters, and many more complex cases

Advanced Filter vs. AutoFilter

When you compare with AutoFilter tool, Advanced Filter works differently in the following manner:

AutoFilter is a built-in capability that is applied in a single button click. You just have to click on the Filter button on the ribbon, and your Excel filter starts.

Advanced Filter cannot be applied automatically since it has no pre-defined setup, it requires manual configuration of the list range and criteria range manually.

AutoFilter allows filtering data with a maximum of 2 criteria,

With the use of **Advanced Filter,** you can find rows that meet multiple criteria in multiple columns, and the advanced criteria need to be entered in a separate range on your worksheet.

To open the Advanced Filter dialog box, click **Data > Advanced.**

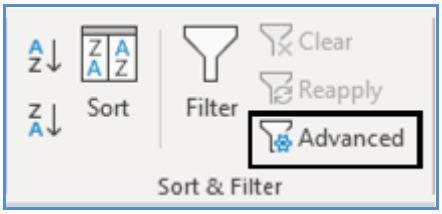


Figure 4.3.9 Advanced Filter

How to use Advanced Filter

Step 1

Arrange the source data

- Add a header row where each column has a unique heading
- Make sure there are no blank rows, Columns within your data set.

Step 2

Set up the criteria range

- Type your criteria, in a separate range on the worksheet. It is more practical to place it on top and separate from the data set with one or more blank rows.
    - The criteria range must have the same column headings as the range that you want to filter.
    - Criteria listed on the **same row** work with the **AND** logic.
    - Criteria entered on **different rows** work with the **OR** logic.
- As an example, to filter records for the *Galaxy Ltd* region with Amount *greater than or equal to 20000*, set up the following criteria range:

	A	B	C	D	E	F	G
1	DATE	INV_NO	CUSTOMER	AMOUNT		CUSTOMER	AMOUNT
2	03-04-2020	1	Milky Way	28,667.00		Galaxy Enterprises	>=20000
3	05-04-2020	2	Sun Moon & Star	23,589.00			

Figure 4.3.10 Set up Criteria

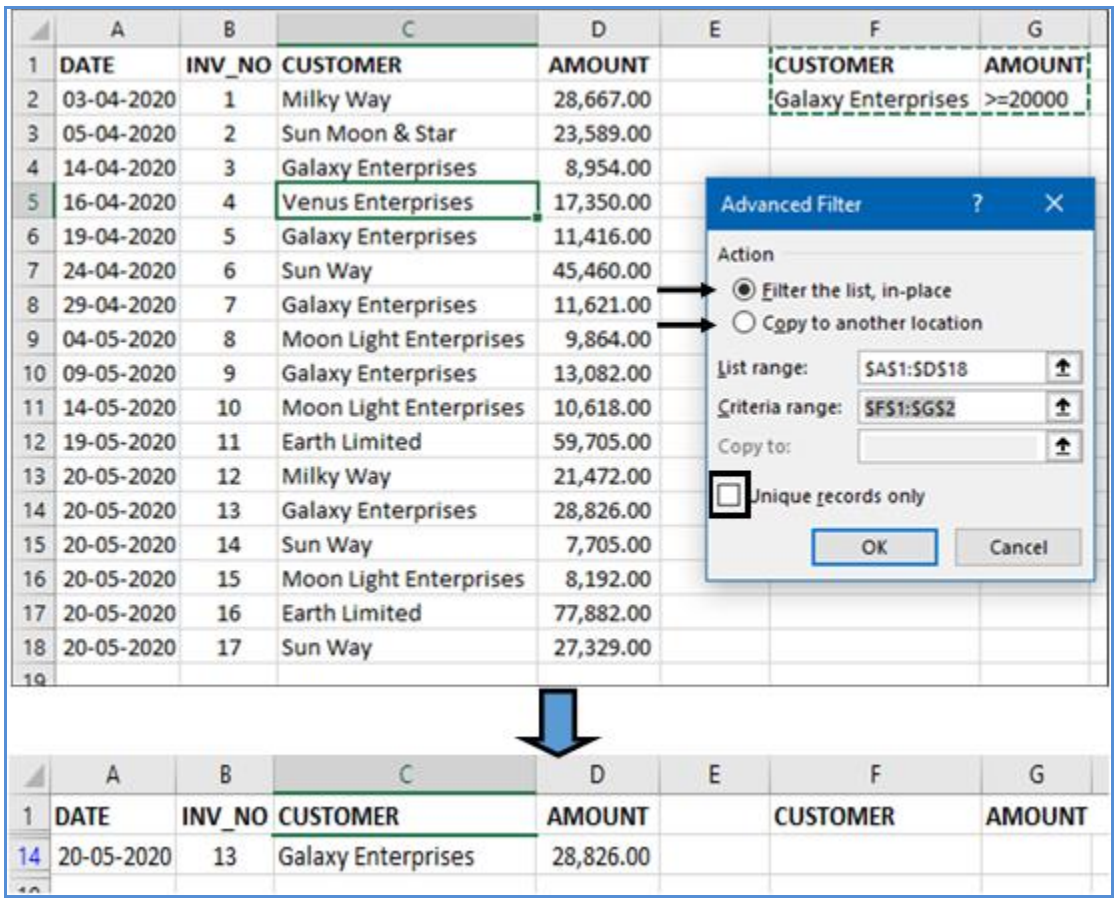
Step 3

Apply Excel Advanced Filter

- In the criteria range in place, apply an advanced filter in this way:
- Select any single cell within your dataset.
- Click on data> Advanced
- Excel Advanced Filter dialog box will appear
    - **Action**. Choose whether to **filter the list in place** or **copy** the results to another location.
    - Selecting "**Filter the list in place**" will hide the rows that don't match your criteria.
    - If you choose "**Copy the results to another location**", select the upper-left cell of the range where you want to paste the filtered rows
    - **List range**. It's the range of cells to be filtered, the column headings should be included.
    - When you select any cell in your data set before clicking the Advanced button, Excel selects the entire list range automatically.
    - **Criteria range**. It's the range of cells in which you **input the criteria**.

Remove Duplicates

- If you, check the box in the lower-left corner of the Advanced Filter dialog window lets you display **unique records** only.



	A	B	C	D	E	F	G
1	DATE	INV_NO	CUSTOMER	AMOUNT		CUSTOMER	AMOUNT
2	03-04-2020	1	Milky Way	28,667.00		Galaxy Enterprises	>=20000
3	05-04-2020	2	Sun Moon & Star	23,589.00			
4	14-04-2020	3	Galaxy Enterprises	8,954.00			
5	16-04-2020	4	Venus Enterprises	17,350.00			
6	19-04-2020	5	Galaxy Enterprises	11,416.00			
7	24-04-2020	6	Sun Way	45,460.00			
8	29-04-2020	7	Galaxy Enterprises	11,621.00			
9	04-05-2020	8	Moon Light Enterprises	9,864.00			
10	09-05-2020	9	Galaxy Enterprises	13,082.00			
11	14-05-2020	10	Moon Light Enterprises	10,618.00			
12	19-05-2020	11	Earth Limited	59,705.00			
13	20-05-2020	12	Milky Way	21,472.00			
14	20-05-2020	13	Galaxy Enterprises	28,826.00			
15	20-05-2020	14	Sun Way	7,705.00			
16	20-05-2020	15	Moon Light Enterprises	8,192.00			
17	20-05-2020	16	Earth Limited	77,882.00			
18	20-05-2020	17	Sun Way	27,329.00			

	A	B	C	D	E	F	G
1	DATE	INV_NO	CUSTOMER	AMOUNT		CUSTOMER	AMOUNT
14	20-05-2020	13	Galaxy Enterprises	28,826.00			

Figure 4.3.11 Setting criteria for advanced filter

Advanced Filter criteria using Formula

In advanced filter for creating more complex conditions, you can Excel functions in the criteria range. It is important to note for criteria based on formulae

- The formula must give a result of **TRUE or FALSE**.
- The criteria range should have a minimum of 2 cells: **heading cell and formula cell**.
- The **heading cell** in the formula-based criteria could be blank or heading different from any of the list range headings.
- If you want the formula to be evaluated for each row of data in the list range, use a relative reference ie without $ to refer to the cell in the first row of data.
- For the formula to be evaluated only for a specific cell or range of cells, use an absolute reference ie with $ to refer to that cell or range.
- When **referencing the list range** in the formula, always use **absolute cell references** ie with $.

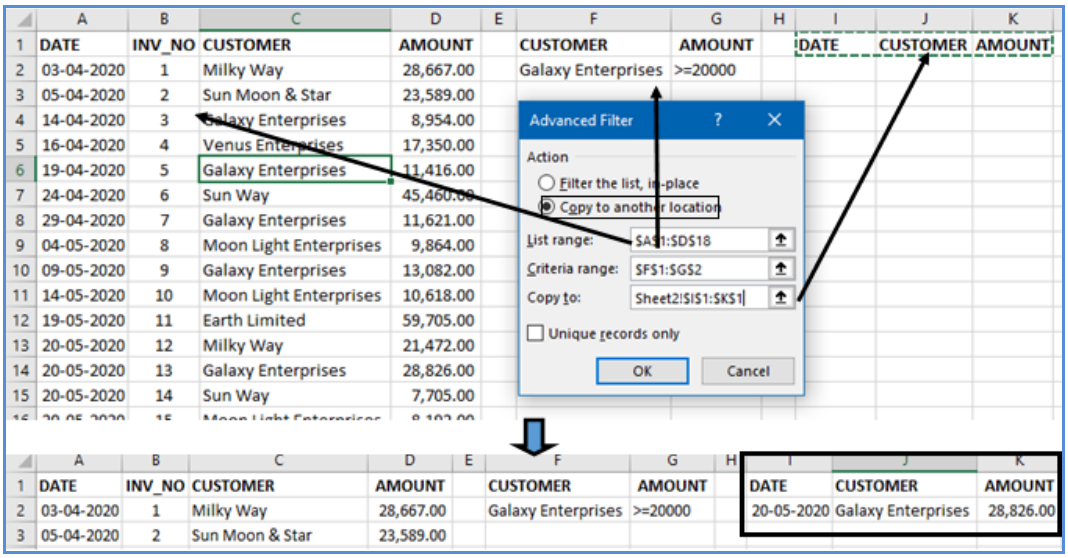


Figure 4.3.12 Copying with Advanced Filter

***CASE STUDY:** In a Bank Investigation you are given a Loan against the FDR list with Rate of Interest on Loan & Rate of interest on FDR as shown in Fig. 1.3.1. It is suspected that an employee has taken a bribe to reduce the interest being charged on some loan accounts. Bank's interest rate circular states that in case of loan against FDR the spread of interest would be 1%. You want to extract at a different location, cases where the interest charged on loan is not 1% more than the Rate of Interest paid on FDR.*

	A	B	C	D	E	F
1	ACC_NO	BORROWER NAME	OUTSTANDING	ROI_LN	ACC_NO_FDR	ROI_FDR
2	155372479	Abhijit	6,74,093.00	13.00	598492374	12.00
3	154216698	Akhil	19,89,808.00	10.00	547622850	9.00
4	153744046	Aravind	3,77,294.00	10.00	545366397	8.00
5	152983361	Arul	3,84,401.00	9.00	583084278	8.00
6	157308461	Ashish	10,44,960.00	11.00	593695151	10.00
7	154666457	Azam	7,10,888.00	10.00	540914226	9.50
8	159965064	Biren	3,80,344.00	13.00	570294241	12.00
9	153132065	Bishakha	9,45,627.00	7.00	598320416	6.00
10	159968598	Charles	6,00,243.00	6.00	576371671	5.00
11	155632102	Dayal	4,70,210.00	12.00	581094800	11.50
12	155509141	Deepak	5,99,276.00	7.00	580429851	6.00
13	157758384	Jayesh	8,48,811.00	8.00	590708529	7.00
14	154987277	Joji	6,46,972.00	8.00	597047061	7.00
15	155978002	Mitali	16,47,208.00	9.00	514466758	8.00

Figure 4.3.13 Loan data

Strategy:

You can find records matching certain criteria by using **Filter** under **Data Tab** and using **Advanced Filter**

Since spread ie Rate of Interest Charged on Loan & ROI Paid on FDR should be 1, you can set up criteria to extract all those cases where the spread is not 1 ie **D2-F2<>1** if the condition is **true** will give us all anomalies, then you can investigate further.

	A	B	C	D	E	F	G	H	I
1	ACC_NO	BORROWER NAME	OUTSTANDING	ROI_LN	ACC_NO_FDR	ROI_FDR		CRITERIA	
2	155372479	Abhijit	6,74,093.00	13.00	598492374	12.00		=(D2-F2)<>1	
3	154216698	Akhil	19,89,808.00	10.00	547622850	9.00			
4	153744046	Aravind	3,77,294.00	10.00	545366397	8.00			

Figure 4.3.14 Set up Criteria

1. Select any one cell within data and select **Data** > **Advanced**.
2. In the Advanced Filter dialog box give the following settings:
   - Select "**Copy to another location"**
   - List Range "$A$1:$F$15
   - Criteria Range: "$H$1:$H$2
   - Copy to $J$1 ie where you want the entire extraction done, If you want only specified columns then you will have to give the headers at the destination location

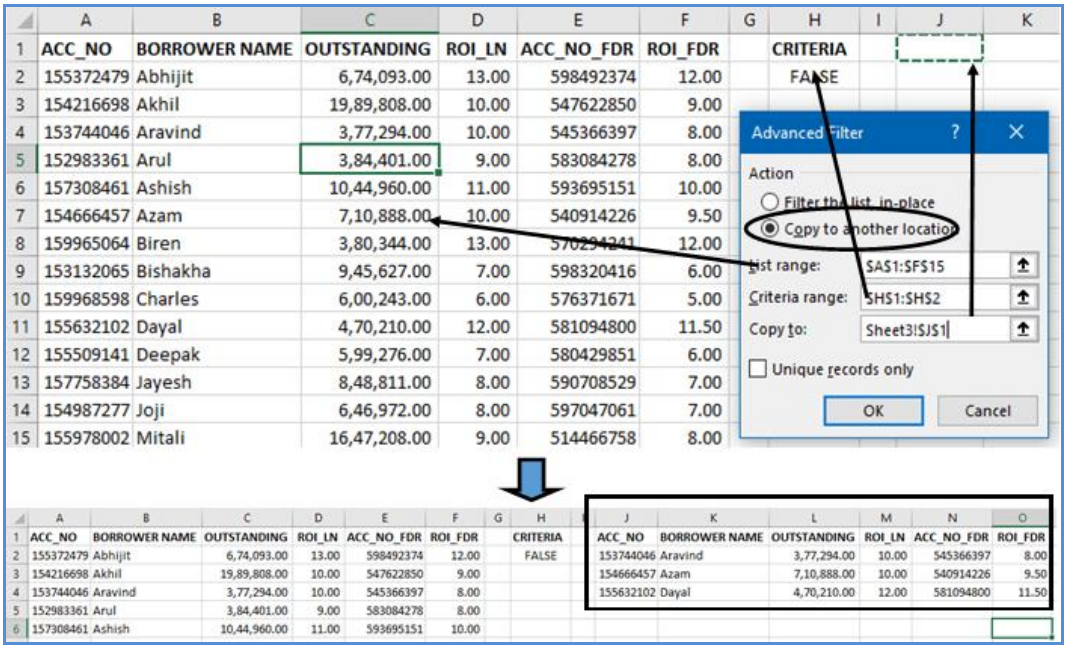


Figure 4.3.15 Extraction with formula

3. Your list of all cases, where the spread is other than 1 are extracted.

Gist: You have extracted a list of cases where the spread of ROI charged on loan and ROI paid on FDR is other than 1%.

Commands learnt: Data > Filter- Advanced options

4.4 OUTLINING DATA

4.4.1 GROUPING/ UNGROUPING DATA

Worksheets with complex information and large content can sometimes feel overwhelming and even become difficult to read. Microsoft Excel provides an easy way to organize data in groups allowing you to hide and show rows or columns with related content to create a compact and understandable view. You can create an outline of up to eight levels. You can also summarize different groups using the Subtotal command and create an outline for your worksheet.

Grouping in Excel is good for structured worksheets that have column headings, no blank rows or columns, and a Subtotal row for each subset of rows. If your data is properly organized, you can easily group it.

An Outline can be created manually or automatically.

Manual - You identify each group of related data.

Automatic - This is typically used when you have a table that contains totals and subtotals.

4.4.2 GROUPING DATA MANUALLY

If your worksheet contains two or more levels of information, you can group rows manually through the following steps.

- Select the rows or columns you want to group
- Select the Data tab on the Ribbon, then click the Group command.
- The selected rows or columns will be grouped.
- You can repeat the steps above as many times as you like.

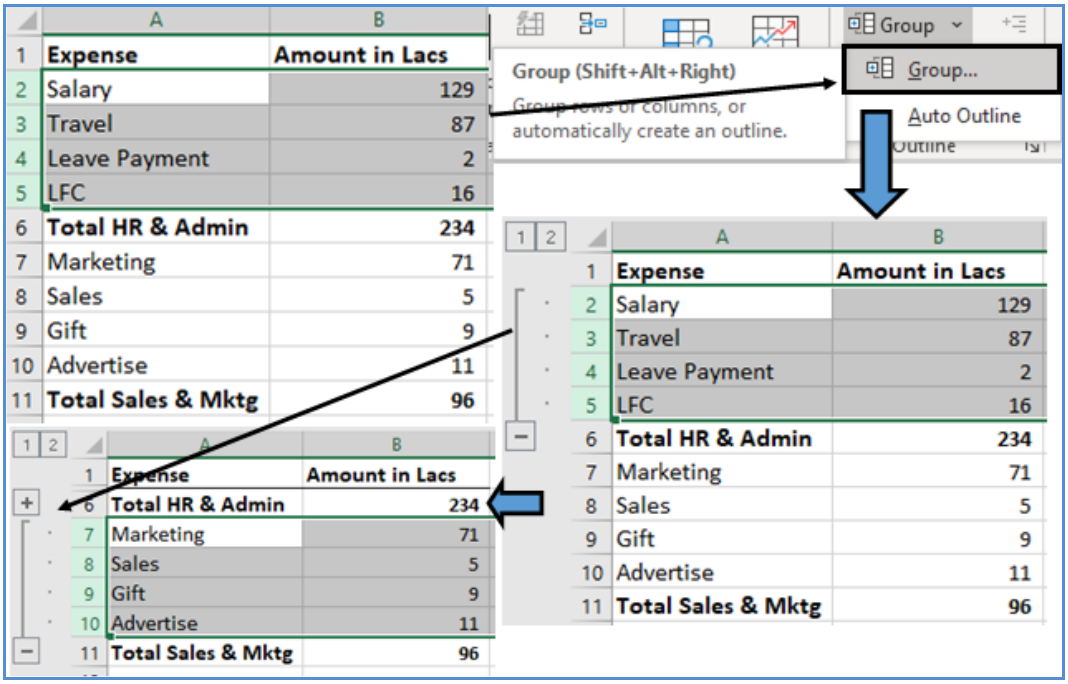


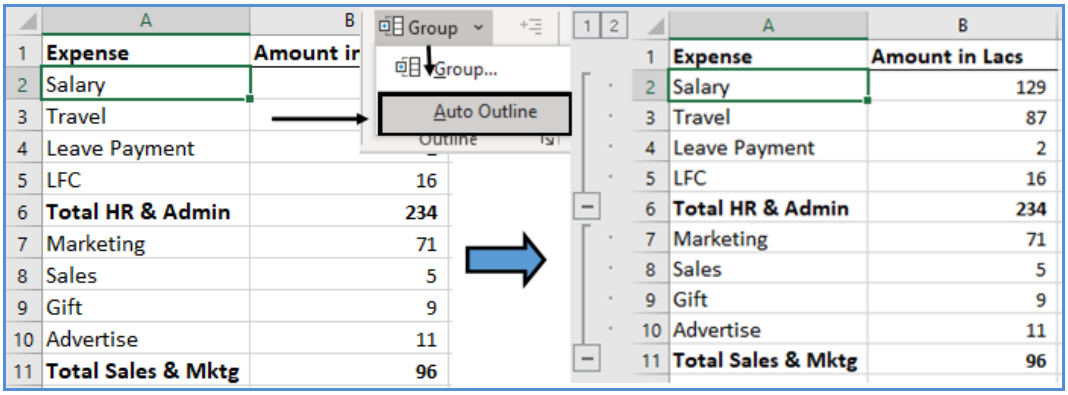
Figure 4.4.1 Grouping data

- To hide/collapse a group, click the Hide Detail button.
- The group will be hidden. To show a hidden group, click the Show Detail button.

4.4.3 AUTO GROUPING DATA (CREATE AN OUTLINE)

If your dataset contains Totals and subtotals ie that have been created out of functions Subtotal or Sum, the fastest way would be to let Excel group rows for you automatically.

- Select any cell in one of the rows you want to group.
- Go to the **Data tab > Outline group**,
- Click the arrow under Group,
- Select Auto Outline.



	A	B
1	Expense	Amount ir
2	Salary	
3	Travel	
4	Leave Payment	
5	LFC	16
6	**Total HR & Admin**	234
7	Marketing	71
8	Sales	5
9	Gift	9
10	Advertise	11
11	**Total Sales & Mktg**	96

	A	B
1	Expense	Amount in Lacs
2	Salary	129
3	Travel	87
4	Leave Payment	2
5	LFC	16
6	**Total HR & Admin**	234
7	Marketing	71
8	Sales	5
9	Gift	9
10	Advertise	11
11	**Total Sales & Mktg**	96

Figure 4.4.2 Auto Outline Data

4.4.4 UNGROUPING DATA

To remove grouping for certain rows without deleting the whole outline.

- **Select the rows** you want to ungroup.
- Go to the **Data tab > Outline group**,
- Click the **Ungroup** button.
- In the **Ungroup** dialog box,
- Select Rows and click OK.

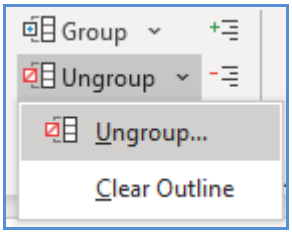


Figure 4.4.3 Ungroup Data

- The same process can be used to **clear outline**

CASE STUDY: Your principal CA Chandiwala gives you an excel sheet containing the Fixed Assets chart as shown in Figure 4.4.4. He wants you to organise it in such a way that when you want to present or print it, you could be a brief outline, but when you want to check details or audit it you can show the details.

	A	B	C	D	E
1	**Asset**	**Opening WDV**	**Addition**	**Depreciation**	**Closing WDV**
2	Building Unit 1	2,67,82,915.00	-	-26,78,292.00	2,41,04,623.00
3	Building Unit 2	5,55,81,592.00	-	-55,58,159.00	5,00,23,433.00
4	BUILDING TOTAL	8,23,64,507.00	-	-82,36,451.00	7,41,28,056.00
5	Canopy	3,41,953.00	1,00,000.00	-44,195.00	3,97,758.00
6	Furniture & Fix.	16,73,323.00	-	-1,67,332.00	15,05,991.00
7	Fixed Fittings	34,56,789.00	-	-3,45,679.00	31,11,110.00
8	FURNITURE & FIXTURE TOTAL	54,72,065.00	1,00,000.00	-5,57,206.00	50,14,859.00
9	FIXED ASSETS TOTAL	8,78,36,572.00	1,00,000.00	-87,93,657.00	7,91,42,915.00

Figure 4.4.4 Fixed assets schedule

Strategy:

You can use **GROUP** from the **OUTLINE** group on the **Data Ribbon**.

Since the data has Totals and Subtotals, you can easily use the **Auto Outline** feature of the group.

1. Select anywhere in the data.
2. Go to Data tab> Group >outline as shown in Figure 4.4.5

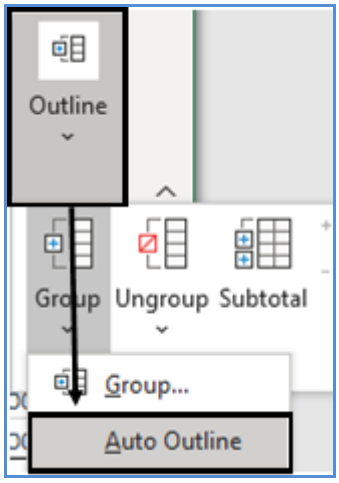
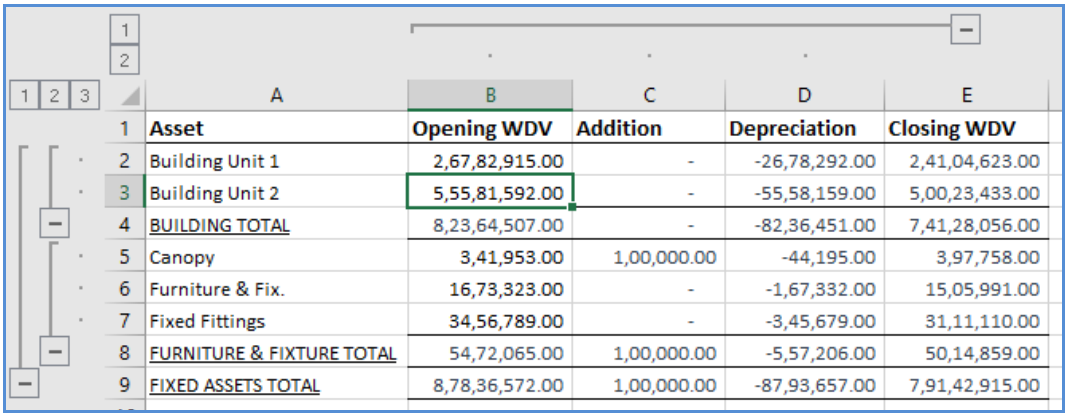


Figure 4.4.5 Auto Outline

3. Click on the **Auto outline**, based on Totals and subtotals Excel will automatically create subtotal as shown in Figure 4.4.6.

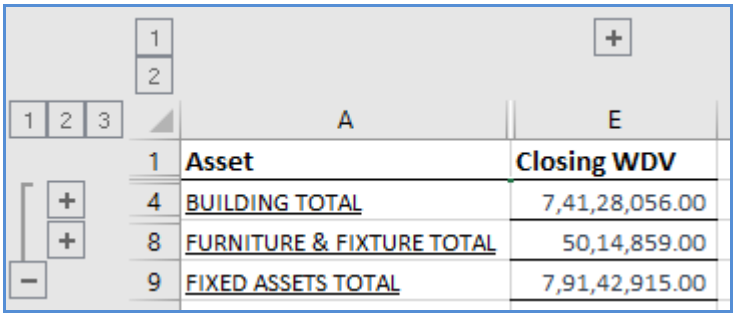


	A	B	C	D	E
1	Asset	Opening WDV	Addition	Depreciation	Closing WDV
2	Building Unit 1	2,67,82,915.00	-	-26,78,292.00	2,41,04,623.00
3	Building Unit 2	5,55,81,592.00	-	-55,58,159.00	5,00,23,433.00
4	BUILDING TOTAL	8,23,64,507.00	-	-82,36,451.00	7,41,28,056.00
5	Canopy	3,41,953.00	1,00,000.00	-44,195.00	3,97,758.00
6	Furniture & Fix.	16,73,323.00	-	-1,67,332.00	15,05,991.00
7	Fixed Fittings	34,56,789.00	-	-3,45,679.00	31,11,110.00
8	FURNITURE & FIXTURE TOTAL	54,72,065.00	1,00,000.00	-5,57,206.00	50,14,859.00
9	FIXED ASSETS TOTAL	8,78,36,572.00	1,00,000.00	-87,93,657.00	7,91,42,915.00

Figure 4.4.6 Grouped Data

4. If you look above and to the left of cell A1, you'll see a series of three small numbers,1,2,3, Excel has created 3 level row grouping and grouping for columns as well.

5. You can expand or collapse depending on the need for the presentation.



	A	E
1	Asset	Closing WDV
4	BUILDING TOTAL	7,41,28,056.00
8	FURNITURE & FIXTURE TOTAL	50,14,859.00
9	FIXED ASSETS TOTAL	7,91,42,915.00

Figure 4.4.7 Collapsed View

Gist: You have created a well organised Fixed asset list that you can expand or collapse depending on the need.

Commands learnt: Data > Group >Auto Outline

4.4.5 SUBTOTALS

Many times you need to show more complex information arranged in Tables, especially if it's split into separate groups, each with its **SUBTOTALS.** As an example, you could have data that could make more sense if you have Quarterly or Yearly or Region-wise subtotals. The Subtotal tool lets us create groups and subtotals all in one click—a feature that can save us lots of time.

Excel's Subtotal feature helps you draw specific information from one or more subsections of a database or list of data.

In addition to finding the subtotal for selected rows of data, you can also use the Subtotal feature to find, among other things, the average value for the selected range of data, the largest and smallest values, or the total number of rows of data included in the range.

CASE STUDY: Your principal CA Chandiwala gives you an excel sheet containing datewise payment made to contractors as shown in Figure 4.4.8 . He wants you to check for cases where total freight paid to any Contractor is more than Rs.1,00,000 so that it could check for TDS compliance under Income Tax.

	A	B	C
1	**Date_Pay**	**Contractor**	**Amount**
2	09-04-2019	Amar	29,300
3	12-04-2019	Anthony	28,800
4	03-05-2019	Amar	28,700
5	22-07-2019	Amar	29,200
6	11-11-2019	Amitabh	21,400
7	05-12-2019	Amitabh	28,200
8	06-12-2019	Amitabh	28,300
9	07-12-2019	Akbar	23,800
10	07-12-2019	Anthony	25,400
11	17-12-2019	Amitabh	21,500
12	01-01-2020	Akbar	22,900
13	02-01-2020	Akbar	27,800
14	15-01-2020	Anthony	23,200
15	30-01-2020	Akbar	8,000
16	03-02-2020	Anthony	27,500
17	21-03-2020	Akbar	15,500
18	21-03-2020	Anthony	22,900
19	24-03-2020	Amar	4,500
20	30-03-2020	Amitabh	15,000

Figure 4.4.8 Contactor payments

Strategy:

You can use **SUBTOTALS** From the **OUTLINE** group on the **Data Ribbon** as shown in Fig.4.4.9

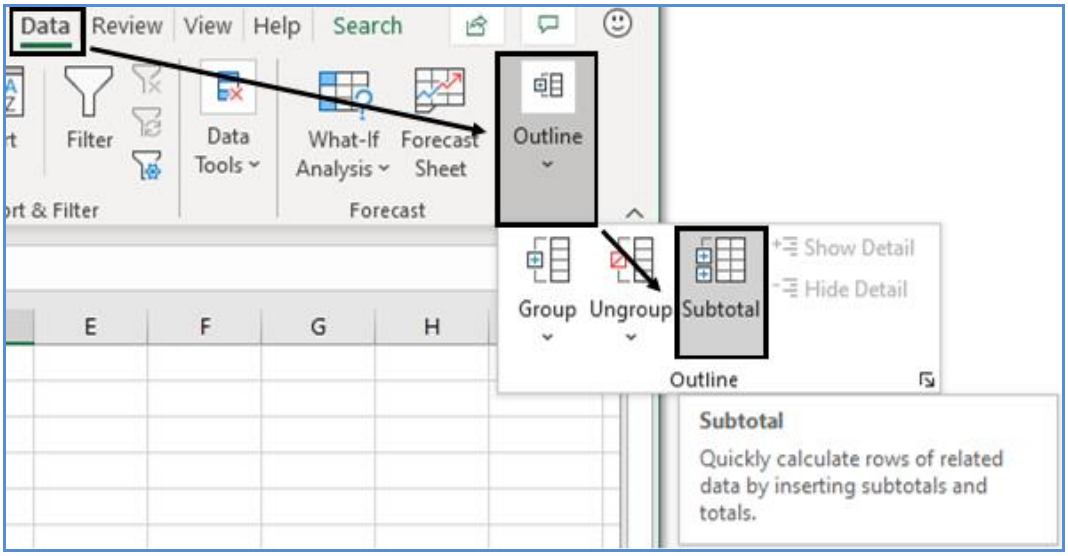
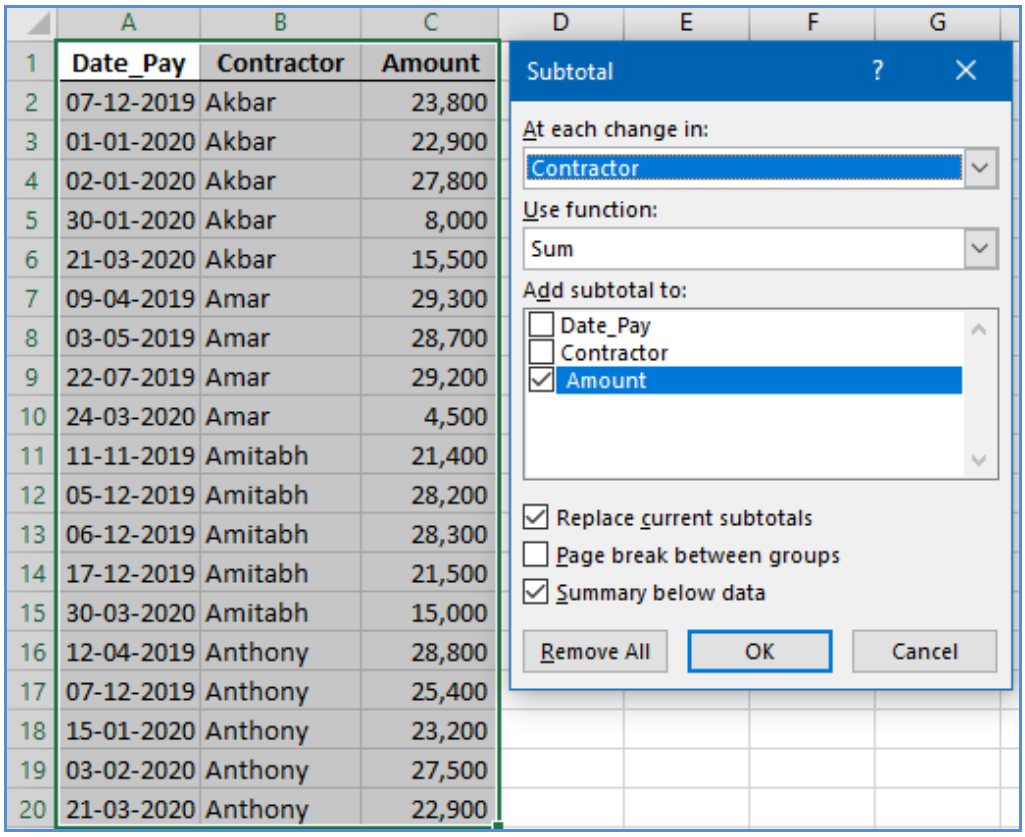


Figure 4.4.9 Subtotal

1. Sort your list by "*Contractor".*
2. Choose **SUBTOTALS** from the **Outline** group on the **Data** Ribbon the dialog box will appear as shown in Figure 4.4.10



	A	B	C
1	Date_Pay	Contractor	Amount
2	07-12-2019	Akbar	23,800
3	01-01-2020	Akbar	22,900
4	02-01-2020	Akbar	27,800
5	30-01-2020	Akbar	8,000
6	21-03-2020	Akbar	15,500
7	09-04-2019	Amar	29,300
8	03-05-2019	Amar	28,700
9	22-07-2019	Amar	29,200
10	24-03-2020	Amar	4,500
11	11-11-2019	Amitabh	21,400
12	05-12-2019	Amitabh	28,200
13	06-12-2019	Amitabh	28,300
14	17-12-2019	Amitabh	21,500
15	30-03-2020	Amitabh	15,000
16	12-04-2019	Anthony	28,800
17	07-12-2019	Anthony	25,400
18	15-01-2020	Anthony	23,200
19	03-02-2020	Anthony	27,500
20	21-03-2020	Anthony	22,900

Figure 4.4.10 Subtotal Dialog Box with options

3. Click the dropdown list arrow to the right of the "**At each change in"** box to display Your column headings.
4. Select *Contractor* ie the heading of the column you want to sort the data.
5. Out of the dropdown list **Use Function** select **SUM,** (there are other operations also)
6. Use the scroll arrows in **add subtotal to** and check the box *Amount*. You can add subtotals to more than one column.
7. Set the options for subtotal.
8. If you want each subtotaled group on a separate page, select **Page Break between groups**.
9. **The summary below Data** adds subtotals and an overall Total below the groups rather than above.
10. To ensure that new subtotals overwrite any existing subtotals select **Replace Current Subtotals.**

11. Click **OK**.

	A	B	C
1	Date_Pay	Contractor	Amount
2	07-12-2019	Akbar	23,800
3	01-01-2020	Akbar	22,900
4	02-01-2020	Akbar	27,800
5	30-01-2020	Akbar	8,000
6	21-03-2020	Akbar	15,500
7		**Akbar Total**	98,000
8	09-04-2019	Amar	29,300
9	03-05-2019	Amar	28,700
10	22-07-2019	Amar	29,200
11	24-03-2020	Amar	4,500
12		**Amar Total**	91,700
13	11-11-2019	Amitabh	21,400
14	05-12-2019	Amitabh	28,200
15	06-12-2019	Amitabh	28,300
16	17-12-2019	Amitabh	21,500
17	30-03-2020	Amitabh	15,000
18		**Amitabh Total**	1,14,400
19	12-04-2019	Anthony	28,800
20	07-12-2019	Anthony	25,400
21	15-01-2020	Anthony	23,200
22	03-02-2020	Anthony	27,500
23	21-03-2020	Anthony	22,900
24		**Anthony Total**	1,27,800
25		**Grand Total**	4,31,900

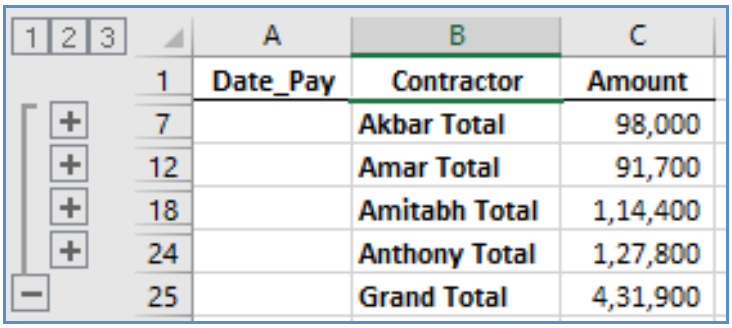
Figure 4.4.11 Contractorwise payments

Excel gives us tools for collapsing or expanding the lists. If you look above and to the left of cell A1, you'll see a series of three small numbers,1,2,3 as shown in Fig.4.4.11 These are the Group & Outline buttons.

There are three types of controls in the outline section:

- **Hide detail buttons** When the rows in a group are visible, a hide detail button "-" appears next to the group.
- **Show detail buttons** When you hide a group of rows, the button next to the group changes to a show detail button "+". Clicking a show detail button restores the rows in that group to the worksheet.
- **Level buttons** Each of the numbered level buttons 1 2 3 represents a level of organization in a worksheet; clicking a level button hides all levels of detail below that of the button you clicked.

12. You can use them to collapse subtotaled data to get the summary report you're looking for as shown in Fig



	A	B	C
1	Date_Pay	Contractor	Amount
7		Akbar Total	98,000
12		Amar Total	91,700
18		Amitabh Total	1,14,400
24		Anthony Total	1,27,800
25		Grand Total	4,31,900

Figure 4.4.12 Contractorwise summary of the amount paid

13. You can then use **Filter** to filter for the case where the Overall amount paid is more than Rs. 1,00,000.

Gist: You have created a list of cases where the aggregate of the amount paid to a contractor is more than Rs.1,00,000 and you can check for TDS compliances in those areas.

Commands learnt: Data > Subtotal

Subtotal can also be used with the following Summary Functions for data analysis:

- **Sum** The sum of the values. This is the default function for numeric source data
- **Count** The number of items. The Count summary function works the same as the COUNTA worksheet function. The count is the default function for source data other than numbers
- **Average** The average of the values.
- **Max** The largest value.
- **Min** The smallest value.
- **Product** The product of the values.
- **Count Nums** The number of rows that contain numeric data. The Count Nums summary function works the same as the COUNT worksheet function.
- **StdDev** An estimate of the standard deviation of a population, where the sample is all of the data to be summarized.
- **StdDevp** The standard deviation of a population, where the population is all of the data to be summarized.
- **Var** An estimate of the variance of a population, where the sample is all of the data to be summarized.
- **Varp** The variance of a population, where the population is all of the data to be summarized.

4.5 SUMMARY

Often you get data which needs to be organised to be understandable, Excels offers a host of feature for orderly organisation of data.

In this chapter, you learned how to you can sort our data, not only with predefined order but you can create your own custom lists and sort according to your custom lists. You have further learnt to Filter the data with text as well as date filters.

You also understood how apart from inbuilt filtering capabilities you can use Advance Filter to create complex criteria for filters. You have further learnt to Outline data using Grouping both manually and Automatically using Auto outline. Creating Quick subtotals has also been addressed in this chapter.

REFERENCES

[1] Michael Alexander, Richard Kusleika, John Walkenbach, 'Excel 2019 Bible', Wiley,2018

[2] Greg Harvey, 'Excel 2019 All-in-One for Dummies, John Wiley & Sons,2018

[3] www.chandoo.org

CHAPTER

5 EXCEL – FORMULA AND FUNCTIONS

LEARNING OBJECTIVES

- Understanding cell references
- Understanding MATHEMATICAL Functions
- Understanding LOGICAL Functions
- Understanding LOOKUP Functions
- Understanding Database Functions
- Understanding various types of ERRORS

5.1 INTRODUCTION TO FUNCTIONS

Formulas are the life of Excel worksheet. A formula identifies the calculation needed to place the result in the cell it is contained within.

A cell containing a formula has two display components –

- the formula itself and
- the resulting value.

The formula can be seen either by double clicking on the particular cell or by clicking once in the cell, then formula is visible in Formula Bar along with the output in that particular cell.

A formula assigns values to a cell or a range of cells, and typically has the format:

= Expression

Example: = SUM(A2:B2)

Where the expression consists of:

- values, such as 2, 9, 14 or 6.67E-11;
- references to other cells, e.g., A1 for a single cell, A1:A3 or A1..A3 for a range;
- arithmetic operators, such as + , - , *, / , and others;
- relational operators, such as > , = , < , and others; and,
- functions, such as SUM(), AVG(), and many others.

5.2 CELL REFERENCE

When a cell contains a formula, it often contains references to other cells. The Cell References are of 3 types:

- **Relative Cell Reference:** In the below example, in cell D2, there is reference of cell A2 (visible in Formula Bar) and the output is value "A" as in cell A2. When the value in cell D2 is copied in cell D3 or cell E2, the reference changes automatically to A3 or B2 respectively. As the reference changes as per change in the location of cell relative to the original position, it is called as "Relative Cell Reference".

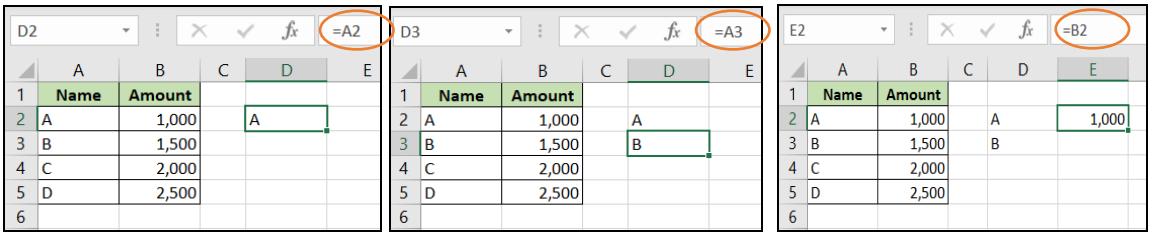


Figure 5.1 – Relative Cell Refence

- **Absolute Cell Reference:** In Absolute Cell Reference, the dollar ($) sign is preceded before Row reference and before Column reference. Because of which, the reference does not change irrespective of copying at any cell in the same sheet /other sheet / sheet in other Excel file. The dollar sign can be placed after double clicking on that cell or while entering cell reference in the cell. The $ sign can be placed individually before Row reference and before Column reference or it can be done using F4 key.

  In the below example, in cell D2, there is reference of cell A2 along with $ sign i.e. $A$2 (visible in Formula Bar). When the D2 cell value is copied to cell D3 or cell E2, the cell reference and the value remains the same i.e. $A$2 and "A" respectively.

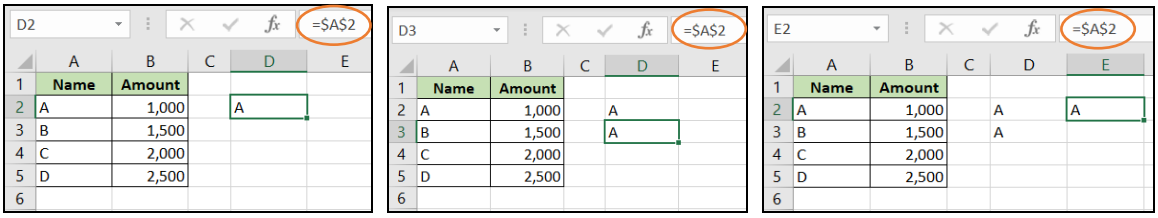


Figure 5.2 – Absolute Cell Refence

- **Mixed Cell Reference:** In case of Mixed Cell Reference, either row reference is fixed or column reference is fixed. It means only one $ sign is used either before Row reference or Column reference.

  In the below example (figure 3.1), in cell D2, there is reference of cell A2 along with $ sign in column reference i.e. $A2 (visible in Formula Bar). When D2 cell value is copied in cell D3, we observe no change in column reference but row reference changes and it becomes $A3 and output is "B". Similarly, when we copy to cell E2, as the row number remains same, we observe no change in row reference and column reference does not change because of $ sign. Hence cell reference remains same i.e. $A2 and the output is "A".

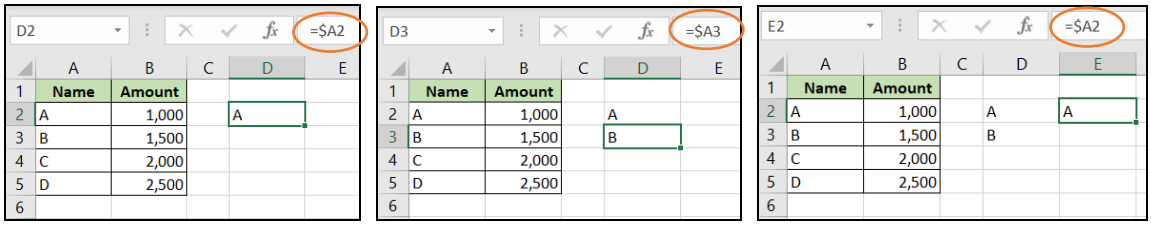


Figure 5.3 – Mixed Cell Refence – Column reference is fixed

In the below example (figure 3.2), in cell D2, there is reference of cell A2 along with $ sign in row reference i.e. A$2 (visible in Formula Bar). When D2 cell value is copied in cell D3, we observe no change in row reference because of $ sign and as column remains same i.e. "D" column, there is no change in column reference. Hence, cell reference remains same i.e. A$2 and output is "A". Similarly, when we copy to cell E2, the row number remains same because of $ sign but column reference changes to "B" column. Hence cell reference becomes B$2 and the output is "1000".

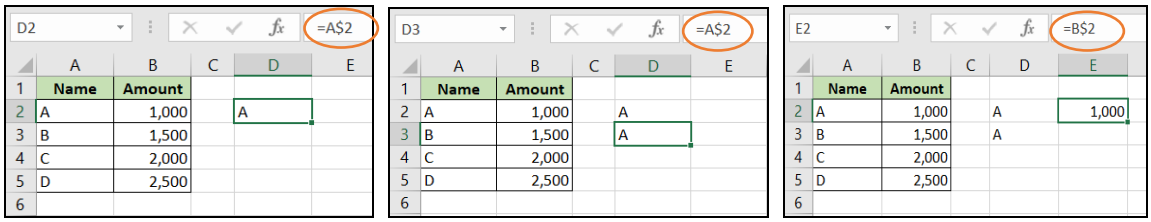


Figure 5.4 – Mixed Cell Refence – Row reference is fixed

Note: Depending on the requirement, we need to use either of these 3 cell references in the formula.

5.3 TEXT FUNCTIONS

There are several Text Functions in Excel – LEFT Function, RIGHT Function, MID function, LEN Function, FIND Function, SEARCH Function, SUBSTITUTE Function, REPLACE Function, TRIM Function, CONCATENATE Function and Concatenate Operator "&".

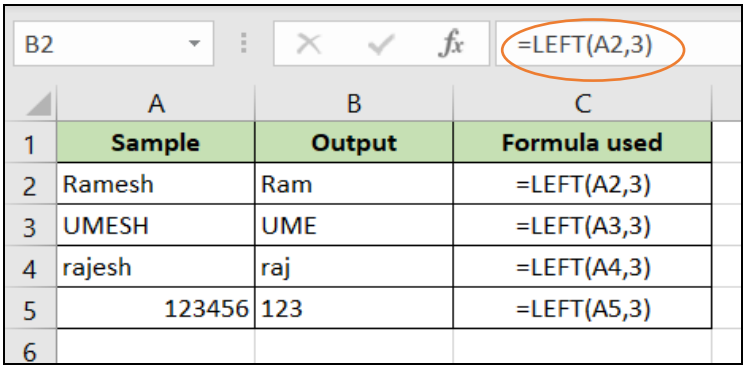
- **LEFT Function:** LEFT Function extracts a given number of characters from the left side (beginning) of the selected text string. The general syntax is as follows:

  = LEFT(text , [num_chars])

  Here the "text " is the text from where we want to extract "number of characters" from left side using LEFT Function.

  (**Note:** The number of characters can be changed as per requirement. In the example, we are extracting 3 characters.)

  For example, in the below table, we want to extract 3 characters from left side of the Sample Column (Column A). We entered the function in cell B2 as "=LEFT(A2,3)" and then copied the function in below rows till row 5. The output is as follows:



B2 =LEFT(A2,3)

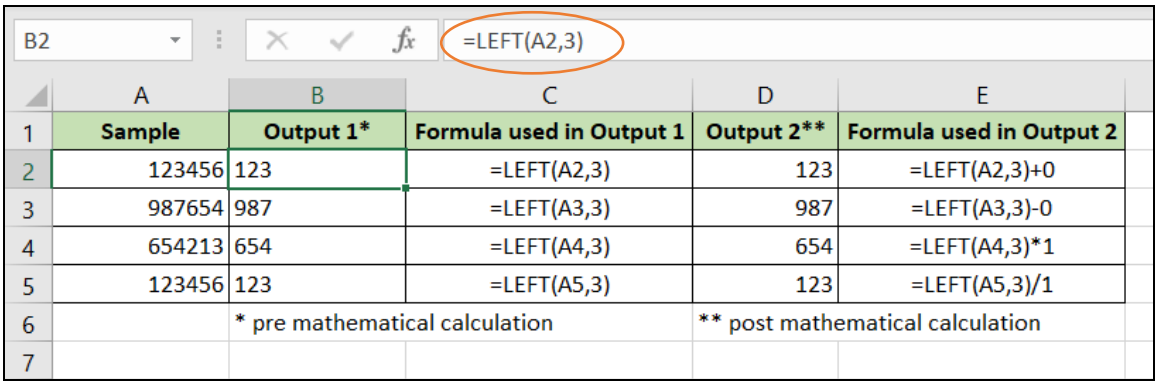
	A	B	C
1	**Sample**	**Output**	**Formula used**
2	Ramesh	Ram	=LEFT(A2,3)
3	UMESH	UME	=LEFT(A3,3)
4	rajesh	raj	=LEFT(A4,3)
5	123456	123	=LEFT(A5,3)
6			

Figure 5.5 – LEFT Function

It can be seen that LEFT Function has extracted 3 characters from the left side of all the cells selected. LEFT Function may also be used for extracting few digits from a given number. In case of cell A5, the value is a number, but LEFT Function being the TEXT Function, it has converted number in Text form and that is why the output is left aligned.

Being output is in Text form, no calculation can be performed on it except count. In case, we want to convert it back to the numeric value in output cell B5, we need to perform **any one** mathematical calculation (i.e. +0, -0, *1, /1 in cell B5 and the final output will be in numerical form.

In the below example, we have numerical values in Sample Column and we want to extract 3 characters from left side. In Output 1 an Output 2 column, we can see the final output (pre and post mathematical calculation).



B2 =LEFT(A2,3)

	A	B	C	D	E
1	**Sample**	**Output 1***	**Formula used in Output 1**	**Output 2****	**Formula used in Output 2**
2	123456	123	=LEFT(A2,3)	123	=LEFT(A2,3)+0
3	987654	987	=LEFT(A3,3)	987	=LEFT(A3,3)-0
4	654213	654	=LEFT(A4,3)	654	=LEFT(A4,3)*1
5	123456	123	=LEFT(A5,3)	123	=LEFT(A5,3)/1
6		* pre mathematical calculation		** post mathematical calculation	
7					

Figure 5.6 – LEFT Function with mathematical calculation

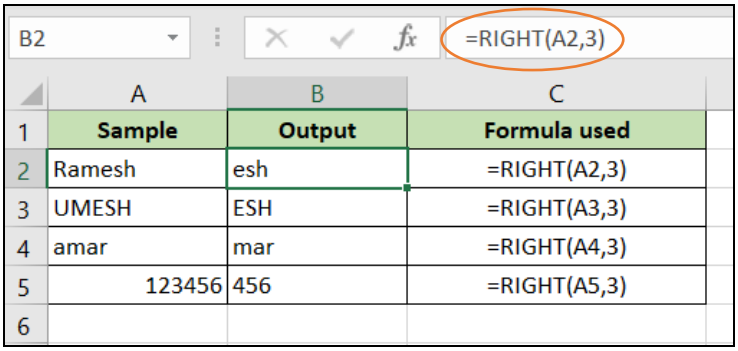
- **RIGHT Function:** Similar to LEFT Function, the RIGHT Function extracts a given number of characters from the right side (end) of the selected text string. The general syntax is as follows:

  = RIGHT(text , [num_chars])

  Here the "text " is the text from where we want to extract "number of characters" from right side using RIGHT Function.

  (**Note:** The number of characters can be changed as per requirement. In the example, we are extracting 3 characters.)

For example, in the below table, we want to extract 3 characters from right side of the Sample Column (Column A). We entered the function in cell B2 as "=RIGHT(A2,3)" and then copied the function in below rows till row 5. The output is as follows:



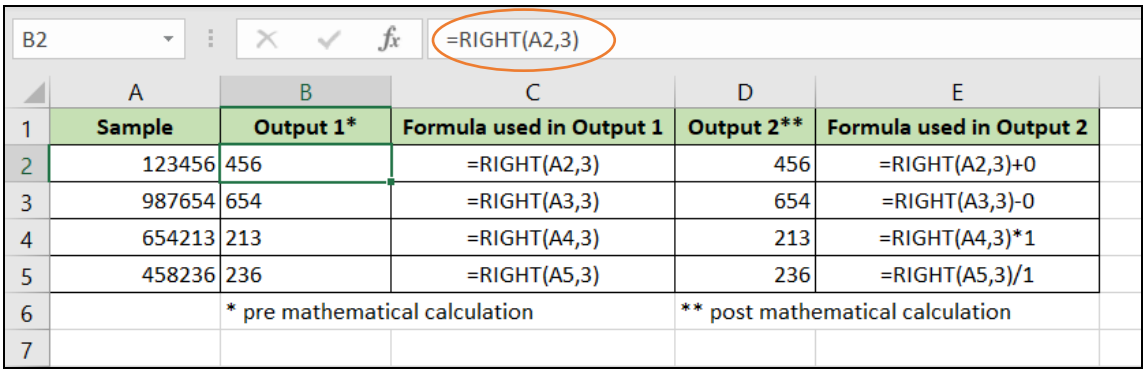
	A	B	C
1	**Sample**	**Output**	**Formula used**
2	Ramesh	esh	=RIGHT(A2,3)
3	UMESH	ESH	=RIGHT(A3,3)
4	amar	mar	=RIGHT(A4,3)
5	123456	456	=RIGHT(A5,3)
6			

Figure 5.7 – RIGHT Function

It can be seen that RIGHT Function has extracted 3 characters from all the cells selected. In case of cell A5, the sample value is a number, but RIGHT Function being the TEXT Function, it has converted number in Text form and that is why the output is left aligned.

Being output is in Text form, no calculation can be performed on it except count. In case, we want to convert numeric value in cell B5, we need to perform **any one** mathematical calculation (i.e. +0, -0, *1, /1 in cell B5 and the final output will be in numerical form.

In the below example, we have numerical values in Sample Column and we want to extract 3 characters from right side. In Output 1 an Output 2 column, we can see the final output (pre and post mathematical calculation).



	A	B	C	D	E
1	**Sample**	**Output 1***	**Formula used in Output 1**	**Output 2****	**Formula used in Output 2**
2	123456	456	=RIGHT(A2,3)	456	=RIGHT(A2,3)+0
3	987654	654	=RIGHT(A3,3)	654	=RIGHT(A3,3)-0
4	654213	213	=RIGHT(A4,3)	213	=RIGHT(A4,3)*1
5	458236	236	=RIGHT(A5,3)	236	=RIGHT(A5,3)/1
6		* pre mathematical calculation		** post mathematical calculation	
7					

Figure 5.8 – RIGHT Function with mathematical calculation

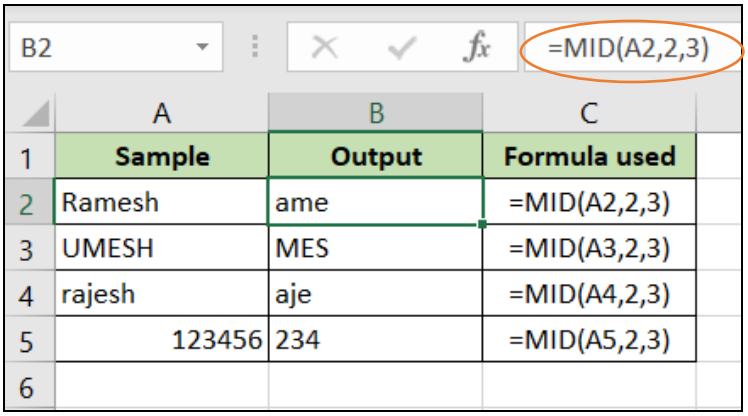
- **MID Function:** Similar to LEFT and RIGHT Function, MID Function also extracts the characters but not from left or right but from middle of a text string, given a starting position and number of characters. The general syntax is as follows:

  = MID(text , start_num, num_chars)

Here the "text " is the text from where we want to extract "number of characters" from the starting number/position specified using MID Function.

Note: The start number and number of characters can be changed as per requirement. In the example, we are extracting 3 characters (num_chars) from 2nd position (start_num).

For example, in the below table, we want to extract 3 characters from 2nd position of the Sample Column (Column A). We entered the function in cell B2 as "=MID(A2,2,3)" and then copied the function in below rows till row 5. The output is as follows:



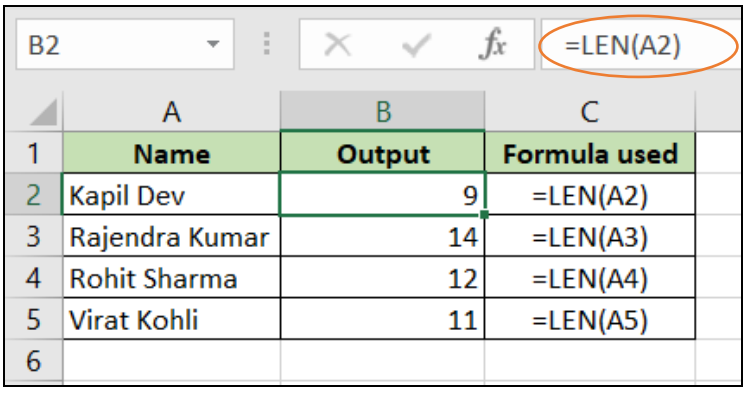
	A	B	C
1	**Sample**	**Output**	**Formula used**
2	Ramesh	ame	=MID(A2,2,3)
3	UMESH	MES	=MID(A3,2,3)
4	rajesh	aje	=MID(A4,2,3)
5	123456	234	=MID(A5,2,3)
6			

Figure 5.9 – MID Function

- **LEN Function:** The LEN is short form of LENGTH. The LEN function returns the number for characters in a text string. We use this function to ascertain the length of characters in a cell. The general syntax is as follows:

  =LEN(text)

  In the below example, we want to know the length (number of characters) in the Name column.



	A	B	C
1	**Name**	**Output**	**Formula used**
2	Kapil Dev	9	=LEN(A2)
3	Rajendra Kumar	14	=LEN(A3)
4	Rohit Sharma	12	=LEN(A4)
5	Virat Kohli	11	=LEN(A5)
6			

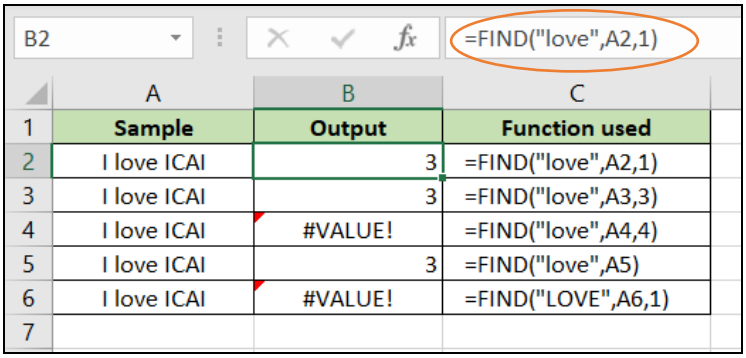
Figure 5.10 – LEN Function

- **FIND Function:** The FIND Function returns the starting position of one text string within another text string. The FIND Function is case sensitive. The general syntax is as follows:

=FIND(find_text, within_text, [start_num])

Here, "find_text" is the text which we want to find in the text ("within_text" – second argument). The third argument i.e. start_num helps to start finding the text from a particular position. If third argument is not given, the function starts looking at the beginning of the larger string.

In the below example, we want to find text "love". The outputs are given below:



	A	B	C
1	**Sample**	**Output**	**Function used**
2	I love ICAI	3	=FIND("love",A2,1)
3	I love ICAI	3	=FIND("love",A3,3)
4	I love ICAI	#VALUE!	=FIND("love",A4,4)
5	I love ICAI	3	=FIND("love",A5)
6	I love ICAI	#VALUE!	=FIND("LOVE",A6,1)
7			

Figure 5.11 – FIND Function

In cell B2, the staring number is 1 and in cell B3, the start number is 3. As the text "love" starts from 3rd position, the output comes as 3.

In case of cell B4, the start number is 4. As the text "love" starts from 3rd position, the function is not able to find the value from 4th position, hence, #value! error came.

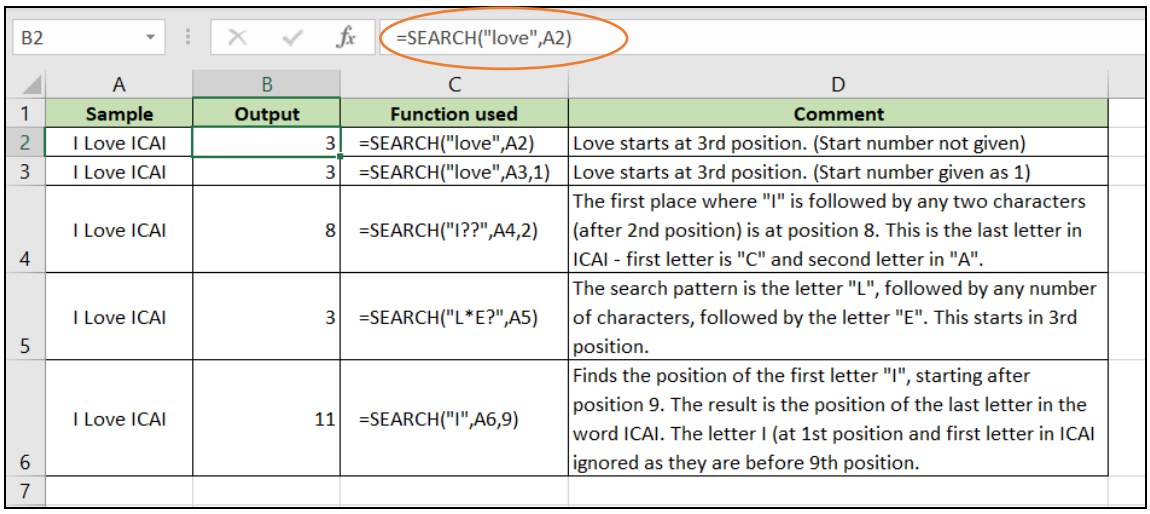
In case of cell B5, the start number is not given and the output comes as 3.

In case of cell B6, the output has #value! error. This error is returned when the text is not found. In this case, we are finding text "LOVE". As the function is case sensitive, the text "LOVE" is not same as "love". Hence, #value! error.

- **SEARCH Function:** The SEARCH Function takes the same arguments as FIND Function. But SEARCH Function is not case sensitive. The general syntax is as follows:

  = SEARCH(find_text, within_text, [start_num])

  The two common wildcards you can use are the asterisk (*) and the question mark (?). An asterisk tells the function to accept any number of characters (including zero characters). A question mark tells the function to accept any single character. It is not uncommon to see more than one question mark together as a wildcard pattern. The examples are as follows:



B2 =SEARCH("love",A2)

	A	B	C	D
1	**Sample**	**Output**	**Function used**	**Comment**
2	I Love ICAI	3	=SEARCH("love",A2)	Love starts at 3rd position. (Start number not given)
3	I Love ICAI	3	=SEARCH("love",A3,1)	Love starts at 3rd position. (Start number given as 1)
4	I Love ICAI	8	=SEARCH("I??",A4,2)	The first place where "I" is followed by any two characters (after 2nd position) is at position 8. This is the last letter in ICAI - first letter is "C" and second letter in "A".
5	I Love ICAI	3	=SEARCH("L*E?",A5)	The search pattern is the letter "L", followed by any number of characters, followed by the letter "E". This starts in 3rd position.
6	I Love ICAI	11	=SEARCH("I",A6,9)	Finds the position of the first letter "I", starting after position 9. The result is the position of the last letter in the word ICAI. The letter I (at 1st position and first letter in ICAI ignored as they are before 9th position.
7				

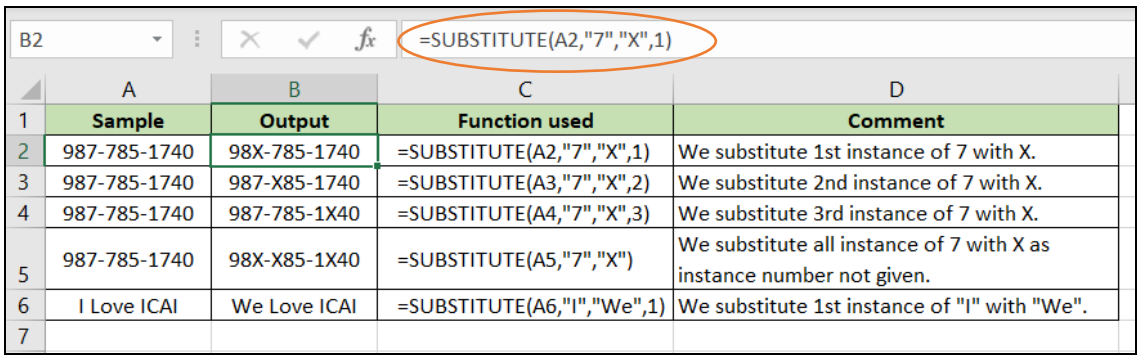
Figure 5.12 – SEARCH Function

- **SUBSTITUTE Function:** This function replaces existing text with new text in a text string. The general syntax is as follows:

  =SUBSTITUTE (text, old_text, new_text, [instance num])

  Here, the "text" is the text in which we want to substitute "old text" with a "new text". It can be done for 1st instance, 2nd instance, 3rd instance or all instances, if instance number is not given.

  The examples are as follows:



B2 =SUBSTITUTE(A2,"7","X",1)

	A	B	C	D
1	**Sample**	**Output**	**Function used**	**Comment**
2	987-785-1740	98X-785-1740	=SUBSTITUTE(A2,"7","X",1)	We substitute 1st instance of 7 with X.
3	987-785-1740	987-X85-1740	=SUBSTITUTE(A3,"7","X",2)	We substitute 2nd instance of 7 with X.
4	987-785-1740	987-785-1X40	=SUBSTITUTE(A4,"7","X",3)	We substitute 3rd instance of 7 with X.
5	987-785-1740	98X-X85-1X40	=SUBSTITUTE(A5,"7","X")	We substitute all instance of 7 with X as instance number not given.
6	I Love ICAI	We Love ICAI	=SUBSTITUTE(A6,"I","We",1)	We substitute 1st instance of "I" with "We".
7				

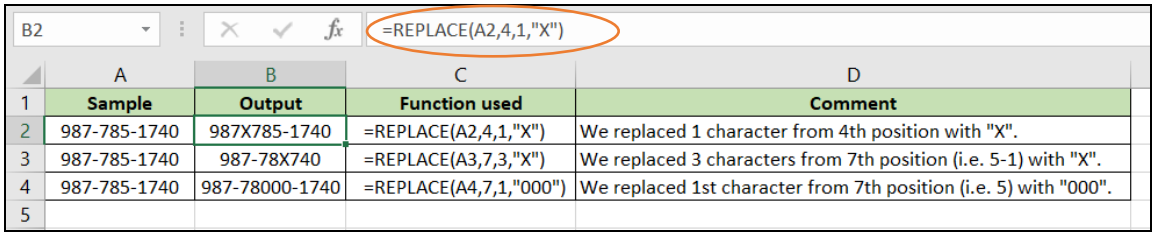
Figure 5.13 – SUBSTITUTE Function

- **REPLACE Function:** This function replaces part of a text string with a different text string. The general syntax is as follows:

  =REPLACE(old_text, start_num, num_chars, new_text)

  Here, the "old_text" is the text where we want to replace specified number of characters from given starting number by new text.

  The examples are as follows:



B2 | =REPLACE(A2,4,1,"X")

	A	B	C	D
1	**Sample**	**Output**	**Function used**	**Comment**
2	987-785-1740	987X785-1740	=REPLACE(A2,4,1,"X")	We replaced 1 character from 4th position with "X".
3	987-785-1740	987-78X740	=REPLACE(A3,7,3,"X")	We replaced 3 characters from 7th position (i.e. 5-1) with "X".
4	987-785-1740	987-78000-1740	=REPLACE(A4,7,1,"000")	We replaced 1st character from 7th position (i.e. 5) with "000".
5				

Figure 5.14 – REPLACE Function

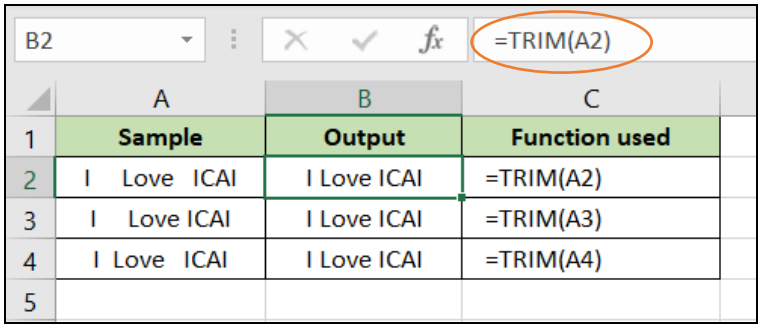
It is to be noted that while writing the function, we need to tell the function that at what position the text starts and how many characters to be replaced. Secondly, the length of text being replaced and the new text may not be the same.

- **TRIM Function:** This function removes all spaces from a text string except for single space between words. The general syntax is as follows:

  =TRIM(text)

  Here, the "text" is the text in which we want to remove extra spaces.

  The examples as follows:



B2 | =TRIM(A2)

	A	B	C
1	**Sample**	**Output**	**Function used**
2	I Love ICAI	I Love ICAI	=TRIM(A2)
3	I Love ICAI	I Love ICAI	=TRIM(A3)
4	I Love ICAI	I Love ICAI	=TRIM(A4)
5			

Figure 5.15 – TRIM Function

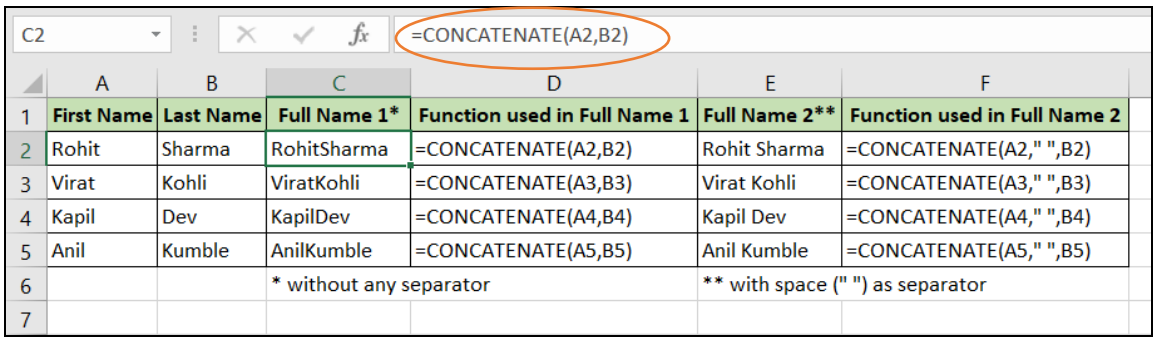
- **CONCATENATE Function and Concatenate Operator "&" :** The CONCATENATE Function joins several text strings into one text string. Similarly "&" (ampersand) can also be used to join several text strings into one text string. The general syntax is as follows:

  = CONCATENATE(text1, [text2], ...)

  = text1 **&** text2

  The CONCATENATE function takes up to 255 arguments. Each argument is a string or a cell reference, and the arguments are separated by commas. The function does not insert anything, such as a space, between the strings. If you need to separate the substrings, as you would with the first name and last name example, you must explicitly insert the separator.

  In the below example, we want to join First Name with Last Name using CONCATENATE Function:

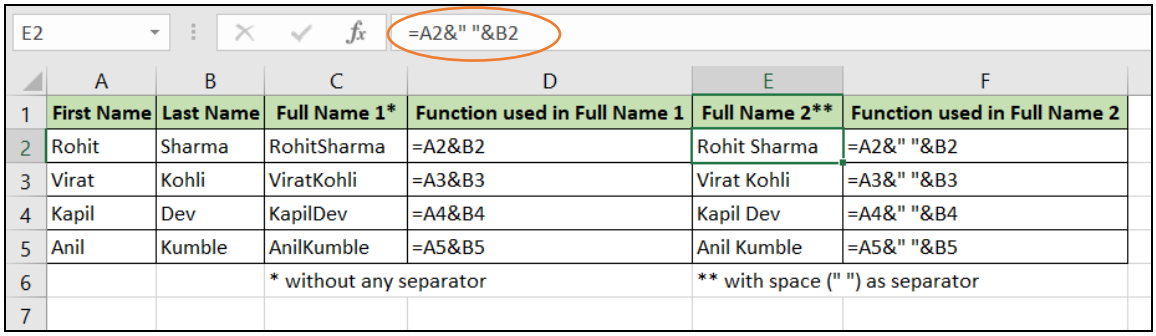


C2 =CONCATENATE(A2,B2)

	A	B	C	D	E	F
1	**First Name**	**Last Name**	**Full Name 1***	**Function used in Full Name 1**	**Full Name 2****	**Function used in Full Name 2**
2	Rohit	Sharma	RohitSharma	=CONCATENATE(A2,B2)	Rohit Sharma	=CONCATENATE(A2," ",B2)
3	Virat	Kohli	ViratKohli	=CONCATENATE(A3,B3)	Virat Kohli	=CONCATENATE(A3," ",B3)
4	Kapil	Dev	KapilDev	=CONCATENATE(A4,B4)	Kapil Dev	=CONCATENATE(A4," ",B4)
5	Anil	Kumble	AnilKumble	=CONCATENATE(A5,B5)	Anil Kumble	=CONCATENATE(A5," ",B5)
6			* without any separator		** with space (" ") as separator	
7						

Figure 5.16 – CONCATENATE Function

In the below example, we want to join First Name with Last Name using "&" (ampersand) operator:



E2 =A2&" "&B2

	A	B	C	D	E	F
1	**First Name**	**Last Name**	**Full Name 1***	**Function used in Full Name 1**	**Full Name 2****	**Function used in Full Name 2**
2	Rohit	Sharma	RohitSharma	=A2&B2	Rohit Sharma	=A2&" "&B2
3	Virat	Kohli	ViratKohli	=A3&B3	Virat Kohli	=A3&" "&B3
4	Kapil	Dev	KapilDev	=A4&B4	Kapil Dev	=A4&" "&B4
5	Anil	Kumble	AnilKumble	=A5&B5	Anil Kumble	=A5&" "&B5
6			* without any separator		** with space (" ") as separator	
7						

(Figure 5.17 – & operator)

5.4 FLASH FILL

Flash Fill is only available in Excel 2013 and later. It is part of AutoFill feature. It automatically fills the data when it senses a pattern in the data. It does not matter whether you are extracting data in next column or several columns away as long as it is withing the same database range.

For example, we can use Flash Fill to separate first and last names from a single column, or combine first and last names from two different columns. It can be used to extract number or text based on a pattern.

In below example, column A contains First Name, column B has Last Name, and we want to fill column C with First and Last names combined. If we establish a pattern by typing the Full Name in column C, Excel's Flash Fill feature will fill in the rest of the names based on the pattern provided by us.

Extract First Name and Last Name from Full Name column:

Steps to be followed for First Name in Column B:

(i) Enter the First Name in cell B2, and press ENTER.

(ii) Start typing the next First Name in cell B3. Excel will sense the pattern we provide, and it shows a preview of the rest of the column filled in with combined text.

(iii) To accept the preview, press ENTER.

Note: Instead of Step ii and iii, we can simply press "Ctrl+E" in cell B3 for Flash Fill the data in below cells.

Similar steps to be followed in column C for extracting Last Name.

The output is as follows:

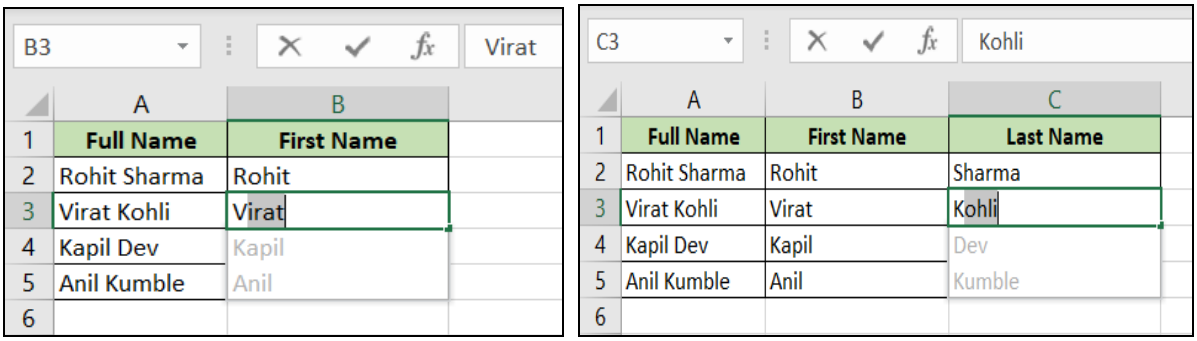


Figure 5.18 – FLASH FILL Function

Concatenate data with Flash Fill

(i) Enter the full name in cell C2, and press ENTER.

(ii) Start typing the next full name in cell C3. Excel will sense the pattern we provide, and it shows a preview of the rest of the column filled in with combined text.

(iii) To accept the preview, press ENTER.

Note: Instead of Step ii and iii, we can simply press "Ctrl+E" in cell C3 for Flash Fill the data in below cells.

The output is as follows:

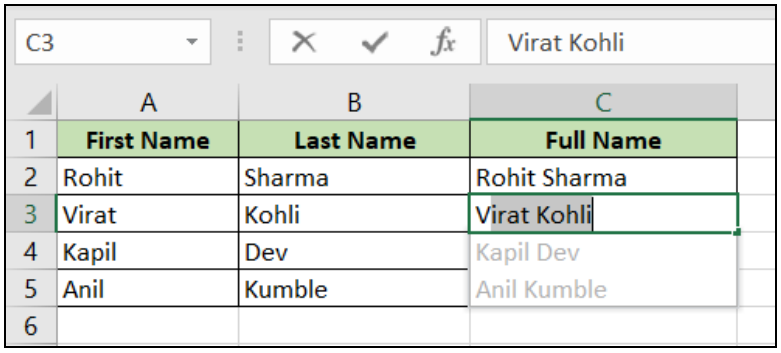


Figure 5.19 – FLASH FILL Function

In case, we want to extract only numbers from PAN (Column A). then we need to type the required number (1234) in cell B2. Then as we type 9.. in cell B3, Excel will sense the pattern we provide, and it shows a preview of the rest of the column filled in with combined text. To accept the preview, press ENTER.

The output is as follows:

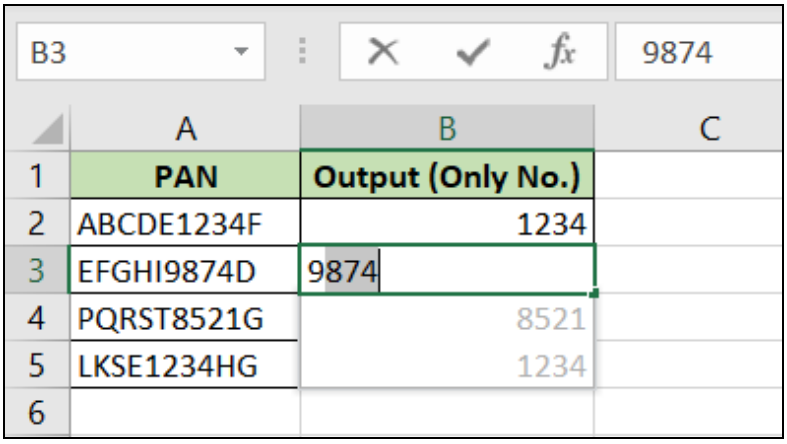


Figure 5.20 – FLASH FILL Function

Note:

- If Flash Fill doesn't generate the preview, it might not be turned on. Then we can go to **Data** > **Flash Fill** to run it manually, or press Ctrl+E.
- To turn Flash Fill on, go to **Tools** > **Options** > **Advanced** > **Editing Options** > check the **Automatically Flash Fill** box.

5.5 ROUNDING FUNCTIONS

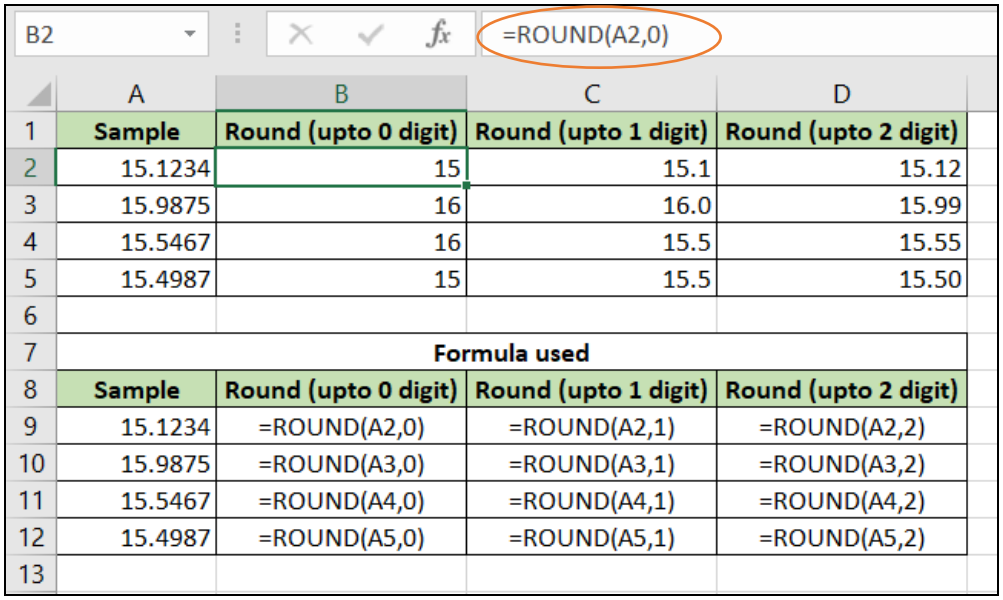
ROUND Function rounds the decimal places of the number by a specified number. There are three Round Functions i.e. ROUND, ROUNDUP and ROUNDDOWN.

- **ROUND Function:** It rounds the number to a specified digit. The general syntax is as follows:

  =ROUND(number, num_digits)

  Here, “number” is the number which will be rounded to specified number of digits (decimal point).

The example for ROUND Function is as follows:



B2 =ROUND(A2,0)

	A	B	C	D
1	**Sample**	**Round (upto 0 digit)**	**Round (upto 1 digit)**	**Round (upto 2 digit)**
2	15.1234	15	15.1	15.12
3	15.9875	16	16.0	15.99
4	15.5467	16	15.5	15.55
5	15.4987	15	15.5	15.50
6				
7	**Formula used**			
8	**Sample**	**Round (upto 0 digit)**	**Round (upto 1 digit)**	**Round (upto 2 digit)**
9	15.1234	=ROUND(A2,0)	=ROUND(A2,1)	=ROUND(A2,2)
10	15.9875	=ROUND(A3,0)	=ROUND(A3,1)	=ROUND(A3,2)
11	15.5467	=ROUND(A4,0)	=ROUND(A4,1)	=ROUND(A4,2)
12	15.4987	=ROUND(A5,0)	=ROUND(A5,1)	=ROUND(A5,2)
13				

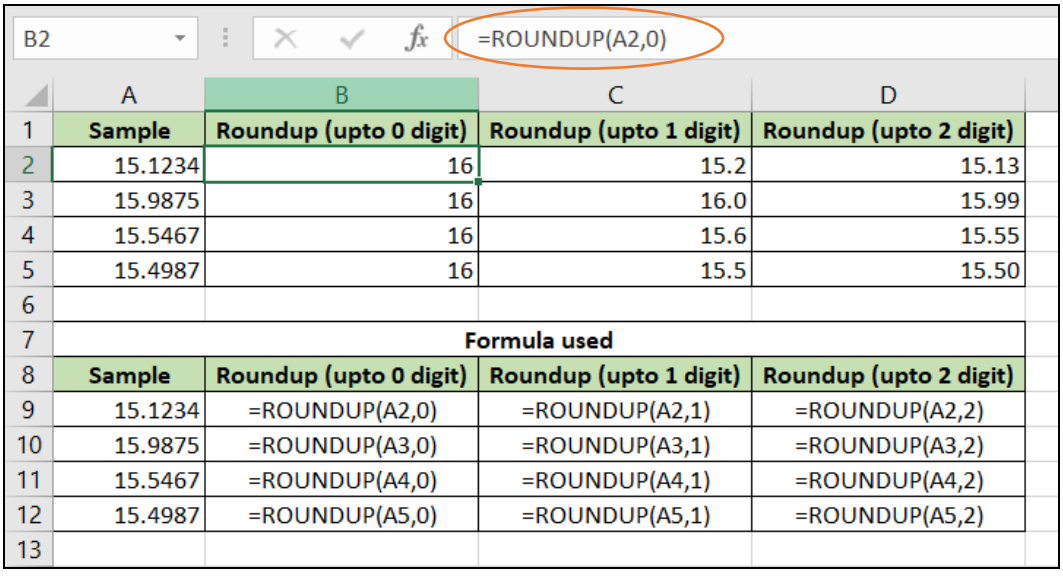
Figure 5.21 – ROUND Function

- **ROUNDUP Function:** The ROUNDUP Function always rounds a number up away from zero (0). The general syntax is as follows:

  =ROUNDUP(number, num_digits)

  Here, "number" is the number which will be rounded up to specified number of digits (decimal point).

  The example for ROUNDUP Function is as follows:



B2 =ROUNDUP(A2,0)

	A	B	C	D
1	**Sample**	**Roundup (upto 0 digit)**	**Roundup (upto 1 digit)**	**Roundup (upto 2 digit)**
2	15.1234	16	15.2	15.13
3	15.9875	16	16.0	15.99
4	15.5467	16	15.6	15.55
5	15.4987	16	15.5	15.50
6				
7	**Formula used**			
8	**Sample**	**Roundup (upto 0 digit)**	**Roundup (upto 1 digit)**	**Roundup (upto 2 digit)**
9	15.1234	=ROUNDUP(A2,0)	=ROUNDUP(A2,1)	=ROUNDUP(A2,2)
10	15.9875	=ROUNDUP(A3,0)	=ROUNDUP(A3,1)	=ROUNDUP(A3,2)
11	15.5467	=ROUNDUP(A4,0)	=ROUNDUP(A4,1)	=ROUNDUP(A4,2)
12	15.4987	=ROUNDUP(A5,0)	=ROUNDUP(A5,1)	=ROUNDUP(A5,2)
13				

Figure 5.22 – ROUNDUP Function

It can be observed that the number after the decimal point is increased by one irrespective of the fact that the next digit is less than 5.

- **ROUNDDOWN Function:** The ROUNDDOWN Function always rounds a number down towards zero (0). The general syntax is as follows:

  =ROUNDDOWN(number, num_digits)

  Here, "number" is the number which will be rounded down to specified number of digits (decimal point).

The example for ROUNDDOWN Function is as follows:

B2 | =ROUNDDOWN(A2,0)

	A	B	C	D
1	**Sample**	**Rounddown (upto 0 digit)**	**Rounddown (upto 1 digit)**	**Rounddown (upto 2 digit)**
2	15.1234	15	15.1	15.12
3	15.9875	15	15.9	15.98
4	15.5467	15	15.5	15.54
5	15.4987	15	15.4	15.49
6				
7	**Formula used**			
8	**Sample**	**Rounddown (upto 0 digit)**	**Rounddown (upto 1 digit)**	**Rounddown (upto 2 digit)**
9	15.1234	=ROUNDDOWN(A2,0)	=ROUNDDOWN(A2,1)	=ROUNDDOWN(A2,2)
10	15.9875	=ROUNDDOWN(A3,0)	=ROUNDDOWN(A3,1)	=ROUNDDOWN(A3,2)
11	15.5467	=ROUNDDOWN(A4,0)	=ROUNDDOWN(A4,1)	=ROUNDDOWN(A4,2)
12	15.4987	=ROUNDDOWN(A5,0)	=ROUNDDOWN(A5,1)	=ROUNDDOWN(A5,2)
13				

Figure 5.23 – ROUNDDOWN Function

5.6 DATE AND TIME FUNCTIONS

In Excel, dates and times are not treated as text entries. Any entry with a format that looks like one the date and time number format being used by the program is converted into a serial number.

For dates, the serial number shows the number of days that have gone by since the start of the 20th century. So that January 1, 1900 uses a serial number of 1. January 2, 1900 uses a serial number 2 and so on.

For times, this serial number is a fraction that represents the number of hours, minutes and seconds that have gone since midnight. This begins with serial number 0.00000000, so that 12:00:00 pm is serial number 0.50000000, while 10:00:00 pm is 0.9166666667, and so on.

There are several DATE and TIME Function. Some of them are as follows:

- **DATE Function:** We can use DATE Function to create a complete date from separate year, month and day information. The general syntax is as follows:

  =DATE(year, month, day)

It can be useful in a scenario where dates don't appear in a proper form. It is possible that in data, one column contains information for year (in either 2 digit or in 4 digit), another column contains information for months (values between 1 to 12) and another column contain information for days (values between 1 to 31). The DATE Function combines individual year, month and day values into a single date.

This can be seen in below examples:

D2 =DATE(A2,B2,C2)

	A	B	C	D	E
1	**Year**	**Month**	**Day**	**Date**	**Function used**
2	1997	5	9	09-05-1997	=DATE(A2,B2,C2)
3	1999	6	27	27-06-1999	=DATE(A3,B3,C3)
4	2000	3	30	30-03-2000	=DATE(A4,B4,C4)
5	2001	10	15	15-10-2001	=DATE(A5,B5,C5)
6	2001	9	21	21-09-2001	=DATE(A6,B6,C6)
7	2007	7	11	11-07-2007	=DATE(A7,B7,C7)
8	2007	-3	11	11-09-2006	=DATE(A8,B8,C8)
9	2009	15	28	28-03-2010	=DATE(A9,B9,C9)
10	2016	3	**-7**	22-02-2016	=DATE(A10,B10,C10)
11	2018	11	**35**	05-12-2018	=DATE(A11,B11,C11)
12					

Figure 5.24 – DATE Function

DATE provides some extra flexibility with the month number. Negative month numbers are subtracted from the specified year. For example, the function "=DATE(2007, -3, 11)" returns the date "11-09-2006", because September 2006 is 3 months before the first month of 2007.

Similarly month numbers greater than 12 works in the same way. Function "=DATE(2009, 15, 28) returns "28-03-2010", because March 2010 is 15 months after the first month of 2009.

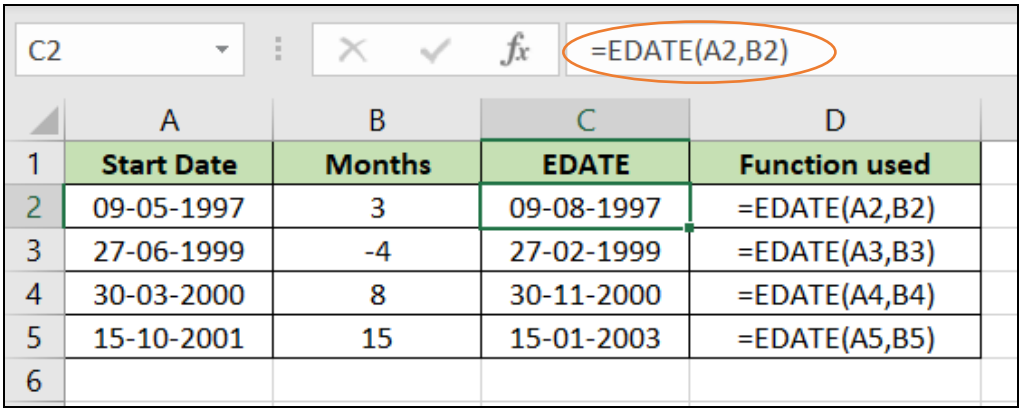
Day numbers work the same way. Negative day numbers are subtracted from the first of the specified month, and numbers that are greater than the last day of the specified month wrap into later months. Thus, =DATE(2016, 3, -7) returns "22-02-2016". Likewise, "=DATE(2018, 11, 35)" returns "05-12-2018".

- **EDATE Function:** This function returns a serial number of the date that is the indicated number of months before or after the start date in the function. The general syntax is as follows:

  =EDATE(start_date, months)

  Here, the start date is the date in which we want to add or reduce number of months.

The example is as follows:



	A	B	C	D
1	**Start Date**	**Months**	**EDATE**	**Function used**
2	09-05-1997	3	09-08-1997	=EDATE(A2,B2)
3	27-06-1999	-4	27-02-1999	=EDATE(A3,B3)
4	30-03-2000	8	30-11-2000	=EDATE(A4,B4)
5	15-10-2001	15	15-01-2003	=EDATE(A5,B5)
6				

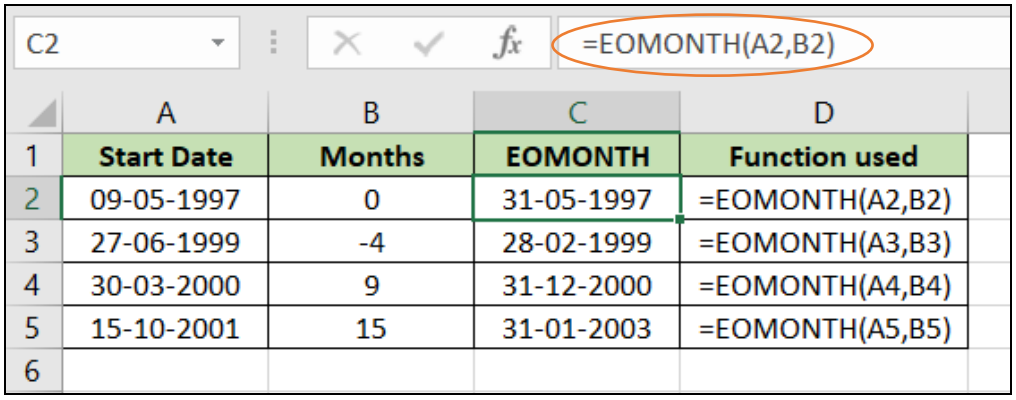
Figure 5.25 – EDATE Function

- **EOMONTH Function:** The EOMONTH function returns the serial number of the last day of the month before or after a specified number of months. The general syntax is follows:

  =EOMONTH(start_date, months)

  Here, the start date is the date for which we want last date of the month after adding or reducing number of months.

  The example is as follows:



	A	B	C	D
1	**Start Date**	**Months**	**EOMONTH**	**Function used**
2	09-05-1997	0	31-05-1997	=EOMONTH(A2,B2)
3	27-06-1999	-4	28-02-1999	=EOMONTH(A3,B3)
4	30-03-2000	9	31-12-2000	=EOMONTH(A4,B4)
5	15-10-2001	15	31-01-2003	=EOMONTH(A5,B5)
6				

(Figure 5.26 – EOMONTH Function)

- **DATEDIF Function:** This function calculates the number of days, months, or years between two dates. This is an undocumented function; that is, we won't see it in the Insert Function dialog box, and we cannot find it in the Excel Help system. The only thing we can do is to remember how to enter it. The general syntax is as follows:

  =DATEDIF(start date, end date, interval)

  Here, start date is the first date and end date is the last date, which should be greater than start date, else we will get error. The interval can be in any one of the following values:

Value	Meaning	Comment
"d"	Days	The count of inclusive days from the start date through the end date.
"m"	Months	The count of complete months between start date and end date. Only those months that fully occur between the dates are counted. For example, if the first date starts after the first of the month, that first month is not included in the count. For the end date, even when it is the last day of the month, that month is not counted.
"y"	Years	The count of complete years between the start date and end date. Only those years that fully occur between the dates are counted. For example, if the first date starts later than January 1, that first year is not included in the count. For the end date, even when it is December 31, that year is not counted.
"yd"	Days excluding years	The count of inclusive days from the start date through the end date, but as if the two dates are in the same year. The year is ignored.
"ym"	Months excluding years	The count of complete months between the dates, but as if the two dates are in the same year. The year is ignored.
"md"	Days excluding months and years	The count of inclusive days from the start date through the end date, but as if the two dates are in the same month and year. The month and year are ignored.

Figure 5.27 – DATEDIF Function - Values

The examples are given below:

C2 =DATEDIF(A2,B2,"d")

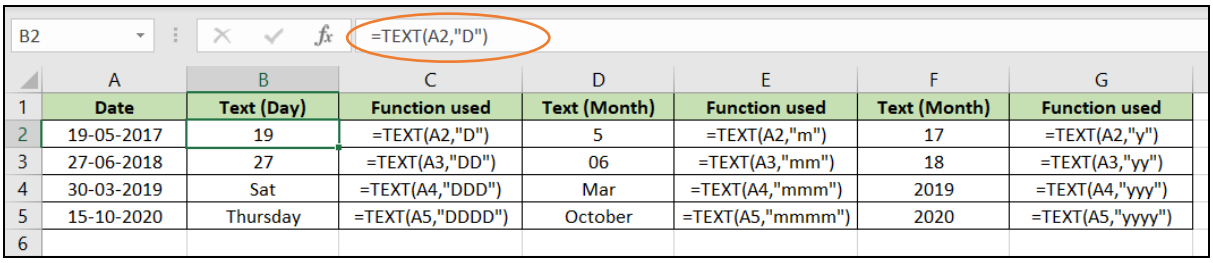
	A	B	C	D	E	F	G	H
1	Start Date	End Date	"d"	"m"	"y"	"yd"	"ym"	"md"
2	09-05-2017	29-05-2017	20	0	0	20	0	20
3	27-06-2019	15-07-2019	18	0	0	18	0	18
4	11-03-2019	13-03-2020	368	12	1	2	0	2
5	15-10-2020	25-12-2020	71	2	0	71	2	10
6								
7								
8			Formula used					
9	Start Date	End Date	"d"	"m"	"y"	"yd"	"ym"	"md"
10	09-05-2017	29-05-2017	=DATEDIF(A2,B2,"d")	=DATEDIF(A2,B2,"m")	=DATEDIF(A2,B2,"Y")	=DATEDIF(A2,B2,"yd")	=DATEDIF(A2,B2,"ym")	=DATEDIF(A2,B2,"md")
11	27-06-2019	15-07-2019	=DATEDIF(A3,B3,"d")	=DATEDIF(A3,B3,"m")	=DATEDIF(A3,B3,"Y")	=DATEDIF(A3,B3,"yd")	=DATEDIF(A3,B3,"ym")	=DATEDIF(A3,B3,"md")
12	01-03-2019	13-03-2020	=DATEDIF(A4,B4,"d")	=DATEDIF(A4,B4,"m")	=DATEDIF(A4,B4,"Y")	=DATEDIF(A4,B4,"yd")	=DATEDIF(A4,B4,"ym")	=DATEDIF(A4,B4,"md")
13	15-10-2020	25-12-2020	=DATEDIF(A5,B5,"d")	=DATEDIF(A5,B5,"m")	=DATEDIF(A5,B5,"Y")	=DATEDIF(A5,B5,"yd")	=DATEDIF(A5,B5,"ym")	=DATEDIF(A5,B5,"md")
14								

(Figure 5.28 – DATEDIF Function)

- **TEXT Function:** The TEXT Function converts a value to text in a specific number format. It is useful in extraction of Day, Month from a give date. The general syntax is as follows:

=TEXT(value, format_text)

The example is as follows:



	A	B	C	D	E	F	G
1	**Date**	**Text (Day)**	**Function used**	**Text (Month)**	**Function used**	**Text (Month)**	**Function used**
2	19-05-2017	19	=TEXT(A2,"D")	5	=TEXT(A2,"m")	17	=TEXT(A2,"y")
3	27-06-2018	27	=TEXT(A3,"DD")	06	=TEXT(A3,"mm")	18	=TEXT(A3,"yy")
4	30-03-2019	Sat	=TEXT(A4,"DDD")	Mar	=TEXT(A4,"mmm")	2019	=TEXT(A4,"yyy")
5	15-10-2020	Thursday	=TEXT(A5,"DDDD")	October	=TEXT(A5,"mmmm")	2020	=TEXT(A5,"yyyy")
6							

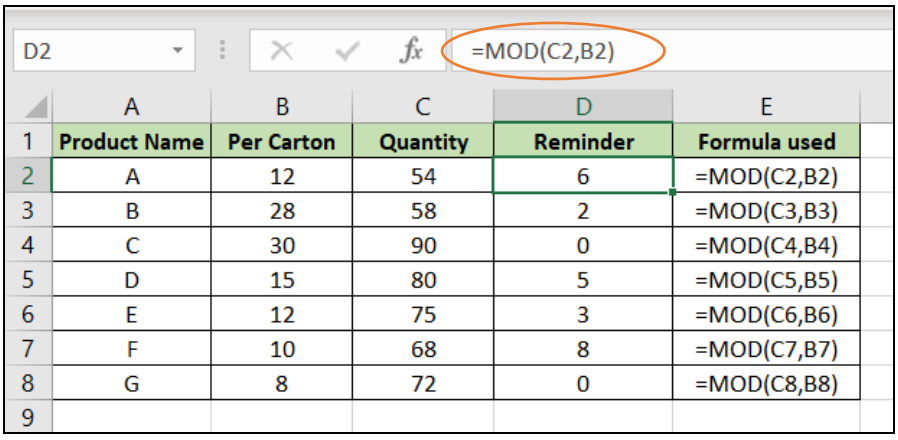
Figure 5.29 – TEXT Function

- **MOD Function:** This function returns the reminder after a number is divided by a divisor. Normally, if we divide a number by some divisor and if the dividend is not completely divisible by the divisor, the quotient is expressed in decimal form. However, sometimes we are interested in obtaining the remainder separately. In such a case, we can make use of Mod function.

  The general syntax is as follows:

  =MOD(number, divisor)

  In the below example, we have various products, which is packed in certain quantity in each carton. We also have a quantity column for each product. We want to calculate reminder value (left over quantity) after packing in carton. The example is as follows:



	A	B	C	D	E
1	**Product Name**	**Per Carton**	**Quantity**	**Reminder**	**Formula used**
2	A	12	54	6	=MOD(C2,B2)
3	B	28	58	2	=MOD(C3,B3)
4	C	30	90	0	=MOD(C4,B4)
5	D	15	80	5	=MOD(C5,B5)
6	E	12	75	3	=MOD(C6,B6)
7	F	10	68	8	=MOD(C7,B7)
8	G	8	72	0	=MOD(C8,B8)
9					

Figure 5.30 – MOD Function

5.7 LOGICAL FUNCTIONS

There are 4 main logical functions in Excel i.e. IF, AND, OR and NOT. However, there are other sub-logical functions derived from these 4 main logical functions, which are XOR, IFERROR, IFNA, NESTEDIF and ISBLANK.

- **IF Function:** The IF Function is one of the most important Logical Excel Function. It helps to check a

condition and give the output either TRUE or FALSE according to the condition. The general syntax is as follows:

= IF(logical_test, [value_if_true], [value_if_false])

Here, the "logical_test" is the condition which will be tested and output will return according to the logic. If logic is TRUE, it will give one value and if logic is FALSE, it will return another value.

In the below example, we are trying to use simple IF function. In the example, we have 2 columns i.e. Name and Amount column. We want to check whether the amount is greater than 9000. We want 4 different output based on one single condition as follows:

C2 =IF(B2>9000,B2,0)

	A	B	C	D	E	F
1	**Name**	**Amount**	**Output 1**	**Output 2**	**Output 3**	**Output 4**
2	A	10,000	10,000	A	A	1,00,000
3	B	20,000	20,000	B	B	2,00,000
4	C	5,000	-	0	NA	1,000
5	D	9,000	-	0	NA	1,800
6						
7			**Formula used**			
8	**Name**	**Amount**	**Output 1**	**Output 2**	**Output 3**	**Output 4**
9	A	10,000	=IF(B2>9000,B2,0)	=IF(B2>9000,A2,0)	=IF(B2>9000,A2,"NA")	=IF(B2>9000,B2*10,B2/5)
10	B	20,000	=IF(B3>9000,B3,0)	=IF(B3>9000,A3,0)	=IF(B3>9000,A3,"NA")	=IF(B3>9000,B3*10,B3/5)
11	C	5,000	=IF(B4>9000,B4,0)	=IF(B4>9000,A4,0)	=IF(B4>9000,A4,"NA")	=IF(B4>9000,B4*10,B4/5)
12	D	9,000	=IF(B5>9000,B5,0)	=IF(B5>9000,A5,0)	=IF(B5>9000,A5,"NA")	=IF(B5>9000,B5*10,B5/5)
13						

Figure 5.31 – IF Function

Condition in Output 1: If Amount > 9000, return the same amount if TRUE and return 0 if FALSE.

Condition in Output 2: If Amount > 9000, return the Name of the person if TRUE and return 0 if FALSE.

Condition in Output 3: If Amount > 9000, return the Name of the person if TRUE and return "NA" if FALSE.

Condition in Output 4: If Amount > 9000, if TRUE, multiply Amount by 10 and if FALSE, divide amount by 5.

- **NESTED IF Function:** We have seen IF Function, which helps in testing one condition and give output. In case, we have multiple conditions and output based on those conditions, then we have to use nested IF Function. In Nested IF Function, we use multiple IF within IF based on conditions. The general syntax is as follows:

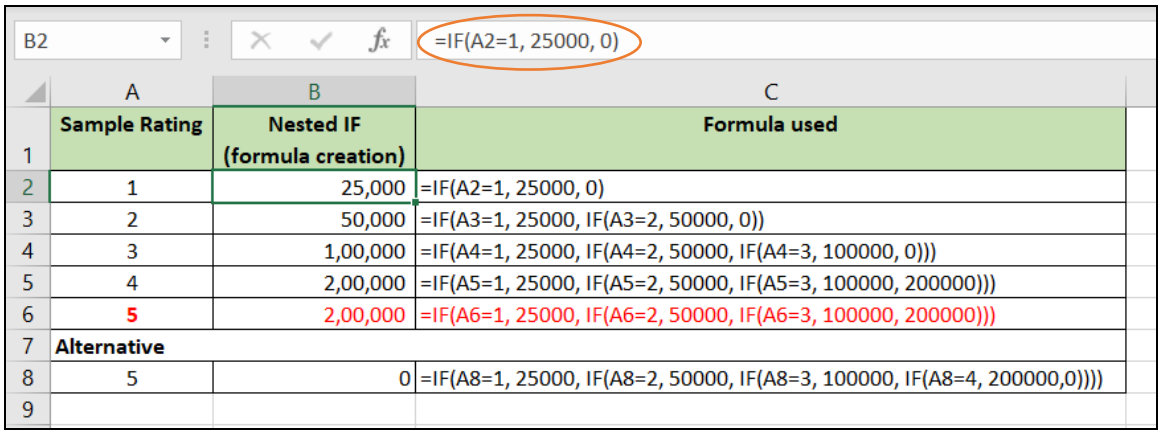
  =IF(logical_test, [value_if_true], IF(logical_test, [value_if_true], [value_if_false]))

  Example 1: In the below example, we have a list of Employees and their rating. The company will give bonus to employees based on Rating to each employee. The criteria for bonus is as follows:

  - If Rating = 1, then ₹ 25,000/- Bonus

- If Rating = 2, then ₹ 50,000/- Bonus
- If Rating = 3, then ₹ 1,00,000/- Bonus
- If Rating = 4, then ₹ 2,00,000/- Bonus

We need to use multiple IF function to get the output. The formula build up is as follows:



B2 =IF(A2=1, 25000, 0)

	A	B	C
1	**Sample Rating**	**Nested IF (formula creation)**	**Formula used**
2	1	25,000	=IF(A2=1, 25000, 0)
3	2	50,000	=IF(A3=1, 25000, IF(A3=2, 50000, 0))
4	3	1,00,000	=IF(A4=1, 25000, IF(A4=2, 50000, IF(A4=3, 100000, 0)))
5	4	2,00,000	=IF(A5=1, 25000, IF(A5=2, 50000, IF(A5=3, 100000, 200000)))
6	**5**	2,00,000	=IF(A6=1, 25000, IF(A6=2, 50000, IF(A6=3, 100000, 200000)))
7	**Alternative**		
8	5	0	=IF(A8=1, 25000, IF(A8=2, 50000, IF(A8=3, 100000, IF(A8=4, 200000,0))))
9			

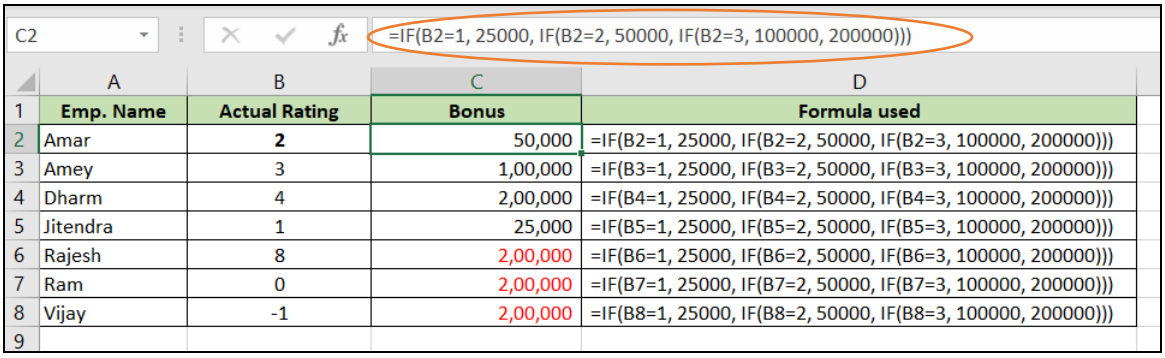
Figure 5.32 – Nested IF formula build up

In the above formula, we can see that the first 3 formulas are not complete as we have 4 conditions to check from each employee. We have to use the formula, which covers all the conditions to be tested i.e. the 4th one. But as the rating changes to 5, still it gives output as 200000 as it is a value if false. Hence, an alternative is given for any rating other than 1, 2, 3 or 4. In the final 5th formula, we have used another IF for 4th condition, and for any other value, it will return as 0.

Formula build up:

- **IF rating = 1, then output will be 25000, else 0.**
- But we want to cover all conditions. Hence time to add another condition.
- If rating = 1, then output will be 25000, for else add another IF condition.
- **IF rating = 1, then output will be 25000, IF rating = 2, then output will be 50000, else 0.**
- Still 2 conditions to be added in formula.
- **IF rating = 1, then output will be 25000, IF rating = 2, then output will be 50000, IF rating = 3, then output will be 100000, else 0.**
- Still 1 condition to be added in formula.
- **IF rating = 1, then output will be 25000, IF rating = 2, then output will be 50000, IF rating = 3, then output will be 100000, else 200000.**
- In last step, we want output 200000, if it is neither 1, nor 2, nor 3. But it has a risk that if the rating is any value other than 1, 2 and 3, then it will return value as 200000, which may not be correct in all scenario. It can be seen in case of Rating =5 in above table.

- The output as per above formula is as follows:



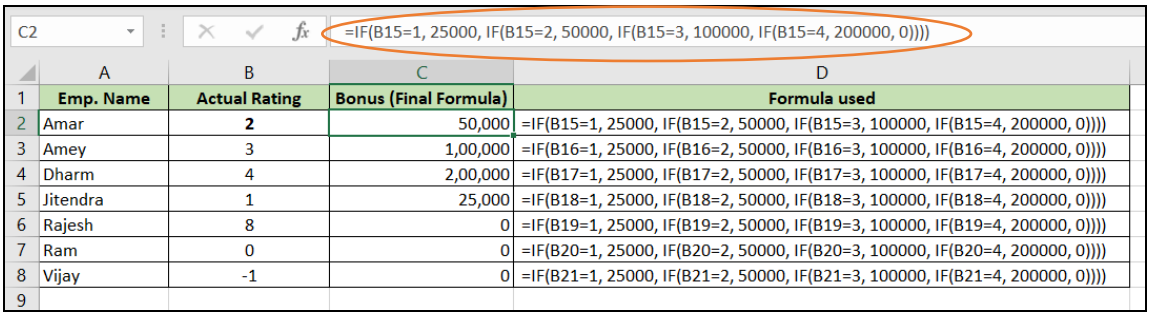
C2 =IF(B2=1, 25000, IF(B2=2, 50000, IF(B2=3, 100000, 200000)))

	A	B	C	D
1	Emp. Name	Actual Rating	Bonus	Formula used
2	Amar	2	50,000	=IF(B2=1, 25000, IF(B2=2, 50000, IF(B2=3, 100000, 200000)))
3	Amey	3	1,00,000	=IF(B3=1, 25000, IF(B3=2, 50000, IF(B3=3, 100000, 200000)))
4	Dharm	4	2,00,000	=IF(B4=1, 25000, IF(B4=2, 50000, IF(B4=3, 100000, 200000)))
5	Jitendra	1	25,000	=IF(B5=1, 25000, IF(B5=2, 50000, IF(B5=3, 100000, 200000)))
6	Rajesh	8	2,00,000	=IF(B6=1, 25000, IF(B6=2, 50000, IF(B6=3, 100000, 200000)))
7	Ram	0	2,00,000	=IF(B7=1, 25000, IF(B7=2, 50000, IF(B7=3, 100000, 200000)))
8	Vijay	-1	2,00,000	=IF(B8=1, 25000, IF(B8=2, 50000, IF(B8=3, 100000, 200000)))
9				

(Figure 5.33 – Nested IF formula with wrong value)

- It can be seen that above formula is returning 200000 (highlighted in red colour manually for reference), which is wrong as it is not as per conditions mentioned above.
- Hence, we need add another IF for 4th condition. The final formula will be:
- **IF rating = 1, then output will be 25000, IF rating = 2, then output will be 50000, IF rating = 3, then output will be 100000, IF rating =4, then output will be 200000, else 0.**

The output as per final NESTED IF function is as follows:



C2 =IF(B15=1, 25000, IF(B15=2, 50000, IF(B15=3, 100000, IF(B15=4, 200000, 0))))

	A	B	C	D
1	Emp. Name	Actual Rating	Bonus (Final Formula)	Formula used
2	Amar	2	50,000	=IF(B15=1, 25000, IF(B15=2, 50000, IF(B15=3, 100000, IF(B15=4, 200000, 0))))
3	Amey	3	1,00,000	=IF(B16=1, 25000, IF(B16=2, 50000, IF(B16=3, 100000, IF(B16=4, 200000, 0))))
4	Dharm	4	2,00,000	=IF(B17=1, 25000, IF(B17=2, 50000, IF(B17=3, 100000, IF(B17=4, 200000, 0))))
5	Jitendra	1	25,000	=IF(B18=1, 25000, IF(B18=2, 50000, IF(B18=3, 100000, IF(B18=4, 200000, 0))))
6	Rajesh	8	0	=IF(B19=1, 25000, IF(B19=2, 50000, IF(B19=3, 100000, IF(B19=4, 200000, 0))))
7	Ram	0	0	=IF(B20=1, 25000, IF(B20=2, 50000, IF(B20=3, 100000, IF(B20=4, 200000, 0))))
8	Vijay	-1	0	=IF(B21=1, 25000, IF(B21=2, 50000, IF(B21=3, 100000, IF(B21=4, 200000, 0))))
9				

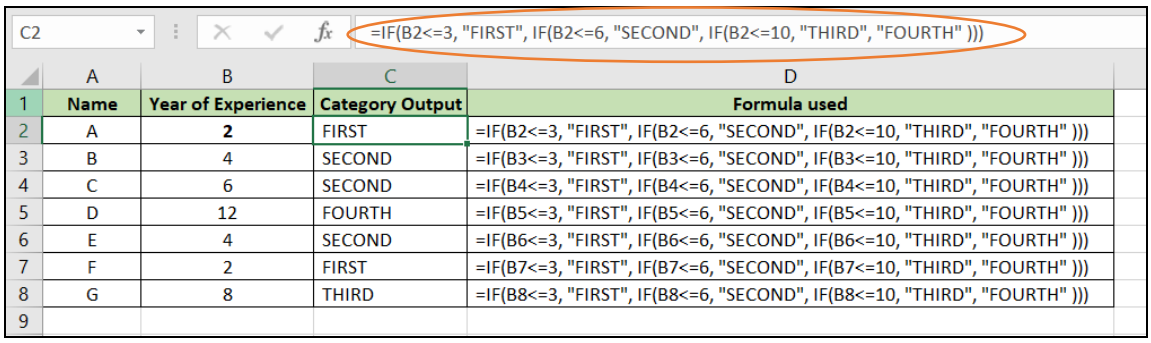
(Figure 5.34 – Nested IF formula with correct output)

It can be seen that we are getting 0 for all the ratings which is other than 1, 2, 3 or 4.

Example 2: In this example, we have a list of Names and their years of experience. We have to categories the Names as per following conditions:

- If years of experience is > 0 and <= 3, then FIRST category.
- If years of experience is > 3 and <= 6, then SECOND category.
- If years of experience is > 6 and <= 10, then THIRD category.
- If years of experience is > 10, then FOURTH category.

The output (as per final formula discussed above) is as follows:



C2 =IF(B2<=3, "FIRST", IF(B2<=6, "SECOND", IF(B2<=10, "THIRD", "FOURTH")))

	A	B	C	D
1	Name	Year of Experience	Category Output	Formula used
2	A	2	FIRST	=IF(B2<=3, "FIRST", IF(B2<=6, "SECOND", IF(B2<=10, "THIRD", "FOURTH")))
3	B	4	SECOND	=IF(B3<=3, "FIRST", IF(B3<=6, "SECOND", IF(B3<=10, "THIRD", "FOURTH")))
4	C	6	SECOND	=IF(B4<=3, "FIRST", IF(B4<=6, "SECOND", IF(B4<=10, "THIRD", "FOURTH")))
5	D	12	FOURTH	=IF(B5<=3, "FIRST", IF(B5<=6, "SECOND", IF(B5<=10, "THIRD", "FOURTH")))
6	E	4	SECOND	=IF(B6<=3, "FIRST", IF(B6<=6, "SECOND", IF(B6<=10, "THIRD", "FOURTH")))
7	F	2	FIRST	=IF(B7<=3, "FIRST", IF(B7<=6, "SECOND", IF(B7<=10, "THIRD", "FOURTH")))
8	G	8	THIRD	=IF(B8<=3, "FIRST", IF(B8<=6, "SECOND", IF(B8<=10, "THIRD", "FOURTH")))
9				

Figure 5.35 – Nested IF formula in case of range scenario

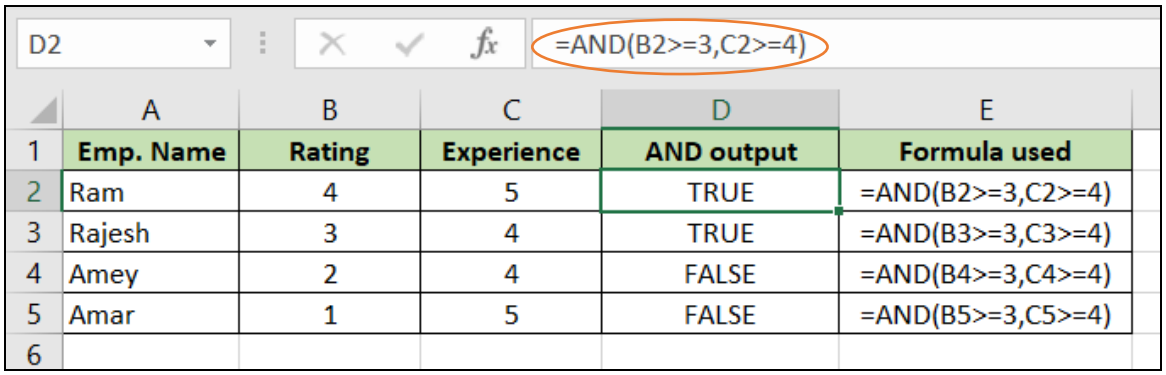
Note: The number of IF functions to be used will depend on the number of conditions to be tested as well as the nature of question. In the first example, we used 4 IF to avoid bonus for any rating other than 1, 2, 3 or 4. In second example, we are using 3 IF as the "else" value will cover all the years of experience > 10.

- **AND Function:** The AND Function is used in a scenario where we need to check more than one condition. If all the logical arguments are valid, it returns value as "True" else it returns value as "False". The general syntax is as follows:

  =AND(logical1, [logical2],…)

  Here, the logical1, logical2 are the conditions which needs to be checked.

  In the example, we have a list of Employees and their Rating and Experience. We want to test 2 condition for each employee. First condition is to test whether Rating is >=3. Second condition is to test whether Experience is >=4. The output after using AND Function is as follows:



D2 =AND(B2>=3,C2>=4)

	A	B	C	D	E
1	Emp. Name	Rating	Experience	AND output	Formula used
2	Ram	4	5	TRUE	=AND(B2>=3,C2>=4)
3	Rajesh	3	4	TRUE	=AND(B3>=3,C3>=4)
4	Amey	2	4	FALSE	=AND(B4>=3,C4>=4)
5	Amar	1	5	FALSE	=AND(B5>=3,C5>=4)
6					

Figure 5.36 – AND Function

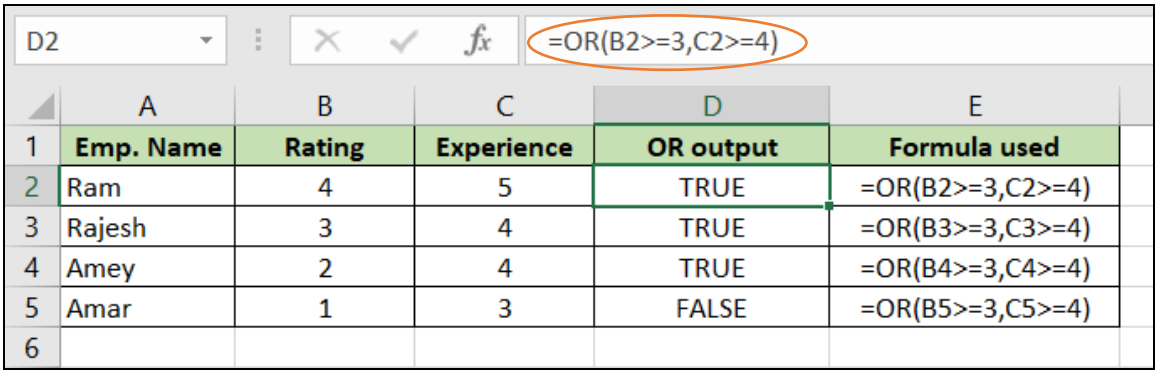
In the first 2 cases, all the conditions are fulfilled. Hence, output is TRUE. In 3rd and 4th case, as all conditions are not fulfilled. Hence, output is FALSE.

- **OR Function:** The OR Function is used in a scenario where we need to check more than one condition. Even if anyone logical argument is valid, it returns value as "True". It will give FALSE with all the arguments are invalid. The general syntax is as follows:

=OR(logical1, [logical2],…)

Here, the logical1, logical2 are the conditions which needs to be checked.

In the example, we have a list of Employees and their Rating and Experience. We want to test 2 condition for each employee. First condition is to test whether Rating is >=3. Second condition is to test whether Experience is >=4. The output after using OR Function is as follows:



	A	B	C	D	E
1	**Emp. Name**	**Rating**	**Experience**	**OR output**	**Formula used**
2	Ram	4	5	TRUE	=OR(B2>=3,C2>=4)
3	Rajesh	3	4	TRUE	=OR(B3>=3,C3>=4)
4	Amey	2	4	TRUE	=OR(B4>=3,C4>=4)
5	Amar	1	3	FALSE	=OR(B5>=3,C5>=4)
6					

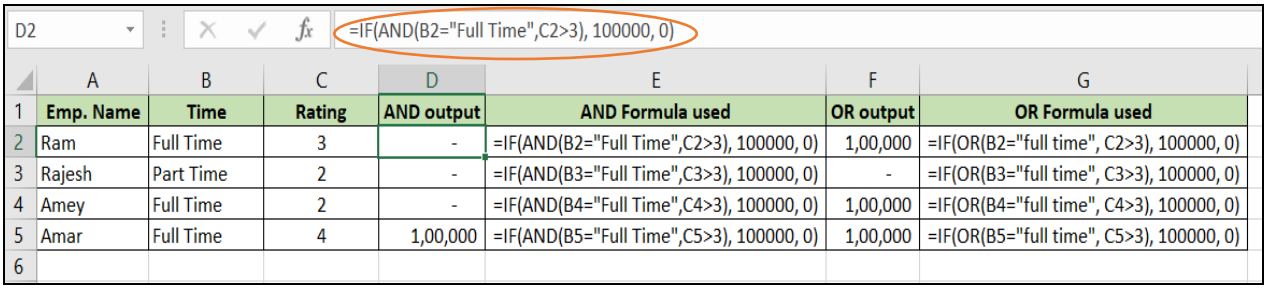
Figure 5.37 – OR Function

In first 3 cases, all or anyone conditions are fulfilled. Hence, output is TRUE. In 4th case, as all conditions are not fulfilled. Hence, output is FALSE.

In another example, we are going to test AND and OR Function alongwith IF Function. We have a list of employees and their working time and Rating.

- AND condition:- If employee is working FULL TIME and Rating is >3, then employee will get bonus of ₹ 100000, else No Bonus i.e. 0 or "-".
- OR condition:- If employee is either working FULL TIME or Rating is >3, then employee will get bonus of ₹ 100000, else No Bonus i.e. 0 or "-".

The output in both the scenario is as follows:



	A	B	C	D	E	F	G
1	**Emp. Name**	**Time**	**Rating**	**AND output**	**AND Formula used**	**OR output**	**OR Formula used**
2	Ram	Full Time	3	-	=IF(AND(B2="Full Time",C2>3), 100000, 0)	1,00,000	=IF(OR(B2="full time", C2>3), 100000, 0)
3	Rajesh	Part Time	2	-	=IF(AND(B3="Full Time",C3>3), 100000, 0)	-	=IF(OR(B3="full time", C3>3), 100000, 0)
4	Amey	Full Time	2	-	=IF(AND(B4="Full Time",C4>3), 100000, 0)	1,00,000	=IF(OR(B4="full time", C4>3), 100000, 0)
5	Amar	Full Time	4	1,00,000	=IF(AND(B5="Full Time",C5>3), 100000, 0)	1,00,000	=IF(OR(B5="full time", C5>3), 100000, 0)
6							

(Figure 5.38 – IF, AND and OR Function together)

In AND scenario, all the conditions are fulfilled only in 4th case. Hence, only Amar will get bonus of ₹ 100000/-

In OR scenario, except in case of Rajesh, all other employees are fulfilling anyone or all the conditions. Hence, they will get bonus of ₹ 100000/-.

It is to be noted that this is not case sensitive function. Hence, while testing we can type anyone of the following - FULL TIME, Full Time, full time.

- **XOR Function:** XOR Function will return TRUE only if majority condition is FALSE. XOR returns TRUE or FALSE depending on the number of conditions. XOR is a variation of OR. In case of OR Function, only one condition needs to be TRUE. In case of XOR Function, each condition's logical value and the number of conditions both determine if XOR returns TRUE or FALSE. The general syntax is as follows:

  =XOR(logical1, [logical2],…)

  Here, the logical1, logical2 are the conditions which needs to be checked.

  Example 1: We have a list of Employees and their Rating and Experience. We want to test 2 condition for each employee. First condition is to test whether Rating is >3. Second condition is to test whether Experience is >4. The output after using XOR Function is as follows:

D2 =XOR(B2>3,C2>4)

	A	B	C	D	E
1	**Emp. Name**	**Rating**	**Experience**	**XOR**	**Formula used**
2	Ram	4	5	FALSE	=XOR(B2>3,C2>4)
3	Rajesh	3	4	FALSE	=XOR(B3>3,C3>4)
4	Amey	**2**	5	TRUE	=XOR(B4>3,C4>4)
5	Amar	**1**	8	TRUE	=XOR(B5>3,C5>4)
6					

(Figure 5.39 – XOR Function – 2 condition)

  - In this example, in first and second case, FALSE is returned as all the conditions are either met or not met.
  - In third and fourth case, TRUE is returned as one condition is not fulfilled out of 2 conditions.

  Example 2: We have a list of Employees and their Rating, Experience and Amount. We want to test 3 condition for each employee. First condition is to test whether Rating is >3. Second condition is to test whether Experience is >4. Third condition is to test whether Amount >2000. The output after using XOR Function is as follows:

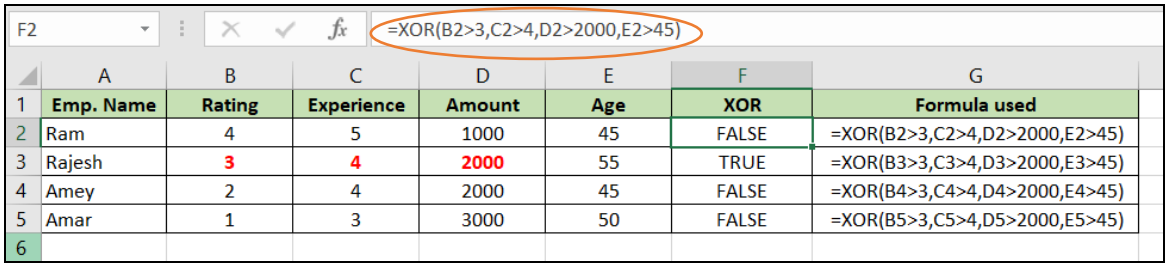
E2 =XOR(B2>3,C2>4,D2>2000)

	A	B	C	D	E	F
1	**Emp. Name**	**Rating**	**Experience**	**Amount**	**XOR**	**Formula used**
2	Ram	4	5	1000	FALSE	=XOR(B2>3,C2>4,D2>2000)
3	Rajesh	3	4	2000	FALSE	=XOR(B3>3,C3>4,D3>2000)
4	Amey	**2**	5	**2000**	TRUE	=XOR(B4>3,C4>4,D4>2000)
5	Amar	1	8	3000	FALSE	=XOR(B5>3,C5>4,D5>2000)
6						

Figure 5.40 – XOR Function – 3 condition

- In this example, in first and fourth case, FALSE is returned as all or majority of the conditions are met.
- In second case, FALSE is returned as all the conditions are not met.
- In third case, TRUE is returned as majority of the conditions are not fulfilled.

Example 3: We have a list of Employees and their Rating, Experience, Amount and Age. We want to test 4 condition for each employee. First condition is to test whether Rating is >3. Second condition is to test whether Experience is >4. Third condition is to test whether Amount >2000. Fourth condition is to test whether Age is >45. The output after using XOR Function is as follows:



F2 =XOR(B2>3,C2>4,D2>2000,E2>45)

	A	B	C	D	E	F	G
1	Emp. Name	Rating	Experience	Amount	Age	XOR	Formula used
2	Ram	4	5	1000	45	FALSE	=XOR(B2>3,C2>4,D2>2000,E2>45)
3	Rajesh	3	4	2000	55	TRUE	=XOR(B3>3,C3>4,D3>2000,E3>45)
4	Amey	2	4	2000	45	FALSE	=XOR(B4>3,C4>4,D4>2000,E4>45)
5	Amar	1	3	3000	50	FALSE	=XOR(B5>3,C5>4,D5>2000,E5>45)
6							

(Figure 5.41 – XOR Function – 4 condition)

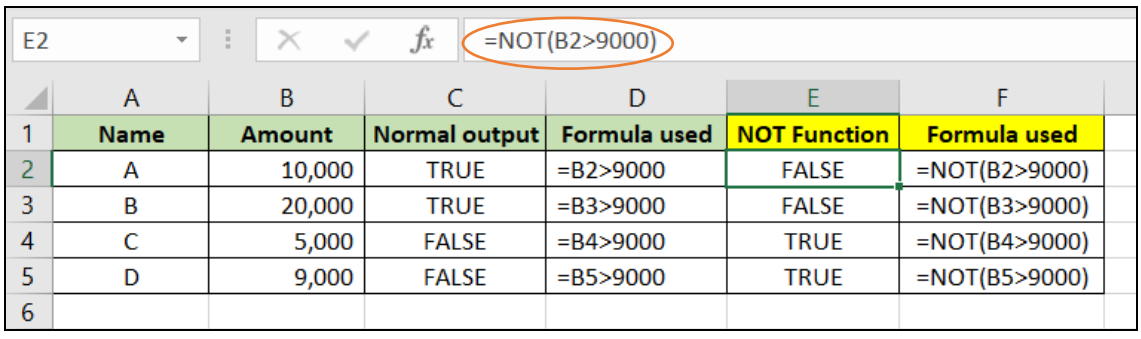
- In this example, in first and fourth case, FALSE is returned as all or majority of the conditions are met.
- In second case, TRUE is returned as majority of the conditions are not fulfilled.
- In third case, FALSE is returned as all the conditions are not met.

- **NOT Function:** The NOT Function returns the opposite of a given value. It means this function changes FALSE to TRUE or TRUE to FALSE. The general syntax is as follows:

  =NOT(logical)

  Here, the logical is the condition which needs to be tested.

  In the below example, we are trying to check whether the amount is greater than 9000. The output using NOT Function is as follows:



E2 =NOT(B2>9000)

	A	B	C	D	E	F
1	Name	Amount	Normal output	Formula used	NOT Function	Formula used
2	A	10,000	TRUE	=B2>9000	FALSE	=NOT(B2>9000)
3	B	20,000	TRUE	=B3>9000	FALSE	=NOT(B3>9000)
4	C	5,000	FALSE	=B4>9000	TRUE	=NOT(B4>9000)
5	D	9,000	FALSE	=B5>9000	TRUE	=NOT(B5>9000)
6						

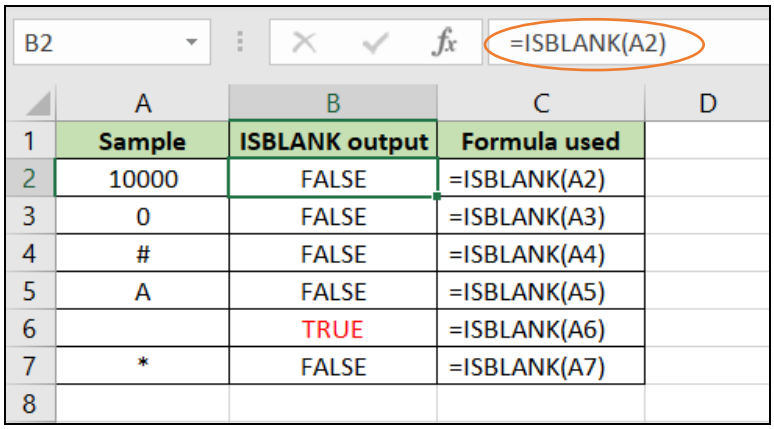
Figure 5.42 – NOT Function

- **ISBLANK Function:** This function checks whether a reference is to an empty cell. If empty, it returns TRUE, else FALSE. The general syntax is as follows:

  =ISBLANK(value)

  Here, the value is the reference to a cell to check whether empty or not.

  The example is as follows:



	A	B	C	D
1	**Sample**	**ISBLANK output**	**Formula used**	
2	10000	FALSE	=ISBLANK(A2)	
3	0	FALSE	=ISBLANK(A3)	
4	#	FALSE	=ISBLANK(A4)	
5	A	FALSE	=ISBLANK(A5)	
6		TRUE	=ISBLANK(A6)	
7	*	FALSE	=ISBLANK(A7)	
8				

Figure 5.43 – ISBLANK Function

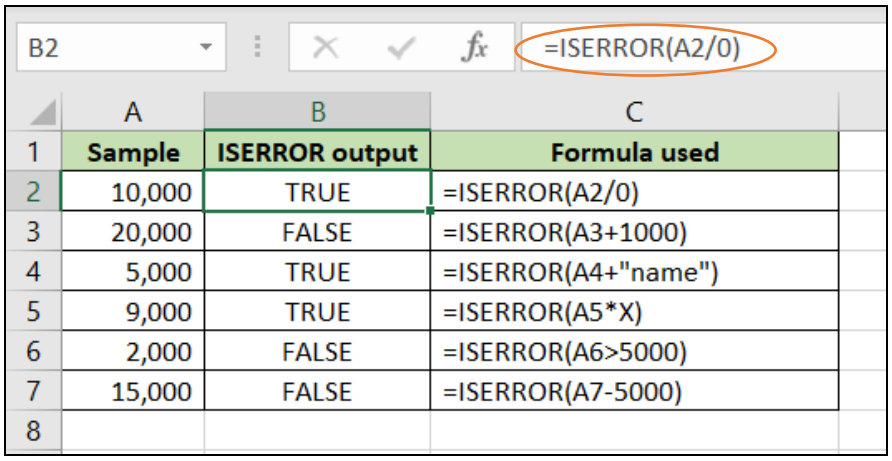
(Here "TRUE" is highlighted manually in red colour for reference)

- **ISERROR Function:** This function will return value as TRUE in case output is expected to be an Error, else False. The general syntax is as follows:

  =ISERROR(value)

  Here, the value is the expected value as per condition. If this value is expected to be an error, output will be TRUE, else FALSE.

  The example is as follows:



	A	B	C
1	**Sample**	**ISERROR output**	**Formula used**
2	10,000	TRUE	=ISERROR(A2/0)
3	20,000	FALSE	=ISERROR(A3+1000)
4	5,000	TRUE	=ISERROR(A4+"name")
5	9,000	TRUE	=ISERROR(A5*X)
6	2,000	FALSE	=ISERROR(A6>5000)
7	15,000	FALSE	=ISERROR(A7-5000)
8			

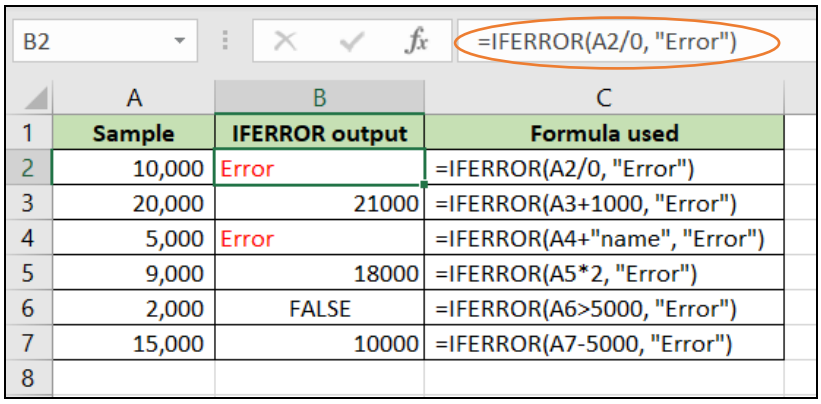
(Figure 5.44 – ISERROR Function)

- **IFERROR Function:** While working in Excel, we may get errors in output cell. In order to identify the errors, we can use IFERROR Function. This function will return "value if error" in case output value is expected0 to be an error, otherwise, it will return the value. The general syntax is as follows:

  = IFERROR(value, value_if_error)

  Here, the value is the expected value as per condition. If this value is an error, output value will be replaced by value_if_error.

  The example is as follows:



	A	B	C
1	**Sample**	**IFERROR output**	**Formula used**
2	10,000	Error	=IFERROR(A2/0, "Error")
3	20,000	21000	=IFERROR(A3+1000, "Error")
4	5,000	Error	=IFERROR(A4+"name", "Error")
5	9,000	18000	=IFERROR(A5*2, "Error")
6	2,000	FALSE	=IFERROR(A6>5000, "Error")
7	15,000	10000	=IFERROR(A7-5000, "Error")
8			

Figure 5.45 – IFERROR Function

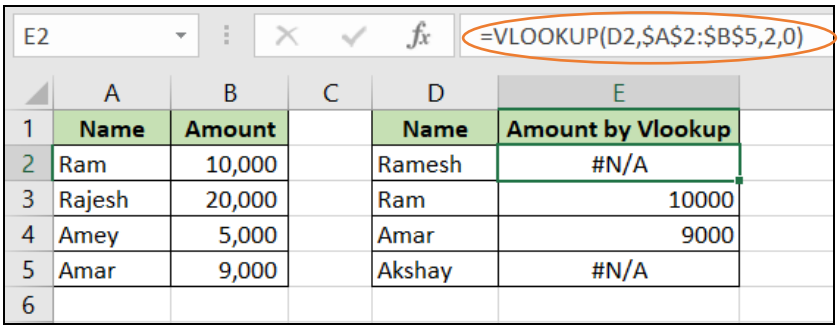
  In the above example, "Error" will be returned only if the value is expected to be an error, else the value is returned. (Here "Error" is highlighted manually in red colour for reference)

- **ISNA Function:** The ISNA Function checks a value and if it is #N/A, then it will return TRUE, else it will return FALSE. The general syntax is as follows:

  =ISNA(value)

  Here, value is checked for #N/A error.

  In the first example, we have a list of Names and Amount in one table and in another table, we want to search names and get the amount, if name matches. The output using VLOOKUP Function is as follows:

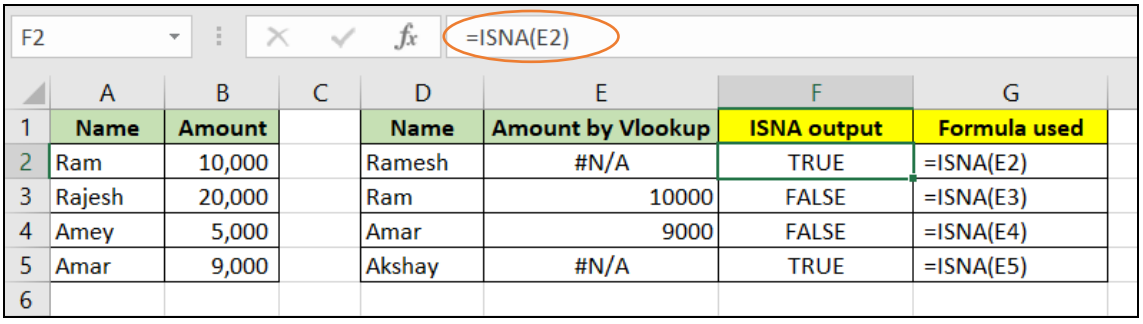


	A	B	C	D	E
1	**Name**	**Amount**		**Name**	**Amount by Vlookup**
2	Ram	10,000		Ramesh	#N/A
3	Rajesh	20,000		Ram	10000
4	Amey	5,000		Amar	9000
5	Amar	9,000		Akshay	#N/A
6					

(Figure 5.46 – VLOOKUP Function)

Note: VLOOKUP Function will be discussed in detail in later part of the chapter.

In the above table, we got 2 #N/A. This output using VLOOKUP, we want to test with ISNA Function. The output is as follows:



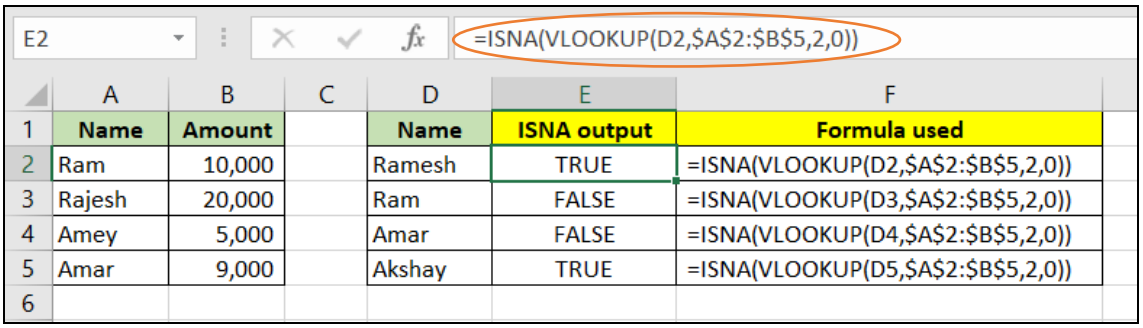
F2 =ISNA(E2)

	A	B	C	D	E	F	G
1	**Name**	**Amount**		**Name**	**Amount by Vlookup**	**ISNA output**	**Formula used**
2	Ram	10,000		Ramesh	#N/A	TRUE	=ISNA(E2)
3	Rajesh	20,000		Ram	10000	FALSE	=ISNA(E3)
4	Amey	5,000		Amar	9000	FALSE	=ISNA(E4)
5	Amar	9,000		Akshay	#N/A	TRUE	=ISNA(E5)
6							

(Figure 5.47 – ISNA Function)

In the above table, after using ISNA Function, we got output TRUE where we got #N/A error, else FALSE.

In the below table, we are using one single formula using ISNA and VLOOKUP Function to get the output. The output is as follows:



E2 =ISNA(VLOOKUP(D2,$A$2:$B$5,2,0))

	A	B	C	D	E	F
1	**Name**	**Amount**		**Name**	**ISNA output**	**Formula used**
2	Ram	10,000		Ramesh	TRUE	=ISNA(VLOOKUP(D2,$A$2:$B$5,2,0))
3	Rajesh	20,000		Ram	FALSE	=ISNA(VLOOKUP(D3,$A$2:$B$5,2,0))
4	Amey	5,000		Amar	FALSE	=ISNA(VLOOKUP(D4,$A$2:$B$5,2,0))
5	Amar	9,000		Akshay	TRUE	=ISNA(VLOOKUP(D5,$A$2:$B$5,2,0))
6						

(Figure 5.48 – ISNA and VLOOKUP Function)

- **IFNA Function:** The IFNA Function returns the value we specify if the expression resolves to #N/A, otherwise it returns the result of the expression. The general syntax is as follows:

  =IFNA(value, value_if_na)

  In the below example, we want "ERROR" text as an output using IFNA Function if output returned as #N/A after using VLOOKUP Function. The final output ("ISNA output" column) is as follows:

F2 =IFNA(E2, "ERROR")

	A	B	C	D	E	F	G
1	Name	Amount		Name	Amount by Vlookup	IFNA output	Formula used
2	Ram	10,000		Ramesh	#N/A	ERROR	=IFNA(E2, "ERROR")
3	Rajesh	20,000		Ram	10,000	10,000	=IFNA(E3, "ERROR")
4	Amey	5,000		Amar	9,000	9,000	=IFNA(E4, "ERROR")
5	Amar	9,000		Akshay	#N/A	ERROR	=IFNA(E5, "ERROR")
6							

(Figure 5.49 – IFNA Function)

In the below example, we want "ERROR" text or the value as an output using IFNA and VLOOKUP Function if output returned as #N/A. The output is as follows:

E2 =IFNA(VLOOKUP(D2,$A$2:$B$5,2,0), "ERROR")

	A	B	C	D	E	F
1	Name	Amount		Name	IFNA output	Formula used
2	Ram	10,000		Ramesh	ERROR	=IFNA(VLOOKUP(D2,$A$2:$B$5,2,0), "ERROR")
3	Rajesh	20,000		Ram	10,000	=IFNA(VLOOKUP(D3,$A$2:$B$5,2,0), "ERROR")
4	Amey	5,000		Amar	9,000	=IFNA(VLOOKUP(D4,$A$2:$B$5,2,0), "ERROR")
5	Amar	9,000		Akshay	ERROR	=IFNA(VLOOKUP(D5,$A$2:$B$5,2,0), "ERROR")
6						

(Figure 5.50 – IFNA and VLOOKUP Function)

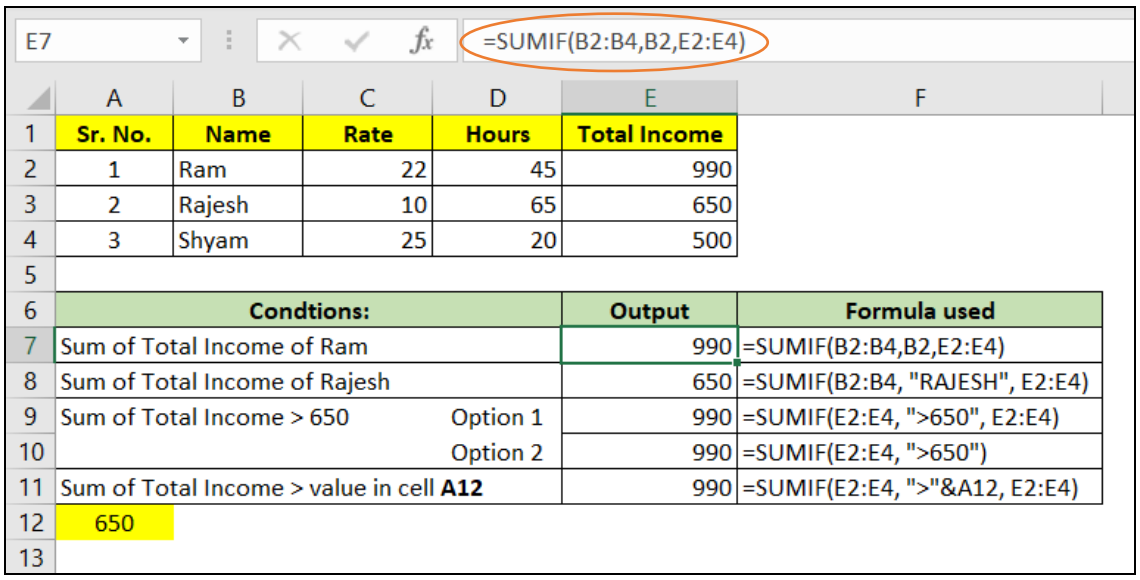
5.8 OTHER FUNCTIONS

- **SUMIF Function:** The SUMIF Function adds the cells specified by a given criteria or condition. The general syntax is as follows:

  =SUMIF(range, criteria, [sum_range])

  Here, "range" is the specified cells which will be added after applying the criteria.

  In the below example, we have a table, where we want to apply SUMIF Function on 4 different scenarios given below. The output is as follows:



E7 | =SUMIF(B2:B4,B2,E2:E4)

	A	B	C	D	E	F
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Total Income**	
2	1	Ram	22	45	990	
3	2	Rajesh	10	65	650	
4	3	Shyam	25	20	500	
5						
6	**Condtions:**				**Output**	**Formula used**
7	Sum of Total Income of Ram				990	=SUMIF(B2:B4,B2,E2:E4)
8	Sum of Total Income of Rajesh				650	=SUMIF(B2:B4, "RAJESH", E2:E4)
9	Sum of Total Income > 650			Option 1	990	=SUMIF(E2:E4, ">650", E2:E4)
10				Option 2	990	=SUMIF(E2:E4, ">650")
11	Sum of Total Income > value in cell **A12**				990	=SUMIF(E2:E4, ">"&A12, E2:E4)
12	650					
13						

(Figure 5.51 – SUMIF Function)

Process followed:

- In first question, we want Sum of Total Income of Ram. Hence –
  - **Range**: B2:B4, where we want to search the criteria
  - **Criteria**: the name "Ram". We have choice to give either cell reference or type name. (it is not case sensitive). Here we are giving cell reference i.e. B2
  - **Sum_range**: E2:E4 i.e. the Total Income column
- In second question, we want Sum of Total Income of Rajesh. Hence –
  - **Range**: B2:B4, where we want to search the criteria
  - **Criteria**: the name "Rajesh. We have choice to give either cell reference or type name. (it is not case sensitive). Here we are typing name i.e. "Rajesh"
  - **Sum_range**: E2:E4 i.e. the Total Income column
- In third question, we want Sum of Total Income which is > 650. Hence –
  - **Range**: E2:E4, where we want to search the criteria
  - **Criteria**: ">650"
  - **Sum_range**: E2:E4 i.e. the Total Income column.

  It can be observed that [sum_range] is in box bracket. It means that it is optional if range and sum range is same. Hence, alternatively, we can ignore sum range value.
- In fourth question, we want Sum of Total Income which is > than value in cell **A12**. It can be any value. For the same sake of convenience, we are taking value as 650. Hence –

- **Range**: E2:E4, where we want to search the criteria
- **Criteria**: ">"&A12. While using symbol and text in the criteria, it should be placed inside " ". But here, we want to use cell reference, which needs to be connected with ampersand (**&** operator).
- **Sum_range**: E2:E4 i.e. the Total Income column. Though optional, it is always advisable to give the sum range.

- **SUMIFS Function:** The SUMIFS Function sums a range of cells based on more than one condition or criteria. The general syntax is as follows:

  =SUMIFS(sum_range, criteria_range1, criteria1,….)

  Here, the sum_range is the range which we want to sum the cells based on criteria range and criteria. We can apply more than one criteria.

  The SUMIFS function works on AND concept. It means it will give output only if all the criteria are met.

  In the below example, we have a table, where we want to apply SUMIFS Function on 2 different scenarios given below. The output is as follows:

E7 =SUMIFS(E2:E4, B2:B4, "ram", C2:C4, ">22")

	A	B	C	D	E	F
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Total Income**	
2	1	Ram	22	45	990	
3	2	Rajesh	10	65	650	
4	3	RAM	25	75	1875	
5						
6	**Condtions:**				**Output**	**Formula used**
7	Sum of Total Income of Ram if rate > 22				1875	=SUMIFS(E2:E4, B2:B4, "ram", C2:C4, ">22")
8	Sum of Total Income > 500 & < 900				650	=SUMIFS(E2:E4, E2:E4, ">500", E2:E4, "<"&900)
9						

(Figure 5.52 – SUMIFS Function)

Process followed:

- In first question, we want Sum of Total Income of Ram if the rate is > 22. Hence –
  - **Sum_range:** E2:E4
  - **Criteria_range1**: B2:B4, where we want to search the 1st criteria
  - **Criteria1**: the name "Ram". We have choice to give either cell reference or type name. (it is not case sensitive). Here we have typed name i.e. "ram"
  - **Criteria_range2**: C2:C4, where we want to search the 2nd criteria
  - **Criteria2**: ">22"
- In second question, we want Sum of Total Income which is >500 and <900. Hence –

- **Sum_range:** E2:E4
- **Criteria_range1**: E2:E4, where we want to search the 1st criteria
- **Criteria1**: ">500".
- **Criteria_range2**: E2:E4, where we want to search the 2nd criteria
- **Criteria2**: "<"&900

In criteria, we have symbol and number to be connected. For the purpose of learning, in first criteria, we have placed both together in " " i.e. ">500". In second criteria, we have used ampersand (&) to connect symbol and number i.e. "<"&900.

Note: It is to be noted that depending on situation, the criteria range and sum range will be selected. Both need not be the same in every situation.

- **COUNTIFS Function:** Similar to SUMIFS Function, COUNTIFS Function also does the count based on more than one criteria. The COUNTIFS Function counts the number of cells in a range that meets all the given criteria. This function also works on AND concept. It means it will count only if all the criteria are met. The general syntax is as follows:

=COUNTIFS(criteria_range1, criteria1,.....)

Here, we can use more than one criteria and criteria range for counting.

The examples are given below:

E14 =COUNTIFS(B2:B11, "RAM", C2:C11, ">35")

	A	B	C	D	E	F
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Total Income**	
2	1	Amar	86	450	38,700	
3	2	Rajesh	25	460	11,500	
4	3	Rajesh	71	330	23,430	
5	4	Ram	67	470	31,490	
6	5	RAM	34	340	11,560	
7	6	Ram	86	340	29,240	
8	7	RAM	31	350	10,850	
9	8	Ramesh	36	390	14,040	
10	9	Shyam	81	490	39,690	
11	10	Vijay	91	320	29,120	
12						
13	**COUNTIFS Condtions:**				**Output**	**Formula used**
14	Count name Ram for Rate > 35				2	=COUNTIFS(B2:B11, "RAM", C2:C11, ">35")
15	Count Total Income > 15000 and < 35000				4	=COUNTIFS(E2:E11, ">15000", E2:E11, "<35000")
16						

Figure 5.53 – COUNTIFS Function

Process followed:

- In first question, we want count number of instances for Ram where rate is > 35. Hence –
  - **Criteria_range1**: B2:B11, where we want to search the 1st criteria
  - **Criteria1**: the name "Ram". We have choice to give either cell reference or type name. (it is not case sensitive). Here we have typed name i.e. "ram"
  - **Criteria_range2**: C2:C11, where we want to search the 2nd criteria
  - **Criteria2**: ">35"
- In second question, we want to count number of instances where Total Income is >15000 and <35000. Hence –
  - **Criteria_range1**: E2:E11, where we want to search the 1st criteria
  - **Criteria1**: ">15000".
  - **Criteria_range2**: E2:E11, where we want to search the 2nd criteria
  - **Criteria2**: "<35000"

  In criteria, we can connect < and 35000 using ampersand (& operator) also as we have done in SUMIFS scenario.

- **AVERAGEIFS Function:** The AVERAGEIFS function returns the average (arithmetic mean) of all cells that meet specific criteria we specify in the function. Like SUMIFS and COUNTIFS, this function also allows to specify several pairs of criteria to select the data that is included in the average. The general syntax is as follows:

  =AVERAGEIFS(average_range, criteria_range1, criteria1,.....)

  Example 1: We want Average Total Income for Ram and Rajesh where Rate is > 50.

D14 =AVERAGEIFS($E$2:$E$11, $B$2:$B$11, A14, $C$2:$C$11, ">50")

	A	B	C	D	E	F	G	H	I	J
1	Sr. No.	Name	Rate	Hours	Total Income					
2	1	Amar	86	450	38,700					
3	2	Rajesh	25	460	11,500					
4	3	Rajesh	71	330	23,430					
5	4	Ram	67	470	31,490					
6	5	RAM	34	340	11,560					
7	6	Ram	86	340	29,240					
8	7	RAM	31	350	10,850					
9	8	Ramesh	36	390	14,040					
10	9	Shyam	81	490	39,690					
11	10	Vijay	91	320	29,120					
12										
13	AVERAGEIFS Condtions: Rate > 50			Output	Formula used					
14	Ram			30,365	=AVERAGEIFS($E$2:$E$11, $B$2:$B$11, A14, $C$2:$C$11, ">50")					
15	Rajesh			23,430	=AVERAGEIFS($E$2:$E$11, $B$2:$B$11, A15, $C$2:$C$11, ">50")					
16										

Figure 5.54 – AVERAGEIFS Function

Process followed:

- As we wanted average Total Income for Ram as well as Rajesh, we created a single function for Ram in cell D14 and then copied for Rajesh in cell D15. –
  - **Average_range:** The Total Income i.e. $E$2:$E$11. This the column for which we want average. We placed $ sign in range to make it absolute cell reference.
  - **Criteria_range1:** The first criteria range is name column i.e. $B$2:$B$11 (again absolute reference has been used)
  - **Criteria1:** The criteria is name "Ram". We have given relative cell reference i.e. A14. It will change to A15, when we will copy for "Rajesh".
  - **Criteria_range2:** The second criteria range is Rate column i.e. $C$2:$C$11, (again absolute reference has been used)
  - Criteria2: The criteria is Rate > 50 i.e. ">50" entered manually. *(We can also use the cell reference for this criteria, if the criteria is available in anyone cell.)*

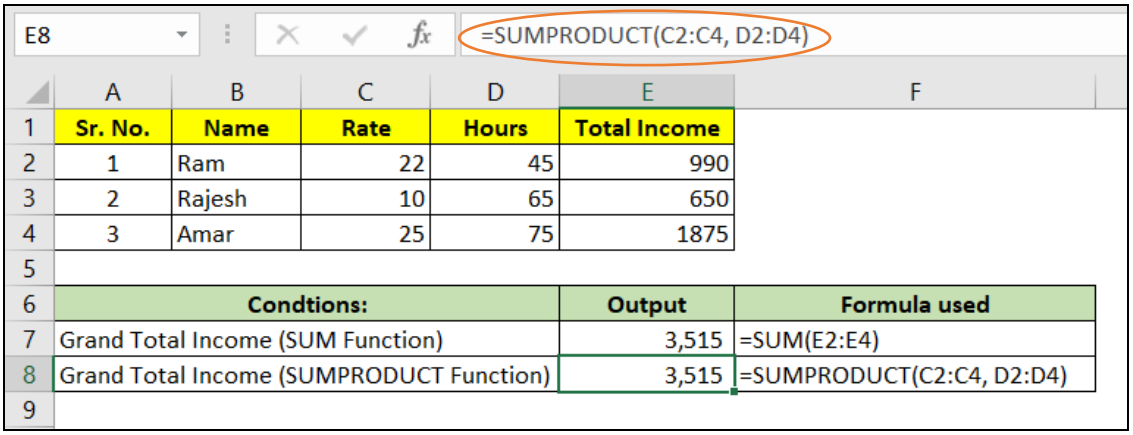
The formula is entered in cell D14 for Ram and then we copied it in cell D15 for Rajesh.

- **SUMPRODUCT Function:** This function returns the sum of the products of corresponding ranges or arrays. The general syntax is as follows:

=SUMPRODUCT(array1, [array2], [array3], ……)

We want to apply SUMPRODUCT Function on ranges. It is to be ensure that all the ranges are of same size, both rows and columns. In this function, up to 255 ranges are allowed, and at least 2 ranges are required. This function works by first multiplying values, by position, across the ranges and then adding all the results.

In the below example, we want SUMPRODUCT using Rate and Hours range.



	A	B	C	D	E	F
1	Sr. No.	Name	Rate	Hours	Total Income	
2	1	Ram	22	45	990	
3	2	Rajesh	10	65	650	
4	3	Amar	25	75	1875	
5						
6	Condtions:				Output	Formula used
7	Grand Total Income (SUM Function)				3,515	=SUM(E2:E4)
8	Grand Total Income (SUMPRODUCT Function)				3,515	=SUMPRODUCT(C2:C4, D2:D4)
9						

Figure 5.55 – SUMPRODUCT Function

Process followed:

- In the above table, if we want Grand Total Income. Then we need to add a new column i.e. "Total Income", where we will multiply rate and hours for each cell. Then we will do sum for Total Income in cell E7 to get output as 3515.
- In case we don't want to follow this long route, we can simply use SUMPRODUCT Function. Here first array will be Rate column i.e. C2:C4 and second array will be Hours column i.e D2:D4. We will get output as 3515 in a single cell in a single step.

5.9 LOOKUP FUNCTIONS

Excel has a group of functions which can be used to extract data from tables (database). A table (Database) is a dedicated matrix of rows and columns that collectively form a cohesive group of data. Tables usually have labels in the top row or the left column that identify the columns and rows of data. The remainder of the table contains the data itself.

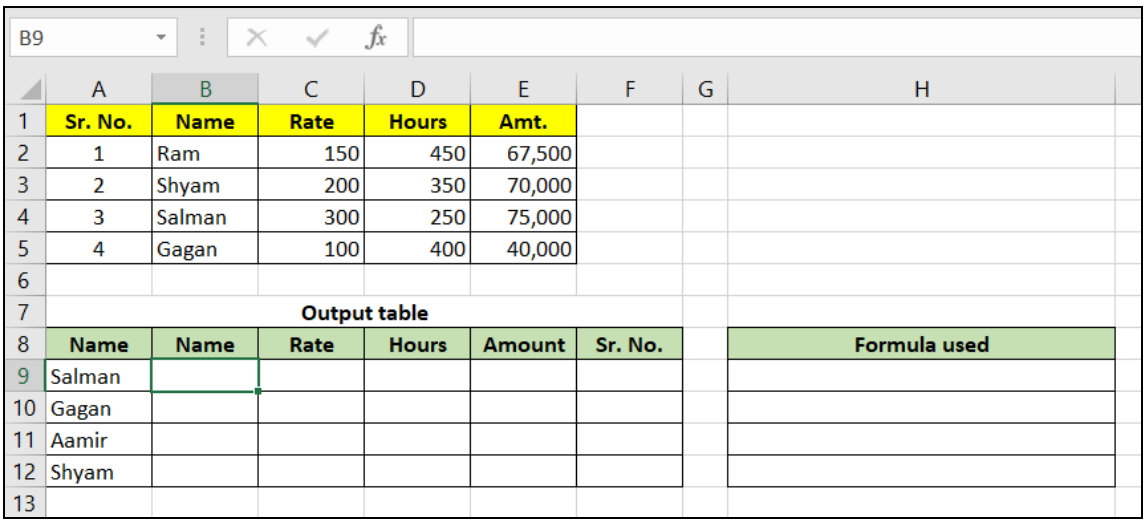
- **VLOOKUP Function:** The VLOOKUP function extracts the data from a particular cell in a table. VLOOKUP Function looks for a value in left most column of a table and then returns a value in the same row from a column specified by us. In short, this function starts by searching down the first column of the table to find a value that we specify. When it finds that value, it moves across a specified number of columns in the given range and returns the value in the target cell. Then general syntax is as follows:

  =VLOOKUP(lookup_value, , table_array, col_index_num, [range_lookup])

  Here, "lookup_value" is the value which we are looking in the "table_array". After matching the value in left most column (i.e. first column in the table_array), it will return the value in the target cell by considering col_index_num specified by us.

  The "range_lookup can be TRUE (Approximate match) or FALSE (Exact match). Selection of TRUE or FALSE will depend on the nature of lookup value and value in table array. Generally, in case of Text or exact number search, we use FALSE. In case table array is in the form of range or slab, in that situation, we prefer TRUE.

  In the below example, we have main table (database table), which has 5 columns i.e. Sr. No., Name, Rate, Hours and Amt. column. We also have Output Table, which has Name column (filled) and other unfilled columns (Name, Rate, Hours, Amount and Sr. No.).



	A	B	C	D	E	F	G	H
1	Sr. No.	Name	Rate	Hours	Amt.			
2	1	Ram	150	450	67,500			
3	2	Shyam	200	350	70,000			
4	3	Salman	300	250	75,000			
5	4	Gagan	100	400	40,000			
6								
7			Output table					
8	Name	Name	Rate	Hours	Amount	Sr. No.		Formula used
9	Salman							
10	Gagan							
11	Aamir							
12	Shyam							
13								

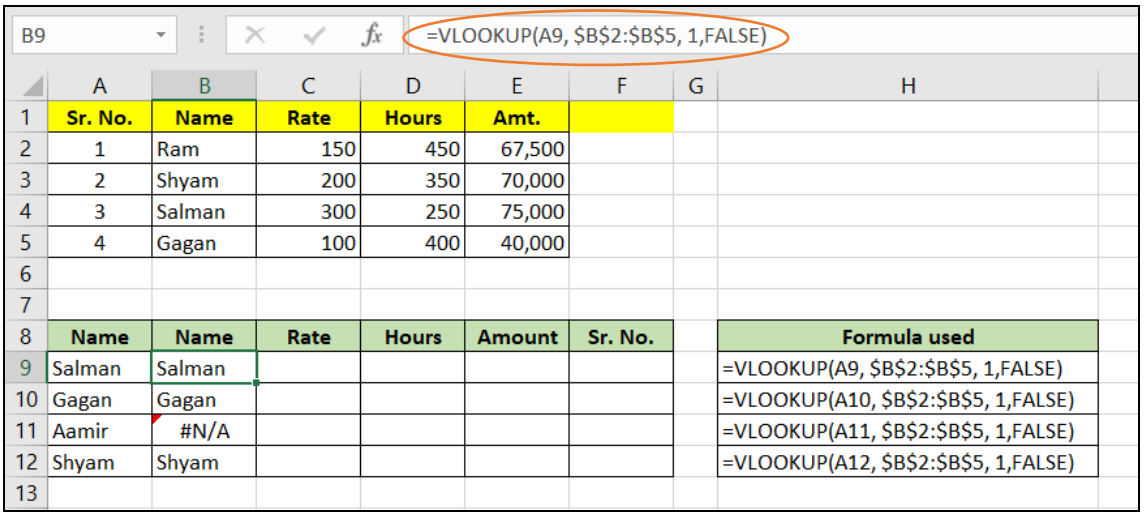
Figure 5.56 – VLOOKUP Function – Table structure

Using VLOOKUP Function, we will search Name column value from Output table in Database table and we will extract values from Name, Rate, Hours, Amount and Sr. No. from Database table.

In this example, for the sake of convenience, we have kept both tables in same sheet, but it is possible that both tables can be in 2 different sheet or file also.

OPTION 1: EXTRACTING INDIVIDUAL COLUMN VALUE

1: Name column using VLOOKUP Function



	A	B	C	D	E	F	G	H
1	Sr. No.	Name	Rate	Hours	Amt.			
2	1	Ram	150	450	67,500			
3	2	Shyam	200	350	70,000			
4	3	Salman	300	250	75,000			
5	4	Gagan	100	400	40,000			
6								
7								
8	Name	Name	Rate	Hours	Amount	Sr. No.		Formula used
9	Salman	Salman						=VLOOKUP(A9, $B$2:$B$5, 1,FALSE)
10	Gagan	Gagan						=VLOOKUP(A10, $B$2:$B$5, 1,FALSE)
11	Aamir	#N/A						=VLOOKUP(A11, $B$2:$B$5, 1,FALSE)
12	Shyam	Shyam						=VLOOKUP(A12, $B$2:$B$5, 1,FALSE)
13								

Figure 5.57 – VLOOKUP Function for Name column

Process followed:

- We are using VLOOKUP Function as the Database table is vertically aligned.

- **lookup_value:** cell A9 i.e. name “Salman”
- **table_array:** range $B$2:$B$5 i.e. the range in which we want to search the lookup value. (absolute range. So that it will not change once we copy in below rows)
- **col_index_num:** column no. 1 as in the table array we have selected only one column and from that particular column we want name if matching.
- **range_lookup:** FALSE (0) as we want exact match.
- If the lookup value is not found, it will return #N/A error. It can be seen that name “Aamir” is not not in the database table. Hence, #N/A error came in cell B11.
- In case, we don’t want error to be shown, we can use IFERROR or IFNA function before VLOOKUP as shown below:
  - **Use of IFERROR**

    =IFERROR(VLOOKUP(A20,$B$2:$B$5,1,0), "Not Available")
  - **Use of IFNA**

    =IFNA(VLOOKUP(A20,$B$2:$B$5,1,0), "Not Available")

In both the case, if error comes, it will show output as “Not Available”, else value will be shown.

B9 =IFNA(VLOOKUP(A9,$B$2:$B$5,1,0), "Not Available")

	A	B	C	D	E	F	G
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Amt.**	**Sr. No.**	
2	1	Ram	150	450	67,500	1	
3	2	Shyam	200	350	70,000	2	
4	3	Salman	300	250	75,000	3	
5	4	Gagan	100	400	40,000	4	
6							
7							
8	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**	
9	Salman	Salman					
10	Gagan	Gagan					
11	Aamir	Not Available					
12	Shyam	Shyam					
13							

Figure 5.58 – IFNA and VLOOKUP Function for error handling

2: Rate column using VLOOKUP Function

Process followed in Figure 5.59 below :

- We are using VLOOKUP Function as the Database table is vertically aligned.
- **lookup_value:** cell A9 i.e. name "Salman"
- **table_array:** range $B$2:$C$5 i.e. the range in which we want to search the lookup value. (absolute range. So that it will not change once we copy in below rows)
- **col_index_num:** column no. 2 as in the table array we have selected 2 column and we want output from 2nd column.
- **range_lookup:** FALSE (0) as we want exact match.

If the lookup value is not found, it will return #N/A error. It can be seen that name "Aamir" is not not in the database table. Hence, #N/A error came in cell C11.

C9 =VLOOKUP(A9,$B$2:$C$5, 2,FALSE)

	A	B	C	D	E	F	G	H
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Amt.**			
2	1	Ram	150	450	67,500			
3	2	Shyam	200	350	70,000			
4	3	Salman	300	250	75,000			
5	4	Gagan	100	400	40,000			
6								
7								
8	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**		**Formula used**
9	Salman	Salman	300					=VLOOKUP(A9,$B$2:$C$5, 2,FALSE)
10	Gagan	Gagan	100					=VLOOKUP(A10,$B$2:$C$5, 2,FALSE)
11	Aamir	#N/A	#N/A					=VLOOKUP(A11,$B$2:$C$5, 2,FALSE)
12	Shyam	Shyam	200					=VLOOKUP(A12,$B$2:$C$5, 2,FALSE)
13								

Figure 5.59 – VLOOKUP Function for Rate column

3: Hours and Amount column using VLOOKUP Function

Process followed in Figure 5.60 below:

- We are using VLOOKUP Function as the Database table is vertically aligned.
- **lookup_value:** cell A9 i.e. name "Salman"
- **table_array:** For Hours, range is $B$2:$D$5 and for Amount column, the range is $B$2:$E$5
- **col_index_num:** For Hours, the column no. 3 as in the table array we have selected 3 column and we want output from 3rd column. Similarly, for Amount, the column no. 4 as in the table array we have selected 4 column and we want output from 4th column.
- **range_lookup:** FALSE (0) as we want exact match.

- If the lookup value is not found, it will return #N/A error. It can be seen that name "Aamir" is not not in the database table. Hence, #N/A error came in cell D11 and E11.

D9 =VLOOKUP(A9,$B$2:$D$5, 3, 0)

	A	B	C	D	E	F	G	H
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Amt.**			
2	1	Ram	150	450	67,500			
3	2	Shyam	200	350	70,000			
4	3	Salman	300	250	75,000			
5	4	Gagan	100	400	40,000			
6								
7								
8	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**		**Formula used**
9	Salman	Salman	300	250	75000			=VLOOKUP(A9,$B$2:$D$5, 3, 0)
10	Gagan	Gagan	100	400	40000			=VLOOKUP(A10,$B$2:$D$5, 3, 0)
11	Aamir	#N/A	#N/A	#N/A	#N/A			=VLOOKUP(A11,$B$2:$D$5, 3, 0)
12	Shyam	Shyam	200	350	70000			=VLOOKUP(A12,$B$2:$D$5, 3, 0)
13								

Figure 5.60 – VLOOKUP Function for Hours and Amount column

4: Sr. No. column using VLOOKUP Function

It can be seen that in the table array, the first column has to be same as lookup value. While using VLOOKUP for Sr. No., when we select Sr. No. and Name column (i.e. A1:B5), the output comes as #N/A because as per pre-requisite, the first column has to be common value i.e. Name.

Hence, we need to shift either Sr. No. column to last column in database table or make Name column as first column (before Sr. No. column). In this example, we have shifted Sr. No. column to last column in below table.

F9 =VLOOKUP(A9,$B$2:$F$5, 5, 0)

	A	B	C	D	E	F	G	H
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Amt.**	**Sr. No.**		
2	1	Ram	150	450	67,500	1		
3	2	Shyam	200	350	70,000	2		
4	3	Salman	300	250	75,000	3		
5	4	Gagan	100	400	40,000	4		
6								
7								
8	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**		**Formula used**
9	Salman	Salman	300	250	75000	3		=VLOOKUP(A9,$B$2:$F$5, 5, 0)
10	Gagan	Gagan	100	400	40000	4		=VLOOKUP(A10,$B$2:$F$5, 5, 0)
11	Aamir	#N/A	#N/A	#N/A	#N/A	#N/A		=VLOOKUP(A11,$B$2:$F$5, 5, 0)
12	Shyam	Shyam	200	350	70000	2		=VLOOKUP(A12,$B$2:$F$5, 5, 0)
13								

Figure 5.61 – VLOOKUP Function for Sr. No. column

Process followed:

- We are using VLOOKUP Function as the Database table is vertically aligned.
- **lookup_value:** cell A9 i.e. name "Salman"
- **table_array:** For new Sr. No. column, range is $B$2:$F$5
- **col_index_num:** column no. 5 as in the table array we have selected 5 column starting from Name column to Sr. No. column and we want output from 5th column.
- **range_lookup:** FALSE (0) as we want exact match.
- If the lookup value is not found, it will return #N/A error. It can be seen that name "Aamir" is not not in the database table. Hence, #N/A error came in cell F11.

OPTION 2: EXTRACTING ALL COLUMN VALUE TOGETHER USING CELL REFERENCE

In the above example, we have used 5 different VLOOKUP Functions and extracted individual column value.

In this example, we want to make a master formula using VLLOKUP and Mixed Cell Reference to get value for all the columns as shown below.

B10 =VLOOKUP($A10,$B$2:$F$5,B$8,0)

	A	B	C	D	E	F
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Amt.**	**Sr. No.**
2	1	Ram	150	450	67,500	1
3	2	Shyam	200	350	70,000	2
4	3	Salman	300	250	75,000	3
5	4	Gagan	100	400	40,000	4
6						
7	**Output table**					
8	**Coulmn No.**	**1**	**2**	**3**	**4**	**5**
9	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**
10	Salman	Salman	300	250	75000	3
11	Gagan	Gagan	100	400	40000	4
12	Aamir	#N/A	#N/A	#N/A	#N/A	#N/A
13	Shyam	Shyam	200	350	70000	2
14						
15						
16	**Master formula created in cell B10 and then copied in entire Output Table**					
17		=VLOOKUP($A10,$B$2:$F$5,B$8,0)				
18						

Figure 5.62 – single VLOOKUP Function for entire table using cell reference

Process followed:

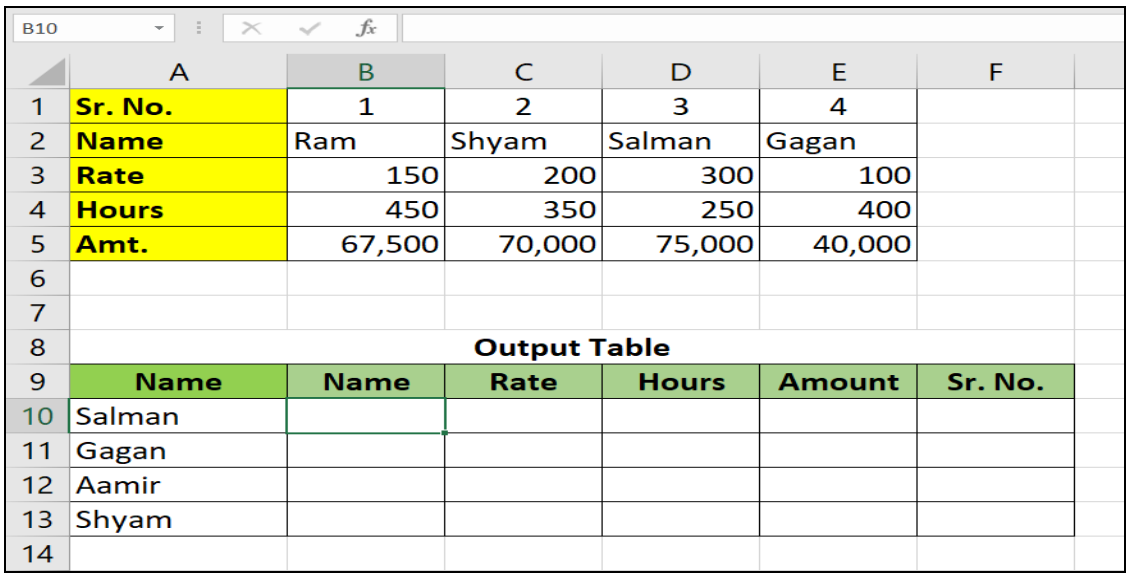
- We are using VLOOKUP Function as the Database table is vertically aligned.
- Above Output table, we are adding a Column No. row and manually enter each column reference number as per its position in Database table above starting from Name column (common column as per lookup value). (Name – 1, Rate – 2, Hours – 3, Amount – 4, Sr. No. 5)
- **lookup_value:** cell $A10 i.e. name "Salman" (mixed cell reference). Here only column reference is fixed as column A should not change as we copy the formula in across the output table.
- **table_array:** In the table array, we are going to select full database table starting from Name column being the common column in both table i.e. $B$2:$F$5
- **col_index_num:** cell B$8 i.e. 1 above Name column in Output table (mixed cell reference). Here only row reference is fixed as $8 (row) should not change as we copy the formula in across the output table.
- **range_lookup:** FALSE (0) as we want exact match.
- If the lookup value is not found, it will return #N/A error. It can be seen that name "Aamir" is not not in the database table. Hence, #N/A error came in cell B12, C12, D12, E12 and F12.

Once the above formula is created, we need to copy this formula in entire output table to get the values. By this method, we can extract values from Database table in each column of Output Table by a single formula.

- **HLOOKUP Function:** Similar to VLOOKUP function, HLOOKUP Function looks for a value in Top row of a table or array of values and then returns a value in the same column from a row specified by us. In short, this function starts by searching across the first row of the table to find a value that we specify. When it finds that value, it moves down the column a specified number of rows in the given range and returns the value in the target cell. Then general syntax is as follows:

  =HLLOKUP(lookup_value, , table_array, row_index_num, [range_lookup])

  In the below example, we have a database table which is horizontally aligned. From this database table, we want extract values in Output Table.



	A	B	C	D	E	F
1	**Sr. No.**	1	2	3	4	
2	**Name**	Ram	Shyam	Salman	Gagan	
3	**Rate**	150	200	300	100	
4	**Hours**	450	350	250	400	
5	**Amt.**	67,500	70,000	75,000	40,000	
6						
7						
8	**Output Table**					
9	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**
10	Salman					
11	Gagan					
12	Aamir					
13	Shyam					
14						

Figure 5.63 – HLOOKUP Function – Table structure

We are going to extract values for each column. Formula used in each column are given below in B17 (Name), B18 (Rate), B19 (Hours), B20 (Amount) and B21 (Sr. No.).

B10 =HLOOKUP(A10, $B$2:$E$2, 1, 0)

	A	B	C	D	E	F	G
1	**Sr. No.**	1	2	3	4		
2	**Name**	Ram	Shyam	Salman	Gagan		
3	**Rate**	150	200	300	100		
4	**Hours**	450	350	250	400		
5	**Amt.**	67,500	70,000	75,000	40,000		
6	**Sr. No.**	1	2	3	4		
7							
8		**Output Table**					
9	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**	
10	Salman	Salman	300	250	75,000	3	
11	Gagan	Gagan	100	400	40,000	4	
12	Aamir	#N/A	#N/A	#N/A	#N/A	#N/A	
13	Shyam	Shyam	200	350	70,000	2	
14							
15							
16	**Formula used in respective column of Output Table**						
17	Name	=HLOOKUP(A10, $B$2:$E$2, 1, 0)					
18	Rate	=HLOOKUP(A10, $B$2:$E$3, 2, 0)					
19	Hours	=HLOOKUP(A10, $B$2:$E$4, 3, 0)					
20	Amonut	=HLOOKUP(A10, $B$2:$E$5, 4, 0)					
21	Sr. No.	=HLOOKUP(A10, $B$2:$E$6, 5, 0)					
22							

Figure 5.64 – HLOOKUP Function for each column

Process followed for Name column using HLOOKUP function:

- We are using HLOOKUP Function as the Database table is horizontally aligned.
- **lookup_value:** cell A10 i.e. name “Salman”
- **table_array:** range $B$2:$E$2 i.e. the horizontal range in which we want to search the lookup value. (absolute range. So that it will not change once we copy in below rows)
- **row_index_num:** row no. 1 as in the table array we have selected only one row and from that particular row we want name if matching.
- **range_lookup:** FALSE (0) as we want exact match.

- If the lookup value is not found, it will return #N/A error. It can be seen that name "Aamir" is not not in the database table. Hence, #N/A error came in cell B12.

Like the above process, we need to follow the process for extracting values for other columns as shown in above table.

Like master formula created in VLOOKUP Function (Option 2), we can also create a master formula in HLOOKUP as well. The formula given in below table:

B11 =HLOOKUP($A11,$B$2:$E$6,B$9,FALSE)

	A	B	C	D	E	F	G
1	**Sr. No.**	1	2	3	4		
2	**Name**	Ram	Shyam	Salman	Gagan		
3	**Rate**	150	200	300	100		
4	**Hours**	450	350	250	400		
5	**Amt.**	67,500	70,000	75,000	40,000		
6	**Sr. No.**	1	2	3	4		
7							
8		**Output Table**					
9	**Row No.**	**1**	**2**	**3**	**4**	**5**	
10	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**	
11	Salman	Salman	300	250	75000	3	
12	Gagan	Gagan	100	400	40000	4	
13	Aamir	#N/A	#N/A	#N/A	#N/A	#N/A	
14	Shyam	Shyam	200	350	70000	2	
15							
16							
17	**Master formula created in cell B11 and then copied in entire Output Table**						
18	=HLOOKUP($A11,$B$2:$E$6,B$9,FALSE)						
19							

Figure 5.65 – single HLOOKUP Function for entire table using cell reference

Process followed:

- We are using HLOOKUP Function as the Database table is horizontally aligned.
- Above Output table, we are adding a **Row No.** row and manually enter each row reference number as per its position in Database table above starting from Name row (common row as per lookup value). (Name – 1, Rate – 2, Hours – 3, Amount – 4, Sr. No. 5)
- **lookup_value:** cell $A11 i.e. name "Salman" (mixed cell reference). Here only column reference is fixed as column A should not change as we copy the formula in across the output table.

- **table_array:** In the table array, we are going to select full database table starting from Name column being the common column in both table i.e. $B$2:$E$6
- **row_index_num:** cell B$9 i.e. 1 above Name column in Output table (mixed cell reference). Here only row reference is fixed as $9 (row) should not change as we copy the formula in across the output table.
- **range_lookup:** FALSE (0) as we want exact match.
- If the lookup value is not found, it will return #N/A error. It can be seen that name "Aamir" is not not in the database table. Hence, #N/A error came in cell B13, C13, D13, E13 and F13.

Once the above formula is created, we need to copy this formula in entire output table to get the values. By this method, we can extract values from Database table in each column of Output Table by a single formula.

- **MATCH Function:** The MATCH Function returns the relative position of an item in an array that matches a specified value in a specified order. In other words, it searches for a specified item in a range of cells and returns the relative position of that item in the range. The general syntax is as follows:

  =MATCH(lookup_value, lookup_array, [match_type])

  Here, the lookup_value can be a number or text, which will be searched in lookup_array (range). The match type can be Less than (1), Exact match (0) and Greater than match (-1). When we are searching Text or specific numbers, then we prefer Exact match i.e. 0.

  In the below example, we want to extract column number using MATCH Function.

B9 | =MATCH(B10,$B$1:$F$1,0)

	A	B	C	D	E	F	G	H
1	Sr. No.	Name	Rate	Hours	Amt.	Sr. No.		
2	1	Ram	150	450	67,500	1		
3	2	Shyam	200	350	70,000	2		
4	3	Salman	300	250	75,000	3		
5	4	Gagan	100	400	40,000	4		
6								
7								
8			Output table					MATCH Function used for Column No.
9	Coulmn No.	1	2	3	#N/A	5		=MATCH(B10,$B$1:$F$1,0)
10	Name	Name	Rate	Hours	Amount	Sr. No.		
11	Salman							
12	Gagan							
13	Aamir							
14	Shyam							
15								

Figure 5.66 – MATCH Function for column number

Process followed for Column No. in row no. 9 using MATCH Function:

- We are using MATCH Function as we want position number of each column heading.
- **lookup_value:** cell B10 i.e. Name column in Output Table
- **lookup_array:** range $B$1:$F$1 i.e. Database table heading in which we want to search the lookup value. (absolute range. So that it will not change once we copy in across the columns)
- **match_type:** Exact Match (0) as we want exact match.
- If the lookup value is not found, it will return #N/A error. It can be seen that in cell E9, we are getting #N/A error as the headings are not matching ("Amount" column in Output table is different from "Amt." column in Database table. Hence, we need to make the heading common. Here we have made both heading as "Amount" as shown in below table.)

Earlier in VLOOKUP, we created a master formula for extracting all the values using a master formula. In that we used Column No. row where we entered column no. manually.

Now we want to extract all the values using VLOOKUP and MATCH function together, where the column number will be extracted using MATCH function as shown in below table:

B9 =VLOOKUP($A9,$B$2:$F$5, MATCH(B$8,$B$1:$F$1,0),0)

	A	B	C	D	E	F	G	H
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**		
2	1	Ram	150	450	67,500	1		
3	2	Shyam	200	350	70,000	2		
4	3	Salman	300	250	75,000	3		
5	4	Gagan	100	400	40,000	4		
6								
7	**Output table**							
8	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**		
9	Salman	Salman	300	250	75000	3		
10	Gagan	Gagan	100	400	40000	4		
11	Aamir	#N/A	#N/A	#N/A	#N/A	#N/A		
12	Shyam	Shyam	200	350	70000	2		
13								
14								
15	**Formual used in cell B9 (VLOOKUP + MATCH)**							
16	=VLOOKUP($A9,$B$2:$F$5, MATCH(B$8,$B$1:$F$1,0),0)							
17								

Figure 5.67 – VLOOKUP and MATCH Function for extracting all column data

Process followed:

- We are using VLOOKUP Function as the Database table is vertically aligned.

- **lookup_value:** cell $A9 i.e. name "Salman" (mixed cell reference). Here only column reference is fixed as column A should not change as we copy the formula in across the Output table.
- **table_array:** In the table array, we are going to select full database table starting from Name column being the common column in both table i.e. $B$2:$F$5
- **col_index_num:** There we want column number using MATCH function as explained in above table.
  - **lookup_value:** cell B$8 i.e. Name column in Output Table (mixed cell reference – row number fixed)
  - **lookup_array:** range $B$1:$F$1 i.e. Database table heading in which we want to search the lookup value. (absolute range. So that it will not change once we copy in across the columns)
  - **match_type:** Exact Match (0) as we want exact match
- **range_lookup:** FALSE (0) as we want exact match.
- If the lookup value is not found, it will return #N/A error. It can be seen that name "Aamir" is not not in the database table. Hence, #N/A error came in cell B11.

Once the above formula is created, we need to copy this formula in entire Output table to get the values. By this method, we can extract values from Database table in each column of Output Table by a single formula.

- **INDEX Function:** The INDEX Function returns a value or reference of the cell at the intersection of a particular row and column in a given range. The general syntax is as follows:

  =INDEX(array, row_num, [column_num])

  Here, the array is the range, where we have to search for value based on intersection of row number and column number. The row number and column can be written manually (can be difficult in large table) or can be derived using MATCH Function.

  Earlier in MATCH Function, we used VLOOKUP and MATCH Function together to extract the value.

  Now we want to use INDEX and MATCH Function together to extract the values as shown in below table:

  Note: As the index Function can extract value from any place in the range, there is no need to shift the "Sr. No." column.

  Process followed in Figure 5.68 below:

  - We are using INDEX Function as INDEX Function can extract value from any place in the given range.
  - **array:** the full table range i.e. $A$1:$E$5 (absolute cell reference).
  - **row_num:** We want row number using MATCH function as explained below.
    - **lookup_value:** cell $A9 i.e. Name column in Output Table (mixed cell reference – column number fixed)

- **lookup_array:** range $B$1:$B$5 i.e. Name column in Database table in which we want to search the lookup value. (absolute range. So that it will not change once we copy in across the table)
- **match_type:** Exact Match (0) as we want exact match

- **column_num:** We want column number using MATCH function as explained below.
  - **lookup_value:** cell B$8 i.e. Name column in Output Table (mixed cell reference – row number fixed)
  - **lookup_array:** range $A$1:$E$1 i.e. Top row in Database table in which we want to search the lookup value. (absolute range. So that it will not change once we copy in across the table)
  - **match_type:** Exact Match (0) as we want exact match
- If the lookup value is not found, it will return #N/A error. It can be seen that name "Aamir" is not not in the database table. Hence, #N/A error came in cell B11.

B9 fx =INDEX($A$1:$E$5, MATCH($A9,$B$1:$B$5,0), MATCH(B$8,$A$1:$E$1,0))

	A	B	C	D	E	F
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Amount**	
2	1	Ram	150	450	67,500	
3	2	Shyam	200	350	70,000	
4	3	Salman	300	250	75,000	
5	4	Gagan	100	400	40,000	
6						
7	**Output table**					
8	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**
9	Salman	Salman	300	250	75000	3
10	Gagan	Gagan	100	400	40000	4
11	Aamir	#N/A	#N/A	#N/A	#N/A	#N/A
12	Shyam	Shyam	200	350	70000	2
13						
14						
15	**Formual used in cell B9 (VLOOKUP + MATCH)**					
16	=INDEX($A$1:$E$5,MATCH($A9,$B$1:$B$5,0),MATCH(B$8,$A$1:$E$1,0))					
17						

Figure 5.68 – INDEX and MATCH Function for extracting all column data

Once the above formula is created, we need to copy this formula in entire Output table to get the values. By this method, we can extract values from Database table in each column of Output Table by a single formula.

- **CHOOSE Function:** The CHOOSE Function chooses a value or action to perform from a list of values based on an index number. The CHOOSE function is ideal for converting a value to a literal. In plain-speak, this means turning a number, such as 4, into a word, such as April. CHOOSE takes up to 30 arguments. The first argument acts as key to the rest of the arguments. In fact, the other arguments do not get processed per se by the function. Instead, the function looks at the value of the first argument and, based on that value, returns one of its other arguments.

  The general syntax is as follows:

  =CHOOSE(index_num, value1, [value2], ...)

  Here, index number must be, or evaluate to, a number. This number in turn indicates which of the following arguments to return. For example, the following returns Three as an Output:

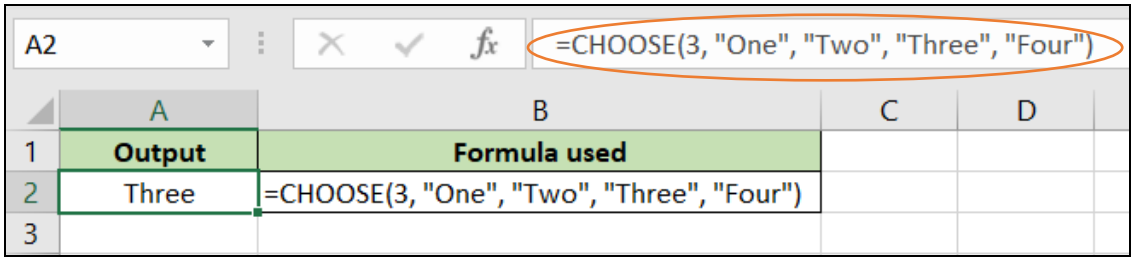


Figure 5.69 – CHOOSE Function for Number

Explanation: The index number is 3. Hence, it is going to return 3rd value as an Output. If index number changes to 2, then it will return 2nd value.

In another example, a list of month numbers from 1 to 12. We want display Month name based on Month number. Here we can use CHOOSE Function as follows:

The function used:

=CHOOSE(B1, "January", "February", "March", "April", "May", "June", "July", "August", "September", "October", "November", "December")

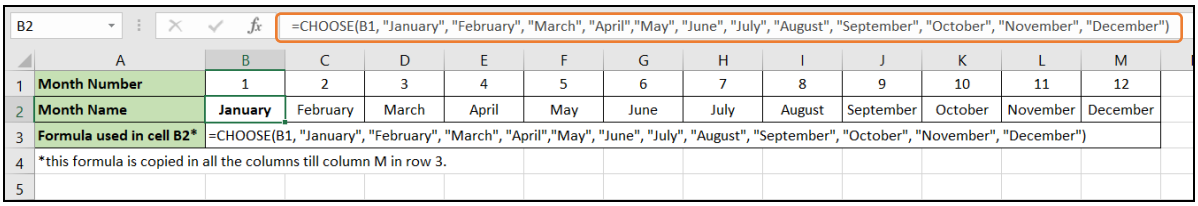
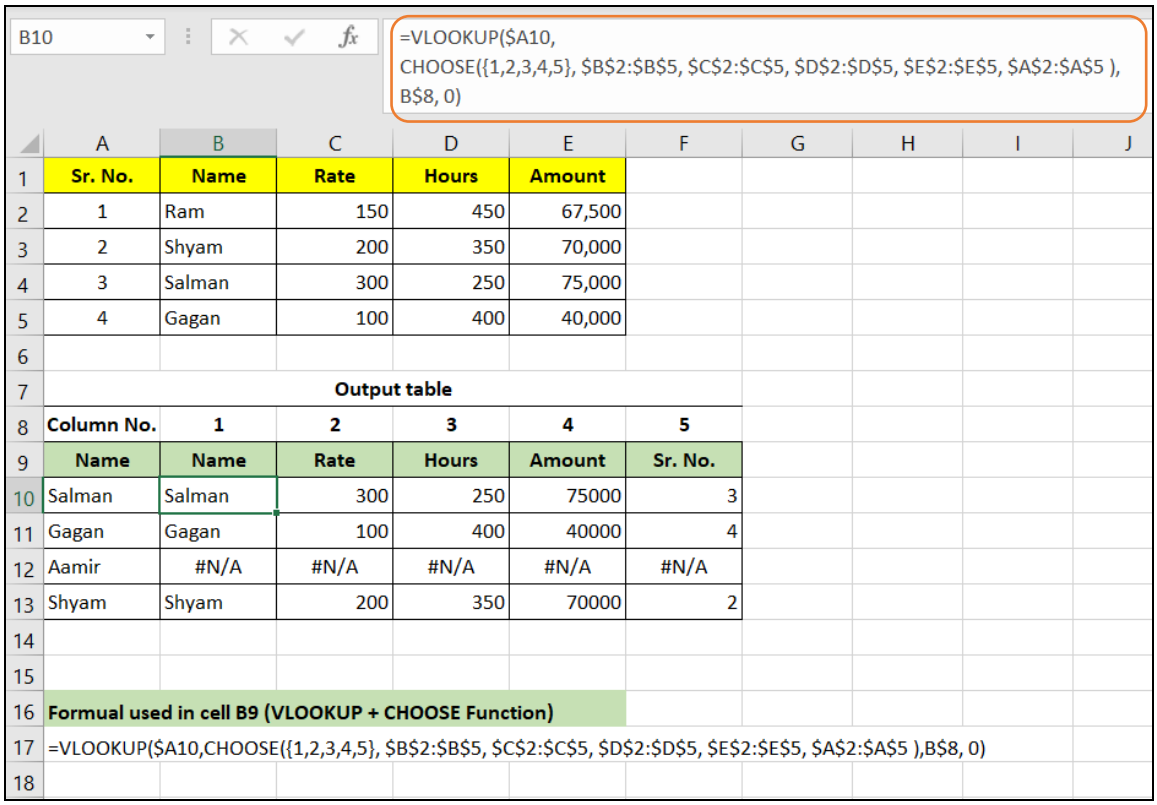


Figure 5.70 – CHOOSE Function for Month Name

Cell B1 contains the value 1, so the first argument starting in the list of possible returned strings (that is, "January") is returned. Likewise other values are returned.

CHOOSE is most often used to return meaningful text that relates to a number, such as returning the name of a month from its numeric value. But CHOOSE is not restricted to returning text strings. We can use it to return numbers as well.

We can also use VLOOKUP and CHOOSE function together to extract values ini Output table from Database table as shown below:



B10 | =VLOOKUP($A10, CHOOSE({1,2,3,4,5}, $B$2:$B$5, $C$2:$C$5, $D$2:$D$5, $E$2:$E$5, $A$2:$A$5), B$8, 0)

	A	B	C	D	E	F	G	H	I	J
1	**Sr. No.**	**Name**	**Rate**	**Hours**	**Amount**					
2	1	Ram	150	450	67,500					
3	2	Shyam	200	350	70,000					
4	3	Salman	300	250	75,000					
5	4	Gagan	100	400	40,000					
6										
7				**Output table**						
8	**Column No.**	**1**	**2**	**3**	**4**	**5**				
9	**Name**	**Name**	**Rate**	**Hours**	**Amount**	**Sr. No.**				
10	Salman	Salman	300	250	75000	3				
11	Gagan	Gagan	100	400	40000	4				
12	Aamir	#N/A	#N/A	#N/A	#N/A	#N/A				
13	Shyam	Shyam	200	350	70000	2				
14										
15										
16	**Formual used in cell B9 (VLOOKUP + CHOOSE Function)**									
17	=VLOOKUP($A10,CHOOSE({1,2,3,4,5}, $B$2:$B$5, $C$2:$C$5, $D$2:$D$5, $E$2:$E$5, $A$2:$A$5),B$8, 0)									
18										

Figure 5.71 – VLOOKUP and CHOOSE Function extracting all column values

Process followed:

- We are using VLOOKUP Function as the Database table is vertically aligned.
- Above Output table, we are adding a Column No. row and manually enter each column reference number as per its position in Database table above starting from Name column (common column as per lookup value). (Name – 1, Rate – 2, Hours – 3, Amount – 4, Sr. No. 5)
- **lookup_value:** cell $A10 i.e. name "Salman" (mixed cell reference). Here only column reference is fixed as column A should not change as we copy the formula in across the output table.
- **table_array:** In the table array, we will use CHOOSE Function as follows:
  - **Index Number:** As we want all 5 column values, our Index number will be 1, 2, 3, 4 and 5. In case, we want only 2 column values, then Index number will 1 and 2.

- **Value 1:** This value will be the common column between both the table, based on which it is going to search. In this scenario, it is Name column i.e. $B$2:$B$5
- **Value 2:** This will be the second column which we want in Output table - Rate column i.e. $C$2:$C$5
- **Value 3:** This will be the third column which we want in Output table - Hours column i.e. $D$2:$D$5
- **Value 4:** This will be the fourth column which we want in Output table - Amount column i.e. $E$2:$E$5
- **Value 5:** This will be the fifth column which we want in Output table – Sr. No. column i.e. $A$2:$A$5

- **col_index_num:** cell B$8 i.e. 1 above Name column in Output table (mixed cell reference). Here only row reference is fixed as $8 (row) should not change as we copy the formula in across the output table.
- **range_lookup:** FALSE (0) as we want exact match.
- If the lookup value is not found, it will return #N/A error. It can be seen that name “Aamir” is not not in the database table. Hence, #N/A error came in cell B12.

Once the above formula is created, we need to copy this formula in entire output table to get the values. By this method, we can extract values from Database table in each column of Output Table by a single formula.

5.10 TYPES OF ERRORS

In Excel, errors are returned when the user is inserting wrong data or asking the Excel to do something that can not be done. It is suggested that when Excel is giving an error, we must correct the error instead of using the error handling functions (like IFERROR, IFNA) to mask the error.

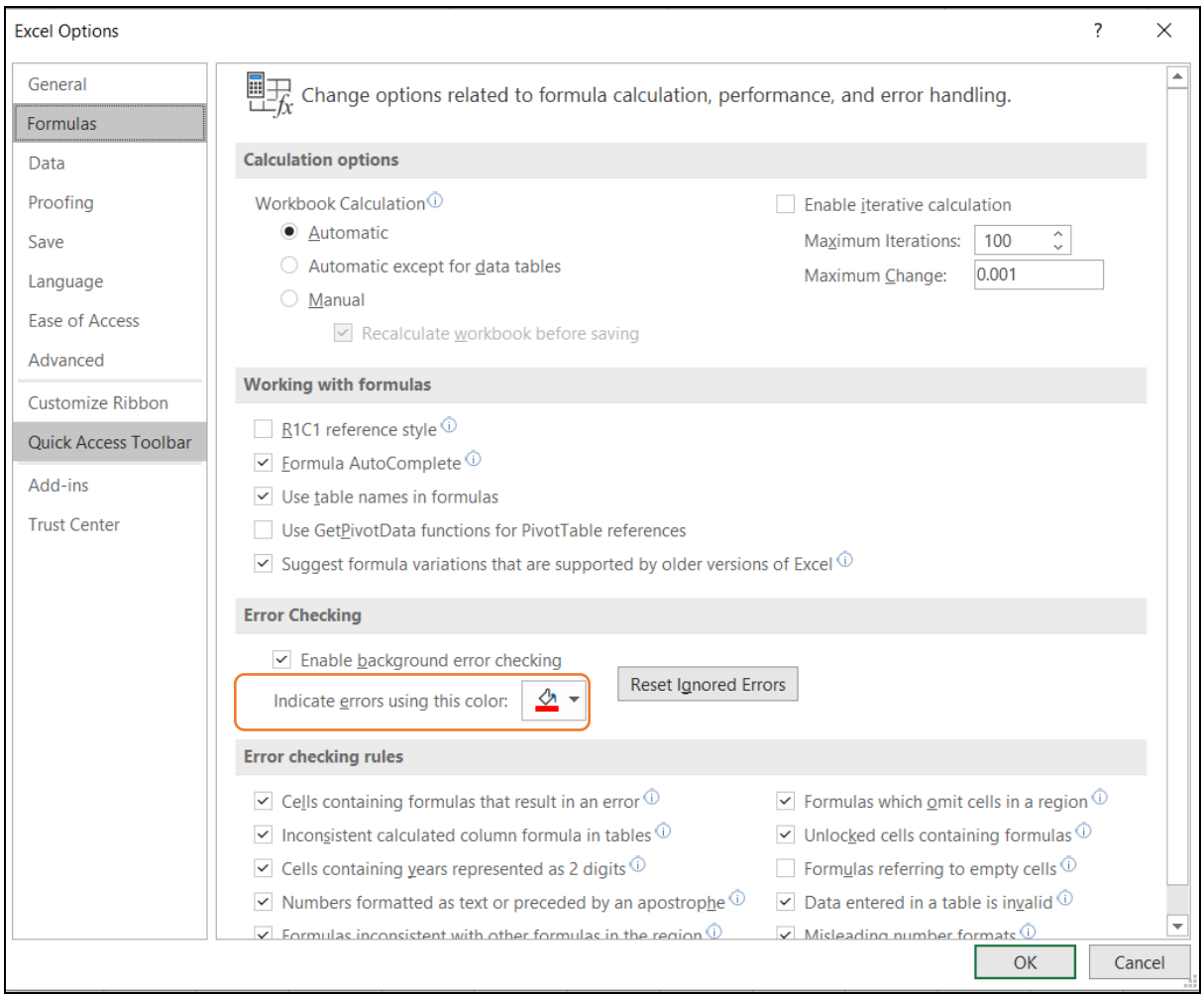
Excel errors are not just errors, but they are also a source of information about what is wrong with the function or the command that has been set for the execution.

Errors occur when we insert some formula in Excel and miss to add the required input in the expected forms, suppose if we have inserted function to add two cells then Excel expect that the cells would have numbers. If either of the cells has text in it, it would give an error. Every Excel function comes with its own terms and conditions and if any of the condition of the function is voided then there exist Excel errors. Every function has a syntax and this must be properly complied with, and if any deviation is observed in entering the syntax then there will be an Excel error.

Understanding Excel errors are important same as we understand the functions. These displayed errors tell us a lot of things. With the proper understanding about the Excel errors, one can easily solve those errors.

We can highlight the error using colours as shown below:

File > Options > Formulas, then select the colour of error highlighting. We have selected red colour for error highlighting.



(Figure 5.72 – Formula Options)

List of Errors:

- **#NAME?** error: This error appears when Excel can't find a named range. Excel assumes that any un-quoted string that isn't a function name is a named range. We will most likely encounter this when we forget to quote a string or mis-type a cell reference. For example:

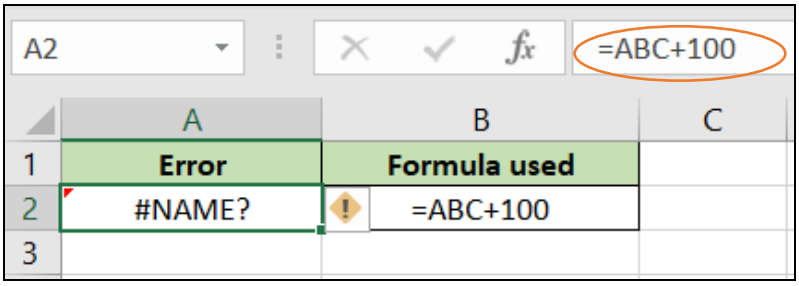


Figure 5.73 – NAME error

We may get #Name? error because the formula that is typed is incorrect. For example

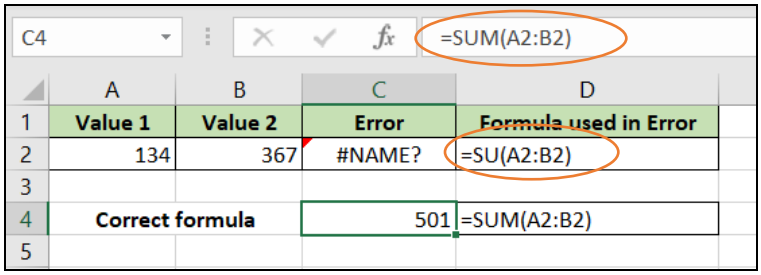


Figure 5.74 – NAME error

- **#DIV/0!** error: This error appears when a mathematical operation attempts to divide by zero or an empty cell. Usually, this occurs because a COUNT or SUM results in zero and another cell is operating on its result. A simple example is dividing any number by zero:

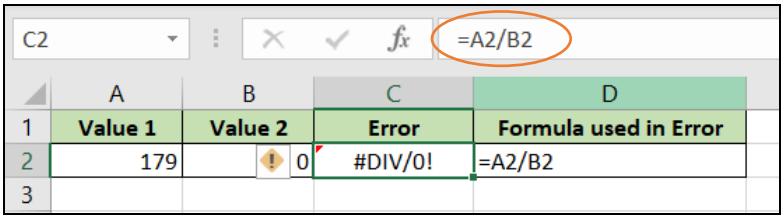


Figure 5.75 – DIVISION error

- **#REF!** error: This error appears when a cell reference is deleted or moved. Excel tries to automatically update all references, but when it can't do so, it replaces the actual cell reference with the error.

  — **Before error**

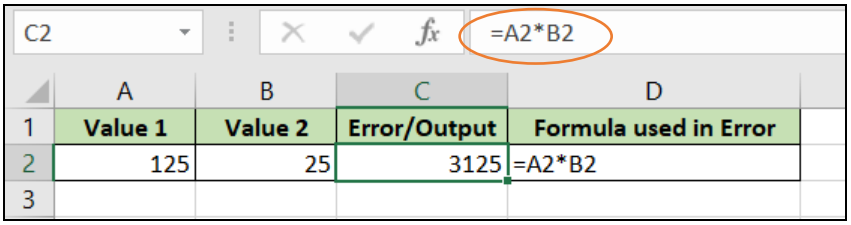


Figure 5.76 – REFERENCE error – before error

  — **After moving Value 1 column on Value 2 column**

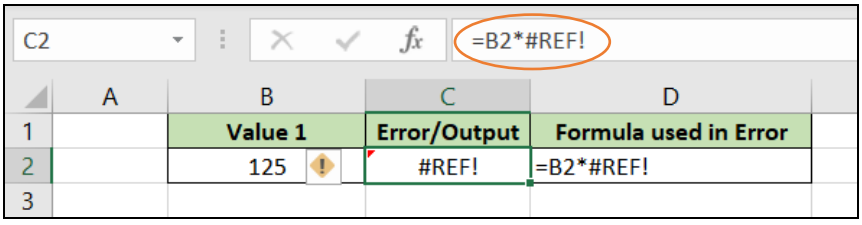


Figure 5.77 – REFERENCE error – after moving column

— **After deleting Value 2 column**

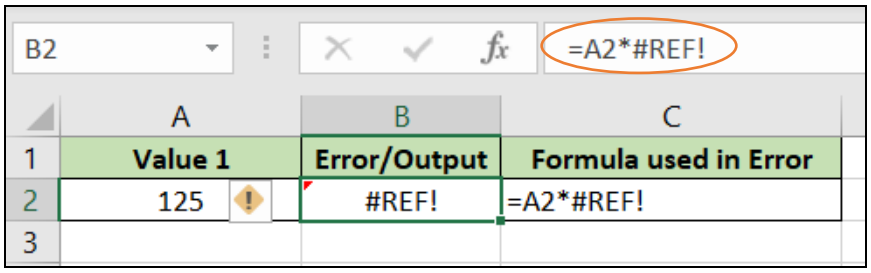


Figure 5.78 – REFERENCE error – after deleting column

- **#NULL!** error: This error is returned when Excel can't figure out the range specified in a cell. This happens when we put a space between function inputs instead of using a comma. For example:

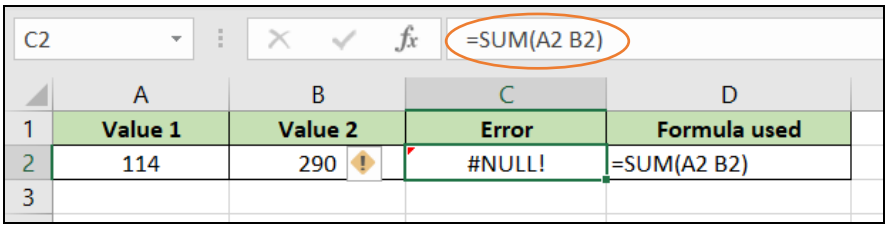


Figure 5.79 – NULL error

- **#N/A** error: This error is returned when a function like MATCH or VLOOKUP cannot find the value in the given range. For example:

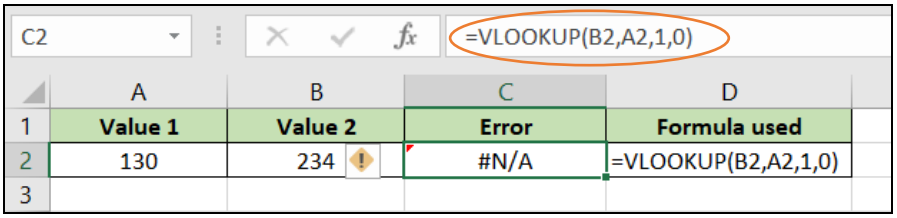


Figure 5.80 – N/A error

- **#VALUE!** error: This error occurs whenever the data type a function is given doesn't match what it is expecting. A simple example would be adding a text value to a number:

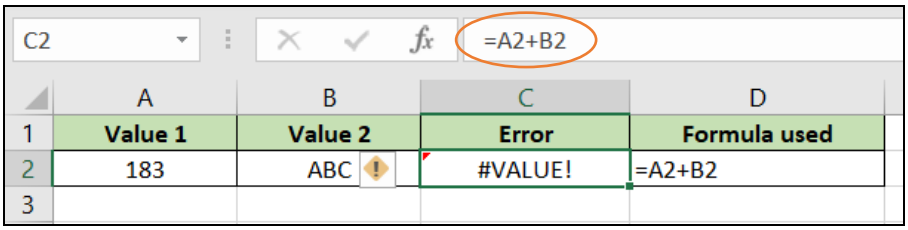


Figure 5.81 – VALUE error

- **#NUM!** error: This error happens when the displayed result is something that is not valid. Suppose if we

have entered a function i.e. 194^238, then the resulting number will be so long which will be not valid and we will get the "#NUM!" error.

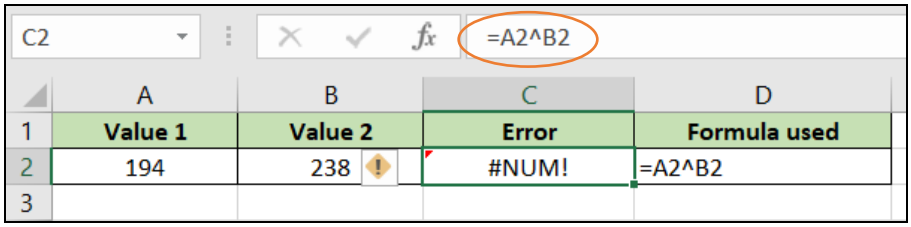


Figure 5.82 – NUM error

- **########** error: This not an error as such but the column is not wide enough to display all the characters in a cell. Hence, we can either extend the width of the column or do format cell.

— **Before applying solution**

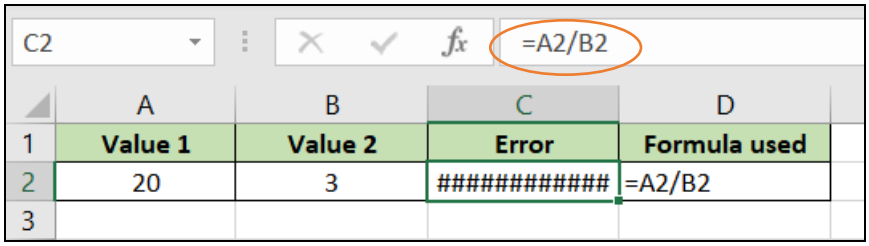


Figure 5.83 – Hash error

— **After expanding C column width**

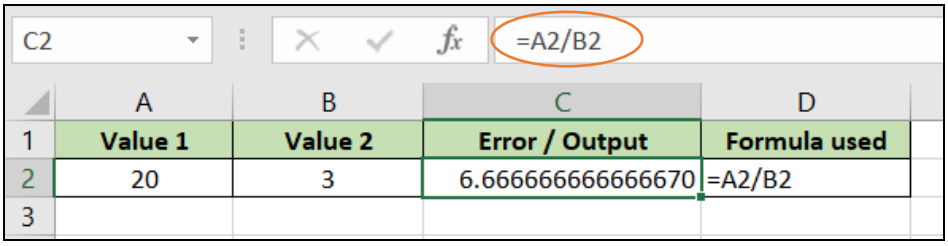


Figure 5.84 – Hash error – after expanding column width

— **After formatting cell C2 (Number format to 2 decimal)**

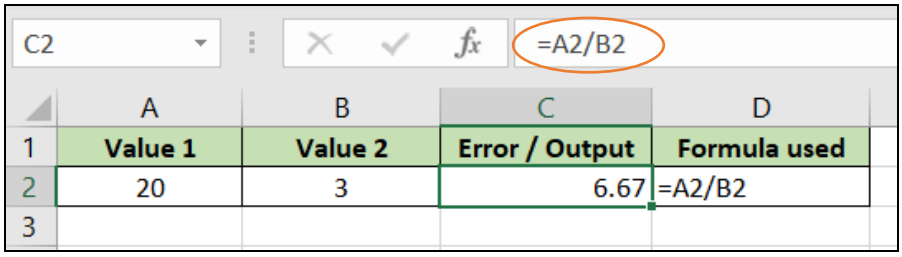


Figure 5.85 – Hash error – after cell formatting

5.11 FORMULA AUDITING TOOLBARS

The Formula Tab in Excel contains Formula Auditing options, mainly Trace Precedents, Trace Dependents, Error Checking, Evaluate Formula and Watch Window.

- **Trace Precedents and Trace Dependents:** With Excel, we can create some fairly complex solutions. A cell can contain a formula that uses values from multitudes of other cells and ranges. Working through long, complex formulas to track down problems can be quite tedious.

  Formulas may contain precedents and may serve as dependents to other formulas:

  - Precedents are cells or ranges that affect the active cell's value.
  - Dependents are cells or ranges affected by the active cell.

  It's all relative. A cell often serves as both a precedent and a dependent.

The example is as follows:

C28 =C17-C26

	A	B	C	D
1		**Profit & Loss A/c**		
2		**Particulars**	**Actual FY2020**	**Formula used**
4				
5		**Sales**	**10,00,000**	
6				
7	**Less :**	Sales Return	10,000	
8				
9		**Net Sales**	**9,90,000**	=C5-C7
10				
11		**Increase / (Decrease) in Stock**	**5,000**	=C13-C12
12		Opening Stock	10,000	
13		Closing Stock	15,000	
14				
15		**Raw Material Consumption**	**50,000**	
16				
17		**Gross Profit**	**9,45,000**	=C9-C15+C11
18				
19	**Less:**	Manufacturing Expenses	4,70,000	
20		Power & Fuel	1,00,000	
21		Repair & Maintenanace	50,000	
22		Salaries	1,00,000	
23		Administrative Exp.	75,000	
24		Depreciation	50,000	
25				
26		**Expenses**	**8,45,000**	=SUM(C19:C24)
27				
28		**Net Profit / (Loss)**	**1,00,000**	=C17-C26
29				

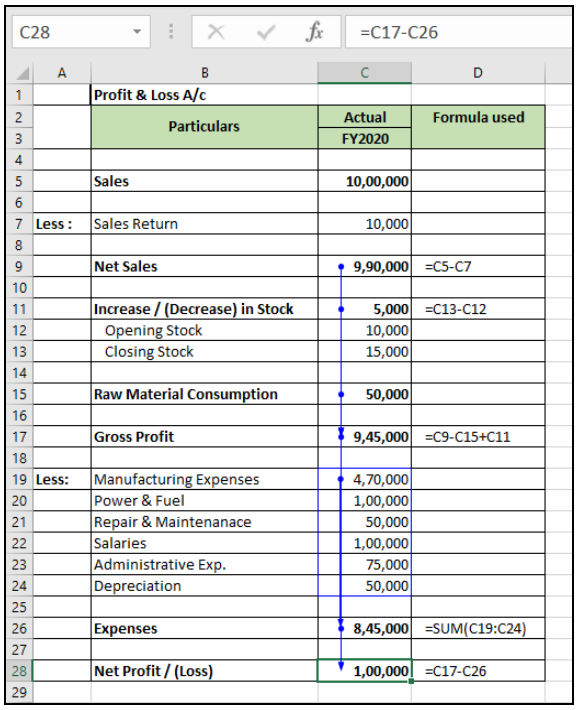
Figure 5.86 – Profit & Loss A/c Table

C28 =C17-C26

	A	B	C	D
1		**Profit & Loss A/c**		
2		**Particulars**	**Actual FY2020**	**Formula used**
4				
5		**Sales**	**10,00,000**	
6				
7	**Less :**	Sales Return	10,000	
8				
9		**Net Sales**	**9,90,000**	=C5-C7
10				
11		**Increase / (Decrease) in Stock**	**5,000**	=C13-C12
12		Opening Stock	10,000	
13		Closing Stock	15,000	
14				
15		**Raw Material Consumption**	**50,000**	
16				
17		**Gross Profit**	**9,45,000**	=C9-C15+C11
18				
19	**Less:**	Manufacturing Expenses	4,70,000	
20		Power & Fuel	1,00,000	
21		Repair & Maintenanace	50,000	
22		Salaries	1,00,000	
23		Administrative Exp.	75,000	
24		Depreciation	50,000	
25				
26		**Expenses**	**8,45,000**	=SUM(C19:C24)
27				
28		**Net Profit / (Loss)**	**1,00,000**	=C17-C26
29				

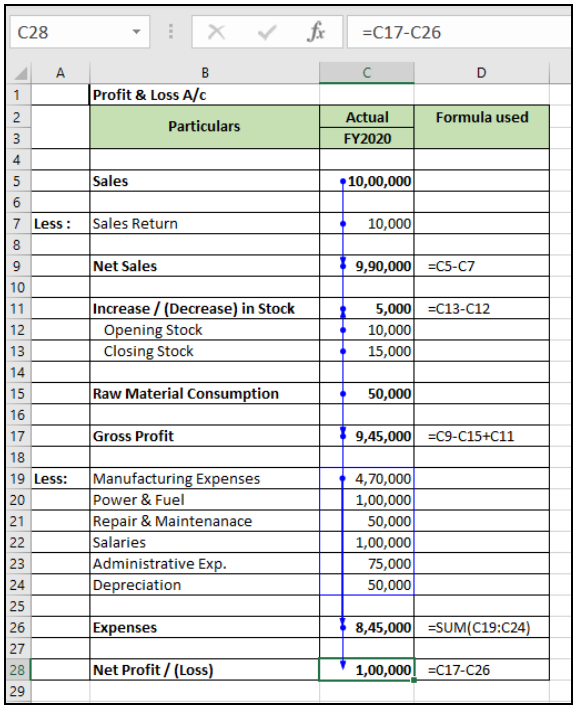
Figure 5.87 – first precedent connection

In the above example, we can see there are many functions/formulas used. If we want to start Trace Precedents from Net Profit/(Loss) and click once, it will be show arrows from C17 and C26 as shown in figure 5.87 above.



C28	=C17-C26		
	A	B	C	D
1		**Profit & Loss A/c**		
2		**Particulars**	**Actual**	**Formula used**
3			**FY2020**	
4				
5		**Sales**	**10,00,000**	
6				
7	**Less :**	Sales Return	10,000	
8				
9		**Net Sales**	**9,90,000**	=C5-C7
10				
11		**Increase / (Decrease) in Stock**	**5,000**	=C13-C12
12		Opening Stock	10,000	
13		Closing Stock	15,000	
14				
15		**Raw Material Consumption**	**50,000**	
16				
17		**Gross Profit**	**9,45,000**	=C9-C15+C11
18				
19	**Less:**	Manufacturing Expenses	4,70,000	
20		Power & Fuel	1,00,000	
21		Repair & Maintenanace	50,000	
22		Salaries	1,00,000	
23		Administrative Exp.	75,000	
24		Depreciation	50,000	
25				
26		**Expenses**	**8,45,000**	=SUM(C19:C24)
27				
28		**Net Profit / (Loss)**	**1,00,000**	=C17-C26
29				

Figure 5.88 – second precedent connection

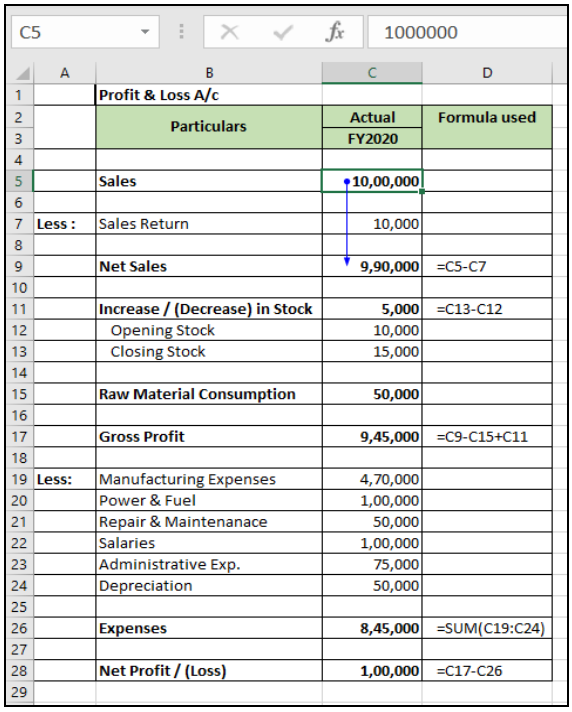


C28	=C17-C26		
	A	B	C	D
1		**Profit & Loss A/c**		
2		**Particulars**	**Actual**	**Formula used**
3			**FY2020**	
4				
5		**Sales**	**10,00,000**	
6				
7	**Less :**	Sales Return	10,000	
8				
9		**Net Sales**	**9,90,000**	=C5-C7
10				
11		**Increase / (Decrease) in Stock**	**5,000**	=C13-C12
12		Opening Stock	10,000	
13		Closing Stock	15,000	
14				
15		**Raw Material Consumption**	**50,000**	
16				
17		**Gross Profit**	**9,45,000**	=C9-C15+C11
18				
19	**Less:**	Manufacturing Expenses	4,70,000	
20		Power & Fuel	1,00,000	
21		Repair & Maintenanace	50,000	
22		Salaries	1,00,000	
23		Administrative Exp.	75,000	
24		Depreciation	50,000	
25				
26		**Expenses**	**8,45,000**	=SUM(C19:C24)
27				
28		**Net Profit / (Loss)**	**1,00,000**	=C17-C26
29				

Figure 5.89 – All precedent connections

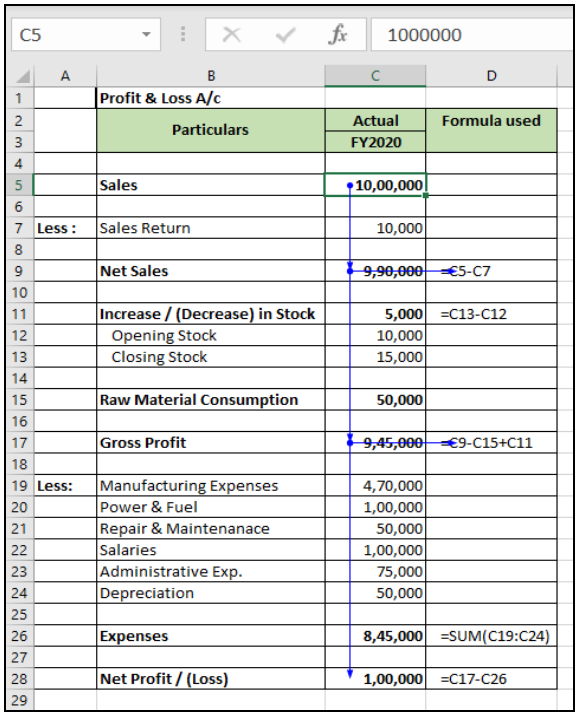
If, without changing cell location, we click again the Trace Precedents, it will take to other cells which are linked to cell C17 and C26 as shown below in Figure 5.88 above. Like this, if we keep on clicking, it will take to last cell as shown in Figure 5.89 above.

In the same manner, we can **Trace Dependents**. In the below Figure 5.90, we clicked at cell C5. Then clicked once on Trace Dependent button in Formula Tab. It will show arrow to all the cells which are dependent on cell C5. Then if we click again on Trace Dependent, it will take next dependent cells as shown in Figure 5.91 below.



	A	B	C	D
1		Profit & Loss A/c		
2		Particulars	Actual FY2020	Formula used
3				
4				
5		Sales	10,00,000	
6				
7	Less :	Sales Return	10,000	
8				
9		Net Sales	9,90,000	=C5-C7
10				
11		Increase / (Decrease) in Stock	5,000	=C13-C12
12		Opening Stock	10,000	
13		Closing Stock	15,000	
14				
15		Raw Material Consumption	50,000	
16				
17		Gross Profit	9,45,000	=C9-C15+C11
18				
19	Less:	Manufacturing Expenses	4,70,000	
20		Power & Fuel	1,00,000	
21		Repair & Maintenanace	50,000	
22		Salaries	1,00,000	
23		Administrative Exp.	75,000	
24		Depreciation	50,000	
25				
26		Expenses	8,45,000	=SUM(C19:C24)
27				
28		Net Profit / (Loss)	1,00,000	=C17-C26
29				

Figure 5.90 – First dependent



	A	B	C	D
1		Profit & Loss A/c		
2		Particulars	Actual FY2020	Formula used
3				
4				
5		Sales	10,00,000	
6				
7	Less :	Sales Return	10,000	
8				
9		Net Sales	9,90,000	=C5-C7
10				
11		Increase / (Decrease) in Stock	5,000	=C13-C12
12		Opening Stock	10,000	
13		Closing Stock	15,000	
14				
15		Raw Material Consumption	50,000	
16				
17		Gross Profit	9,45,000	=C9-C15+C11
18				
19	Less:	Manufacturing Expenses	4,70,000	
20		Power & Fuel	1,00,000	
21		Repair & Maintenanace	50,000	
22		Salaries	1,00,000	
23		Administrative Exp.	75,000	
24		Depreciation	50,000	
25				
26		Expenses	8,45,000	=SUM(C19:C24)
27				
28		Net Profit / (Loss)	1,00,000	=C17-C26
29				

Figure 5.91 – All dependent connections

In Formula Tab, there is option "Remove Arrows" as shown below:

- If we click on option "Remove Arrows", then it will remove all the arrows from the screen.
- If we click on option "Remove Precedent Arrows", then it will remove last precedent arrows.
- If we click on option "Remove Dependent Arrows", then it will remove last dependent arrows.

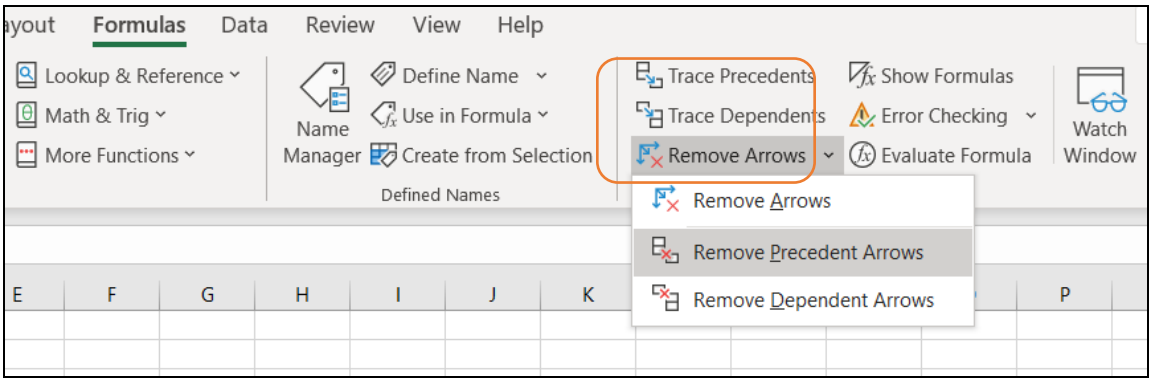


Figure 5.92 – Remove Arrows

- **Evaluate Formula:** The Evaluate Formula dialog box (under Formula Tab) helps to walk through the sequential steps used in calculating a result from a formula. These steps are useful for tracking down errors in formulas that are long or have precedents. For example in below Figure 5.93, we have used

Index Function. Using the Evaluate Formula dialog box makes it easy to see how Excel works out this formula. The step-by-step approach helps to see what is done at each step.

Figure 5.93 shows the Evaluate Formula dialog box at the start of evaluating the formula. To display the Evaluate Formula dialog box, simply click the Evaluate Formula button in the Formula Auditing group of the Formulas tab of the Ribbon. With each successive click of the Evaluate button, the Evaluation box displays the interim results. The Step In and Step Out buttons are enabled during the steps that work on the precedents.

The Evaluate Formula dialog box is great for really seeing how each little step feeds into the final calculated result. Using this dialog box lets us pinpoint exactly where a complex formula has gone sour.

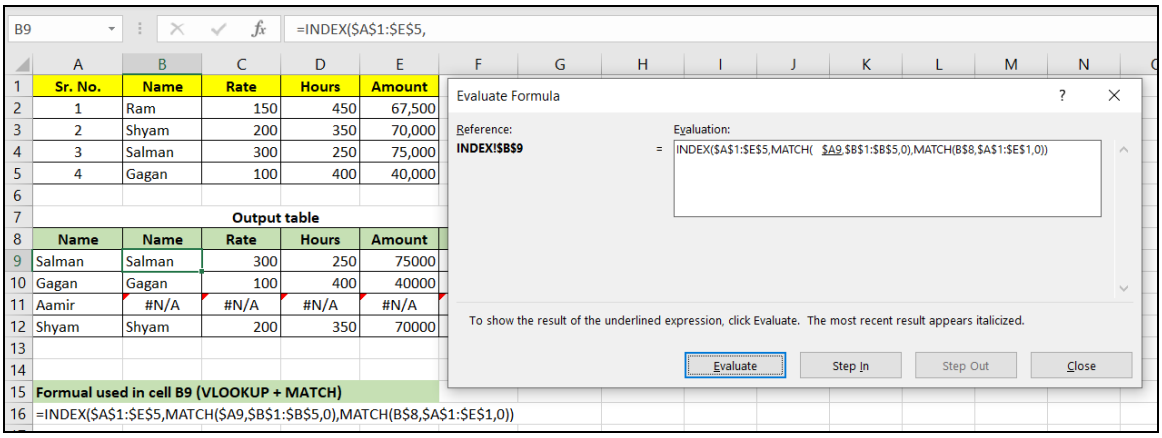


Figure 5.93 – Evaluate Formula

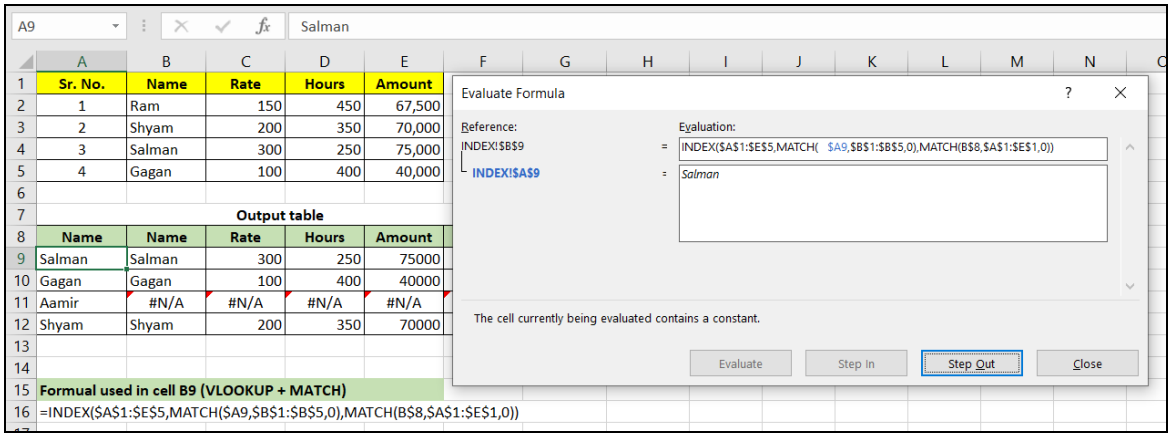
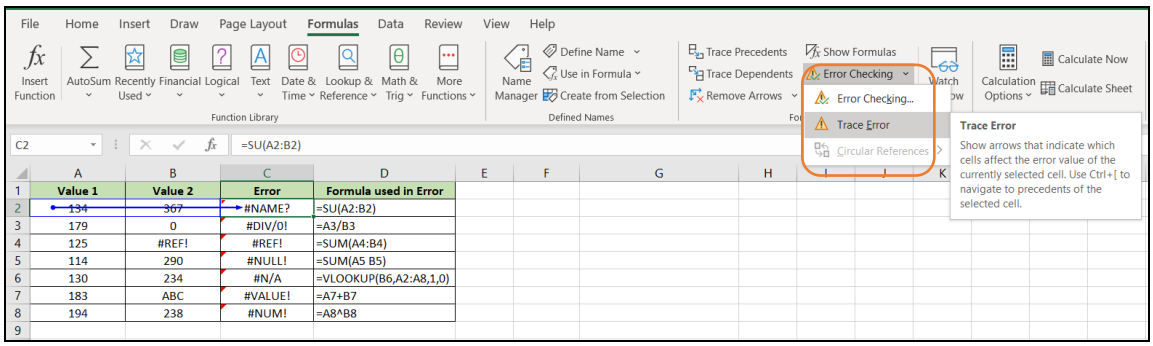


Figure 5.94 – Evaluate Formula – Step in & Step Out

- **Error Checking:** A similar error-hunting tool is the Error Checking dialog box, shown in Figure 5.95 below.



	A	B	C	D
1	Value 1	Value 2	Error	Formula used in Error
2	134	367	#NAME?	=SU(A2:B2)
3	179	0	#DIV/0!	=A3/B3
4	125	#REF!	#REF!	=SUM(A4:B4)
5	114	290	#NULL!	=SUM(A5 B5)
6	130	234	#N/A	=VLOOKUP(B6,A2:A8,1,0)
7	183	ABC	#VALUE!	=A7+B7
8	194	238	#NUM!	=A8^B8
9				

Figure 5.95 – Error Checking Option – Trace Error

First we select the cell which has error, then click on Error Checking option as shown in Figure 5.96 above. It shows many options/buttons to analyze the error and make decisions about it:

- **Help on This Error** : starts the Excel Help system.
- **Show Calculation Steps** : opens the Evaluate Formula dialog box.
- **Ignore Error** : ensures that Excel ignores the error. The cell may still display an error symbol, but Excel does not give a hoot, and we probably won't either, because we clicked the button.
- **Edit in Formula Bar** : places the cursor in the Formula Bar, making it easy to edit the formula.
- **Options** : opens the Excel Options dialog box.
- **Previous and Next** : cycle through the multiple errors on the worksheet, assuming that there is more than one error.

The Error Checking drop-down menu hosts the Trace Error command. Only precedents are pointed out by the tracer lines. This makes it easy to see the cells that feed into a cell that has an error.

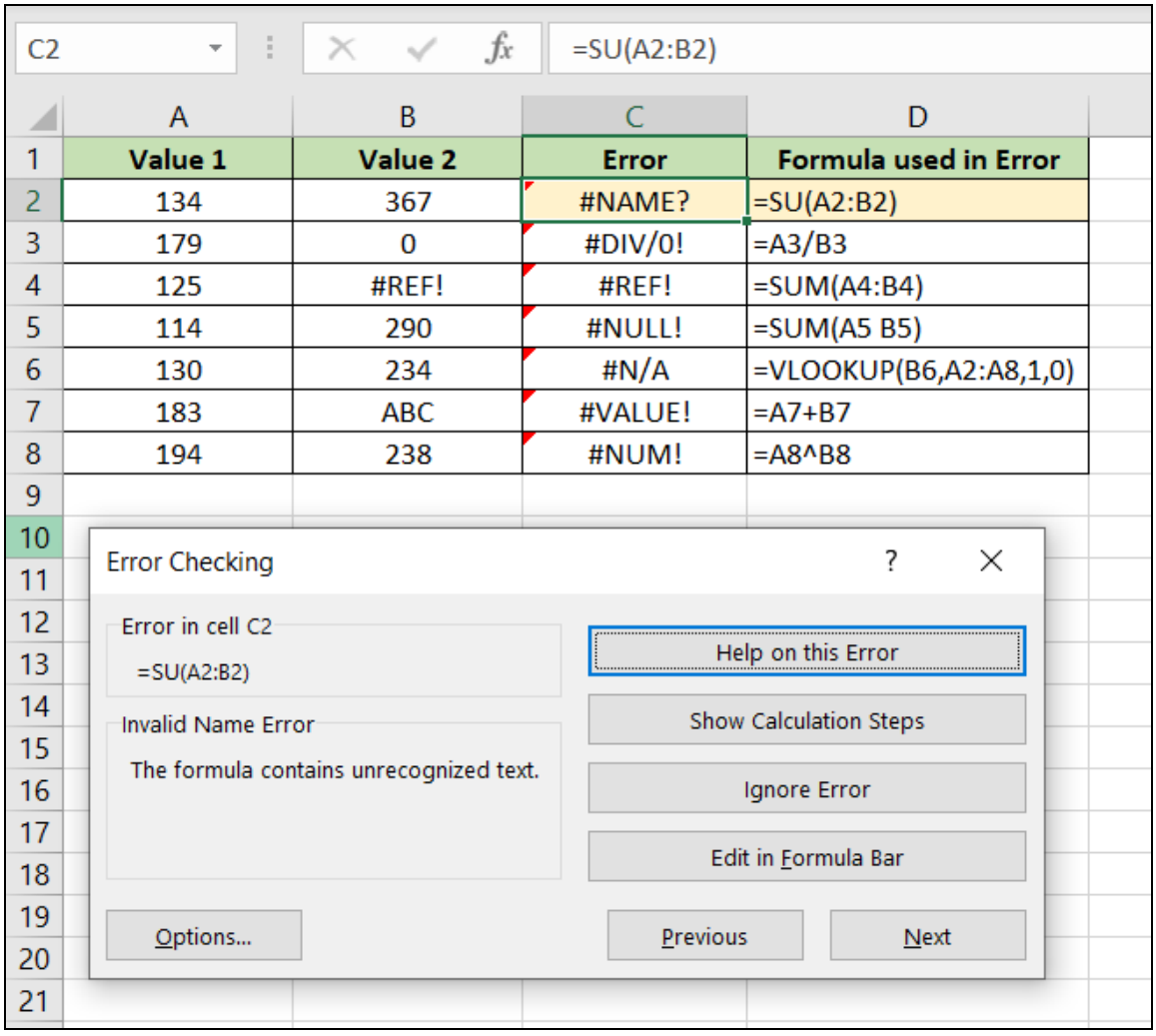


Figure 5.96 – Error Checking Option

- **Using Watch Window:** The Watch Window lets us watch the calculated results of a formula but without the limitation of having the cell be in the viewing area of Excel. This feature is helpful when we're working on correcting formulas that use precedents that are scattered about the worksheet or workbook.

  First, to set up a watch, follow these steps:

  i. Click the Watch Window button on the Formulas tab of the Ribbon.

  ii. In the Watch Window, click the **Add Watch** button. The Add Watch dialog box opens.

  iii. Then specify the cell(s) or type in the cell address or range.

  iv. Click the **Add** button in the Add Watch dialog box to complete setting up the watch.

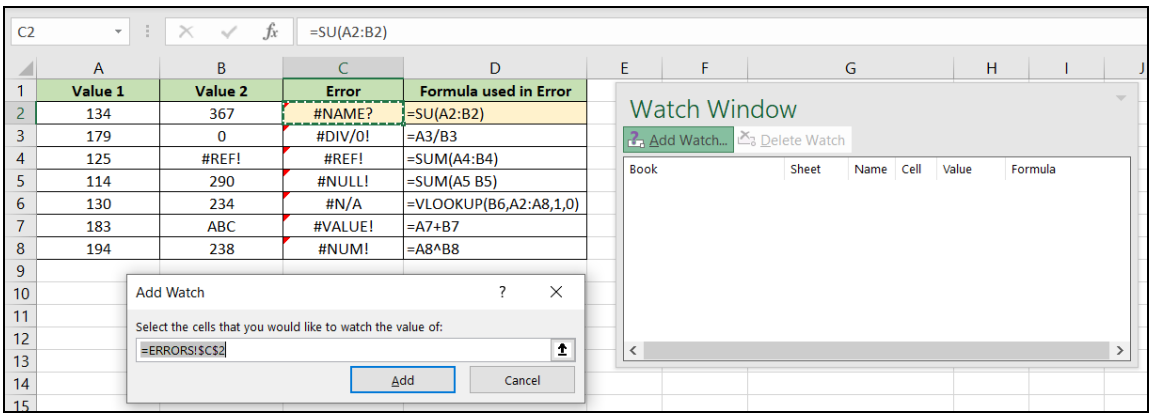


Figure 5.97 – Add Watch option

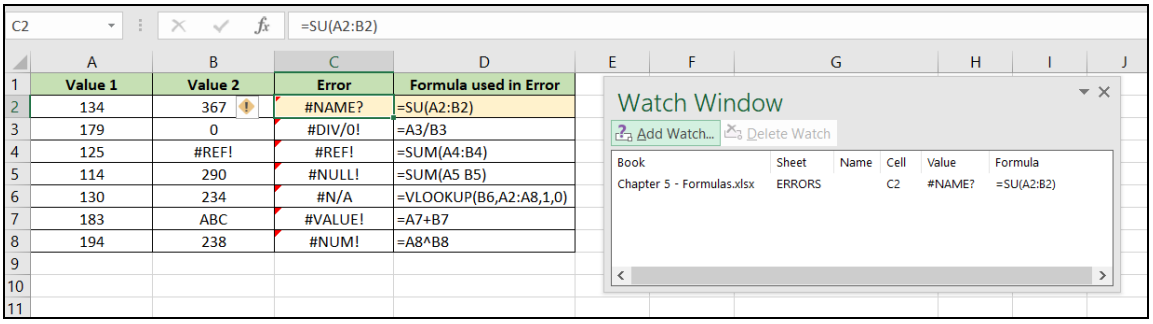


Figure 5.98 – Watch Window – after Add Watch

Figure 5.97 shows the Watch Window with a watch in place. Cell C2 of the ERRORS worksheet is being watched. The Watch Window sits on top of the workbook and stays visible regardless of which worksheet is active. This means, for example, that we could try different values in the formula in ERRORS worksheet and see the result in the calculation in C2, but without having to bounce around the worksheets to see how new values alter the calculated result. In the below figure, we can see that we have corrected the SUM Function and the same is visible in WATCH WINDOW.

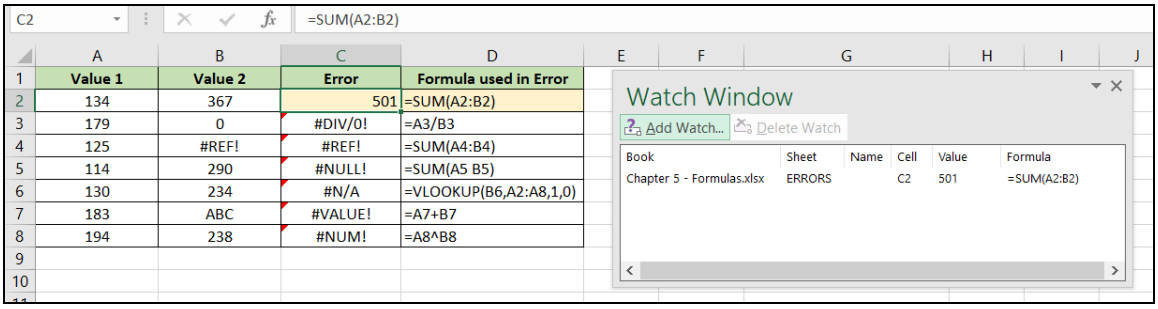


Figure 5.99 – Watch Window – after Add Watch

The Watch Window also allows to delete a watch. To delete a watch, we need to perform following steps:

i. Select a watch from the list of watches in the Watch Window.

ii. Click the Delete Watch button.

5.12 DATABASE FUNCTIONS

A database table has fields and records like we have rows and columns in an Excel worksheet. Here, we are going to discuss Excel's database functions, such as DGET, DSUM, DCOUNT, DAVERAGE.

To use the database functions, we need to put data into a structured format. To make the best of the database functions, we need to get data into a contiguous area of rows and columns. Each row is a record, and each column is a field. The top row contains labels that identify the fields.

A great way to work with the database functions is to name the database area and then enter the name, instead of the range address, in the function. Later, if records are added to the bottom of the database, we have to redefine the named area's range to include the new rows. If we add records to database range by inserting new rows somewhere in the middle, rather than adding them on at the end, Excel automatically adjusts the reference to the named range.

Criteria Area: In database function, we need to define the criteria as well. The criteria are not part of the database function arguments but are somewhere in the worksheet and then referenced by the function. The criteria area can contain a single criterion, or it can contain two or more criteria. Each individual criterion is structured as follows:

- In one cell, enter the field name (header) of the database column that the criterion will apply to.
- In the cell below, enter the value that the field data must meet.

Criteria options: The criteria can include a variety of expressions. The table below shows some examples:

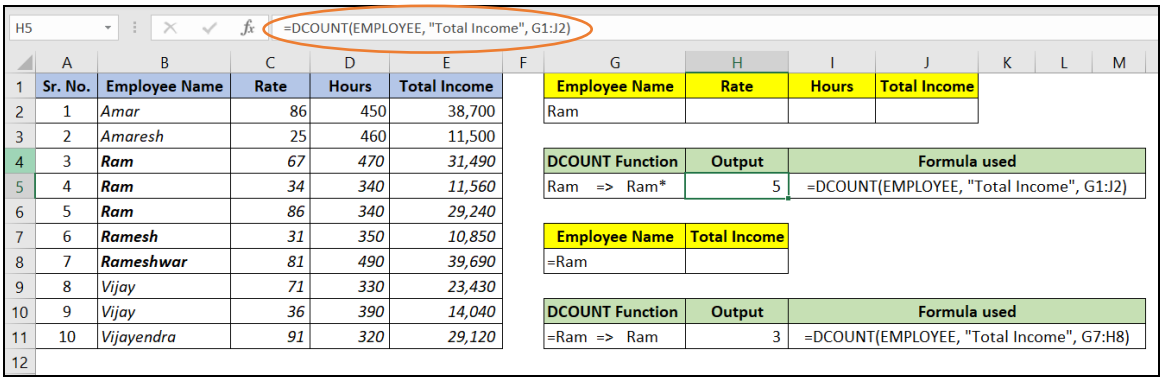
Criteria	Behaviour
Ram	Begins with Ram*
=Ram	Match "RAM" or "ram" or "Ram"
10	Equal to 10
>10	Greater than 10
<>	Not Blank
<>10	Not 10

The criteria range can include more than one row below the headers. When criteria include more than one row, each row is joined with OR logic, and the expressions in a given criteria row are joined with AND logic.

- **DCOUNT Function:** This function counts the cells containing numbers in the field (column) of the records in the database that matches the condition we specify. The general syntax is as follows:

  =DCOUNT(database, field, criteria)

The example is as follows:



	A	B	C	D	E
1	Sr. No.	Employee Name	Rate	Hours	Total Income
2	1	Amar	86	450	38,700
3	2	Amaresh	25	460	11,500
4	3	Ram	67	470	31,490
5	4	Ram	34	340	11,560
6	5	Ram	86	340	29,240
7	6	Ramesh	31	350	10,850
8	7	Rameshwar	81	490	39,690
9	8	Vijay	71	330	23,430
10	9	Vijay	36	390	14,040
11	10	Vijayendra	91	320	29,120
12					

Employee Name	Rate	Hours	Total Income
Ram			

DCOUNT Function	Output	Formula used
Ram => Ram*	5	=DCOUNT(EMPLOYEE, "Total Income", G1:J2)

Employee Name	Total Income
=Ram	

DCOUNT Function	Output	Formula used
=Ram => Ram	3	=DCOUNT(EMPLOYEE, "Total Income", G7:H8)

Figure 5.100 – DCOUNT Function – Begins with and Exact match

In the above figure, we have 2 examples.

Process followed In First example:

- First we gave name to entire database table as “EMPLOYEE”.
- Then we created a Criteria table with 4 columns i.e. Employee Name, Rate, Hours and Total Income (Table Highlighted in Yellow colour)
- Then in cell G2, we typed “Ram” below Employee Name heading in criteria table.
- Then we typed the function in cell H5 as follows:

  =DCOUNT(database, field, criteria)

  where,

  - database : EMPLOYEE
  - field : “Total Income”, (we can select cell also. We also can select any other field heading as well)
  - criteria : the criteria table i.e. G1:J2
- Output: The function returns count as 5 as it matches all the values with begins with “Ram”. So Ramesh and Rameshwar also considered in count.

Process followed In second example (exact match count):

- First we gave name to entire database table as “EMPLOYEE”.
- Then we created a Criteria table with 2 columns i.e. Employee Name and Total Income (Table Highlighted in Yellow colour)
- Then in cell G8, we typed “=Ram” below Employee Name heading in criteria table.
- Then we typed the function in cell H11 as follows:

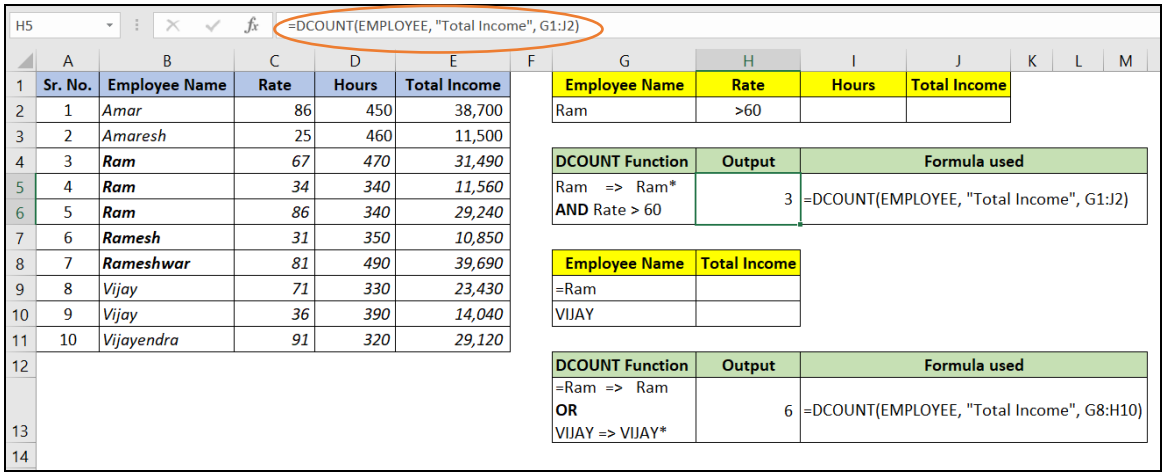
  =DCOUNT(database, field, criteria)

where,

- database : EMPLOYEE
- field : “Total Income”, (we can select cell also. We also can select any other field heading as well)
- criteria : the criteria table i.e. G7:H8

- Output: The function returns count as 3 as it considers only 100% matches for count.

Now we are applying DCOUNT function and AND / OR concept.



Sr. No.	Employee Name	Rate	Hours	Total Income
1	*Amar*	86	450	38,700
2	*Amaresh*	25	460	11,500
3	***Ram***	*67*	*470*	*31,490*
4	***Ram***	*34*	*340*	*11,560*
5	***Ram***	*86*	*340*	*29,240*
6	***Ramesh***	*31*	*350*	*10,850*
7	***Rameshwar***	*81*	*490*	*39,690*
8	*Vijay*	*71*	*330*	*23,430*
9	*Vijay*	*36*	*390*	*14,040*
10	*Vijayendra*	*91*	*320*	*29,120*

Employee Name	Rate	Hours	Total Income
Ram	>60		

DCOUNT Function	Output	Formula used
Ram => Ram* **AND** Rate > 60	3	=DCOUNT(EMPLOYEE, "Total Income", G1:J2)

Employee Name	Total Income
=Ram	
VIJAY	

DCOUNT Function	Output	Formula used
=Ram => Ram **OR** VIJAY => VIJAY*	6	=DCOUNT(EMPLOYEE, "Total Income", G8:H10)

Figure 5.101 – DCOUNT Function – AND / OR

In the above figure, we have 2 examples.

In the first example, the Employee Name is “Ram” and Rate > 60. As both conditions are mentioned in same row, it works on AND concept. It means it will return output only when all the conditions are fulfilled.

- There are 5 names which begins with Ram but only 3 Rates are >60.
- Final output in cell H5 = 3

In second example, we have entered 2 names i.e. “=Ram” and “VIJAY” in 2 rows below Employee Name heading in criteria table. As conditions are in 2 rows, it acts on OR concept.

- There only 3 names which 100% match with “=Ram” condition. There are 3 names which begins with “VIJAY”.
- Final Output in cell H13 = 6

- **DSUM Function:** This function adds the numbers in the field (column) of records in the database that match the conditions we specify. The general syntax is as follows:

  =DSUM (database, field, criteria)

The 2 DSUM example is as follows (Figure 5.102):

H5 =DSUM(EMPLOYEE, "Total Income", G1:J2)

Sr. No.	Employee Name	Rate	Hours	Total Income
1	*Amar*	86	450	38,700
2	*Amaresh*	25	460	11,500
3	***Ram***	*67*	*470*	*31,490*
4	***Ram***	*34*	*340*	*11,560*
5	***Ram***	*86*	*340*	*29,240*
6	***Ramesh***	*31*	*350*	*10,850*
7	***Rameshwar***	*81*	*490*	*39,690*
8	*Vijay*	*71*	*330*	*23,430*
9	*Vijay*	*36*	*390*	*14,040*
10	*Vijayendra*	*91*	*320*	*29,120*

Employee Name	Rate	Hours	Total Income
Ram			

DSUM Function	Output	Formula used
Ram => Ram*	1,22,830	=DSUM(EMPLOYEE, "Total Income", G1:J2)

Employee Name	Total Income
=Ram	

DSUM Function	Output	Formula used
=Ram => Ram	72,290	=DSUM(EMPLOYEE, "Total Income", G7:H8)

Figure 5.102 – DSUM Function – Begins with and Exact match

In the first example, the Employee Name is "Ram". The function will do the sum for all names which begins with name "Ram". It will also include "Ramesh" and "Rameshwar"

— There are 5 names which begins with Ram. Hence, it will do sum of all the corresponding Total Income.

— Final output in cell H5 = 1,22,830

In second example, we have name as "=Ram" below Employee Name heading in criteria table.

— There are only 3 names which 100% match with "=Ram" condition. Hence, it will do sum of all the corresponding Total Income.

— Final Output in cell H11 = 72,290

H5 =DSUM(EMPLOYEE, "Total Income", G1:J2)

Sr. No.	Employee Name	Rate	Hours	Total Income
1	*Amar*	86	450	38,700
2	*Amaresh*	25	460	11,500
3	*Ram*	*67*	*470*	*31,490*
4	*Ram*	*34*	*340*	*11,560*
5	*Ram*	*86*	*340*	*29,240*
6	*Ramesh*	*31*	*350*	*10,850*
7	*Rameshwar*	*81*	*490*	*39,690*
8	*Vijay*	*71*	*330*	*23,430*
9	*Vijay*	*36*	*390*	*14,040*
10	*Vijayendra*	*91*	*320*	*29,120*

Employee Name	Rate	Hours	Total Income
Ram	>60	<>340	

DSUM Function	Output	Formula used
Ram => Ram* **AND** Rate > 60	71,180	=DSUM(EMPLOYEE, "Total Income", G1:J2)

Employee Name	Rate	Hours	Total Income
=Ram	>60	<>340	
VIJAY	>75		

DSUM Function	Output	Formula used
=Ram => Ram **OR** VIJAY => VIJAY*	60,610	=DSUM(EMPLOYEE, "Total Income", G8:J10)

Figure 5.103 – DSUM Function – AND / OR

In the above Figure 5.103, we have 2 examples.

In the first example, we have 3 conditions i.e. the Employee Name is "Ram", Rate > 60 and Hours <> 340. As all the 3 conditions are mentioned in same row, it works on AND concept. It means it will return output only when all the conditions are fulfilled.

— There are 5 names which begins with Ram but only 3 Rates are >60 and only 2 values in Hours columns are <>340.

— Final output in cell H5 = 71,180

In second example, we have entered 2 names i.e. "=Ram" and "VIJAY" in 2 rows below Employee Name heading in criteria table. Also in Rate column, 2 conditions i.e. >60 and >75 written. In Hours column, another condition i.e. <> 340 written. As conditions are in 2 rows, it acts on OR concept and all conditions on same row will act on AND concept.

— There only 3 names which 100% match with "=Ram" condition but only 2 Rates > 60 and only 1 Hour value <>340. Hence, Output – 31,490.

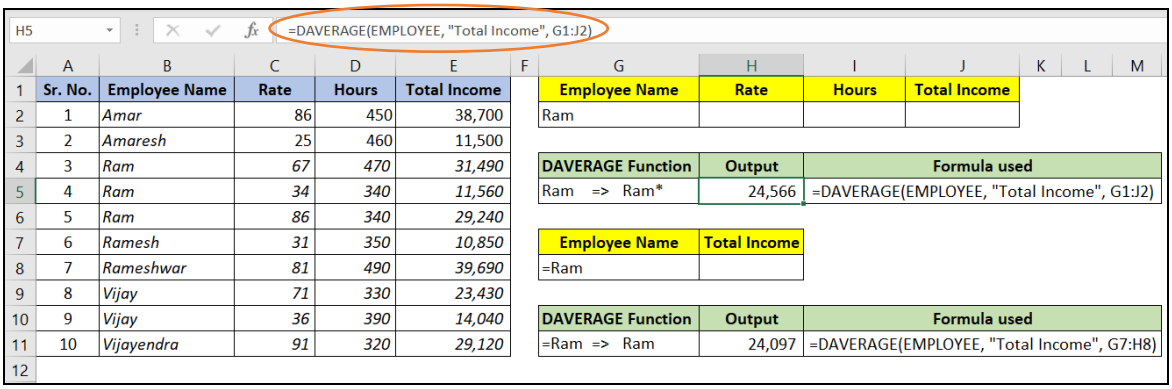
— There are 3 names with begin with "VIJAY" condition but only 1 Rates value > 75. Hence, Output – 29,120

— Final Output in cell H13 = 60,610 (sum of 31490+29120)

- **DAVERAGE Function:** This function averages the values in a column in a list or database that match the conditions we specify. The general syntax is as follows:

=DAVERAGE(database, field, criteria)

The 2 DAVERAGE example is as follows (Figure 5.104):



	A	B	C	D	E
1	Sr. No.	Employee Name	Rate	Hours	Total Income
2	1	Amar	86	450	38,700
3	2	Amaresh	25	460	11,500
4	3	Ram	67	470	31,490
5	4	Ram	34	340	11,560
6	5	Ram	86	340	29,240
7	6	Ramesh	31	350	10,850
8	7	Rameshwar	81	490	39,690
9	8	Vijay	71	330	23,430
10	9	Vijay	36	390	14,040
11	10	Vijayendra	91	320	29,120
12					

Employee Name	Rate	Hours	Total Income
Ram			

DAVERAGE Function	Output	Formula used
Ram => Ram*	24,566	=DAVERAGE(EMPLOYEE, "Total Income", G1:J2)

Employee Name	Total Income
=Ram	

DAVERAGE Function	Output	Formula used
=Ram => Ram	24,097	=DAVERAGE(EMPLOYEE, "Total Income", G7:H8)

Figure 5.104 – DSUM Function – Begins with and Exact match

In the first example, the Employee Name is "Ram". The function will calculate average for all names which begins with name "Ram". It will also include "Ramesh" and "Rameshwar"

— There are 5 names which begins with Ram. Hence, it will do average of all the corresponding Total Income.

— Final output in cell H5 = 24,566

In second example, we have name as "=Ram" below Employee Name heading in criteria table.

— There are only 3 names which 100% match with "=Ram" condition. Hence, it will do average of all the corresponding Total Income.

— Final Output in cell H11 = 24,097

- **DGET Function:** The DGET Function is a unique database function. It does not perform a calculation but checks for duplicate entries. The function returns one of three values:
  - If one record matches the criterion, DGET returns the criterion.
  - If more than one record matches the criterion, DGET returns the #NUM! error.
  - If no records match the criterion, DGET returns the #VALUE! error.

  The general syntax is as follows:

  =DGET(database, field, criteria)

  There are 3 examples in Figure 5.105 as follows:

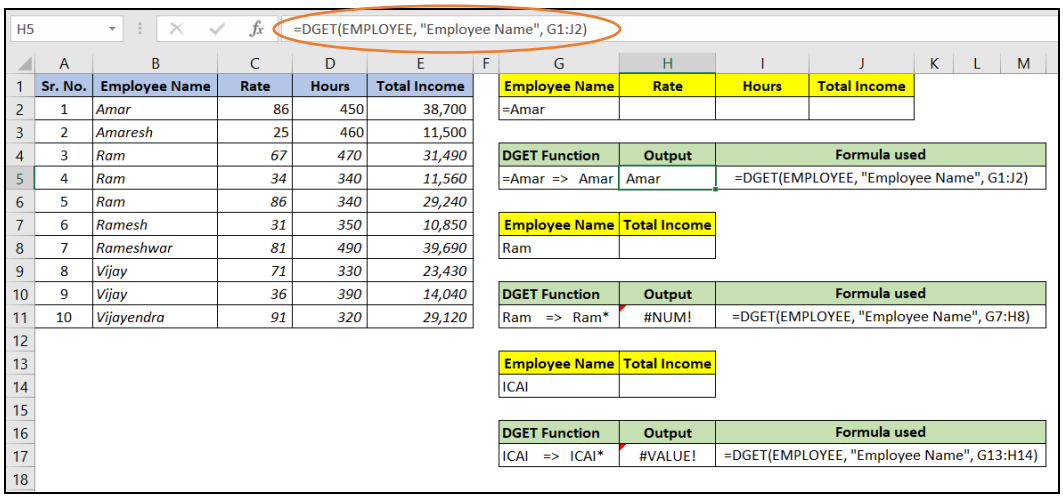


Figure 5.105 – DGET Function

In the first example, the Employee Name is "=Amar". The function will do 100% match and return the name as the field name is "Employee Name" in DGET Function.

- There is only 1 name which is 100% match.
- Hence, final output in cell H5 = Amar

In the second example, the Employee Name is "Ram". The function will search all names which begins with name "Ram". It will also include "Ramesh" and "Rameshwar"

- There are 5 names which begins with Ram.
- As per DGET Function, if more than one record matches the criterion, DGET returns the #NUM! error.
- Hence, final output in cell H11 = #NUM!

In the third example, the Employee Name is "ICAI". The function will search all names which begins with name "ICAI".

- There is no name which begins with ICAI.

— As per DGET Function, if no records match the criterion, DGET returns the #VALUE! error.

— Hence, final output in cell H17 = #VALUE!

In order to avoid errors, we can use IFERROR and IFNA Functions.

5.13 ARRAY FOR FORMULAS AND FUNCTIONS

An array in Excel is a collection of two or more items. The items can be text or numbers and they can reside in a single row or column, or in multiple rows and columns.

For example, if we put list of names into an Excel array format, it would look like:

{ "Ram", "Rajesh", "Shyam", "Amar" }

Then, if we select cells A1 through D1, enter the above array preceded by an equal sign (=) in the Formula Bar and press CTRL + SHIFT + ENTER, we will get the following result:

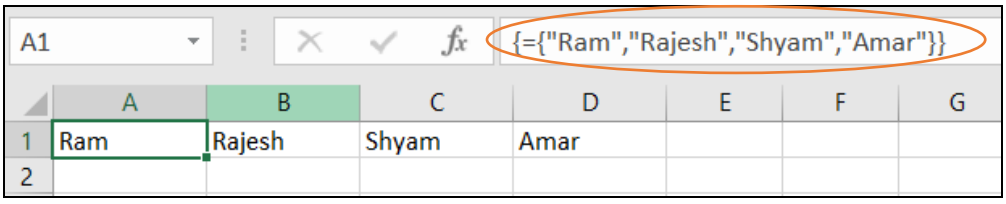


Figure 5.106 – Array example

Note: Beginning with the September 2018 update for Microsoft 365, any formula that can return multiple results will automatically spill them either down, or across into neighbouring cells. This change in behaviour is also accompanied by several new dynamic array functions.

The Types of Arrays in Excel

There are three types of arrays in Excel:

- A **reference array** contains more than one cell. Examples include a range of cells, a worksheet reference, and a defined name.
- An **array created by a formula element**, also called a resultant array, is an array of items created by the array operation.
- An **array constant** is an array of values hard coded into a formula.

5.14 CONCEPT OF ARRAY, DOUBLE UNARY

- **Array Formula:**

  An array formula is a formula that contains an operation (math, comparative, join, or function argument) on an array of items rather than on single items, and, the operation delivers a resultant array of items rather than a single item. This operation is called an array operation and is distinguish from an aggregate operation, which delivers a single item. The resultant array of items (also called an array created by a

formula element) can be used as a formula element in a larger formula, or it can be the final answer that the array formula delivers to a range of cells. The final answer from an array formula can either be a single item or an array of items.

Dynamic array formulas, whether they are using existing functions or the dynamic array functions, only need to be input into a single cell, then confirmed by pressing Enter. Earlier, legacy array formulas require first selecting the entire output range, then confirming the formula with **C**trl+**S**hift+**E**nter. They are commonly referred to as CSE formulas.

Array formulas are available in all versions of Excel 2013, Excel 2010, Excel 2007 and lower.

Advantages of Array Formulas:

i) Array formula can sometimes achieve what no other excel method can achieve without using VBA code.

ii) Array formula can save spreadsheet space by reducing a multiple column solution to a single cell formula.

iii) It is difficult to delete array formulas entered in multiple cells.

iv) Given a desired formula result and the situational demand, an array formula is the best option.

Disadvantages of Array Formulas:

i) Array formulas may slow down formula calculation time for a spreadsheet. This is a concern for spreadsheet that contain many formulas with large numbers of cell references and calculations.

ii) Array formulas sometime require Crtl+Shift+Enter. For beginners and advanced users alike, remembering to use Crtl+Shift+Enter can be difficult.

iii) Other users of spreadsheet may have a hard time using array formulas.

iv) Not as much information is available about array formulas as about other topics in Excel, such as PivotTables, Charts, Filter, and so on.

v) Creating array formulas can sometime be very complicated.

Simple example of Excel array formula

In the below example, we have a list of Employees, their rate and hours. We want to calculate Grand Total Income.

E15 {=SUM(C2:C11*D2:D11)}

	A	B	C	D	E	F
1	**Sr. No.**	**Employee Name**	**Rate**	**Hours**	**Total Income**	**Formula Used**
2	1	Amar	86	450	38,700	=C2*D2
3	2	Amaresh	25	460	11,500	=C3*D3
4	3	Ram	67	470	31,490	=C4*D4
5	4	Shyam	34	340	11,560	=C5*D5
6	5	Rajesh	86	340	29,240	=C6*D6
7	6	Ganesh	31	350	10,850	=C7*D7
8	7	Abdul	81	490	39,690	=C8*D8
9	8	Ramesh	71	330	23,430	=C9*D9
10	9	Vijay	36	390	14,040	=C10*D10
11	10	Vijayendra	91	320	29,120	=C11*D11
12						
13	***Option 1 - Normal Excel Function***		**Grand Total Income**		**2,39,620**	=SUM(E2:E11)
14						
15	***Option 2 - Array Formula***		**Grand Total Income**		**2,39,620**	{=SUM(C2:C11*D2:D11)}
16						

Figure 5.107 – Array Formula in Sum

Option 1 : Normal Excel Function

- We first calculated Total Income in cell E2 as Rate * Hours i.e. C2*D2.
- Then we copied the formula in below cells till E10.
- Then we used the SUM Function in cell E13 as "=SUM(E2:E11).
- Final Output – Grand Total Income = 2,39,,620

Option 2 : Array Formula

If we use Array Formula, we can avoid few staps and directly get the Grand Total Income as follows:

- Type the SUM Function in cell E15 as "=SUM(C2:C11*D2*D11)"
- Then Press the keyboard shortcut CTRL + SHIFT + ENTER to complete the array formula.

  Once we do this, Microsoft Excel surrounds the formula with {curly braces}, which is a visual indication of an array formula.

  What the formula does is multiply the values in each individual row of the specified array (cells C2 through D11), add the sub-totals together, and

  Output the Grand Total Income = 2,39,620

This simple example shows how powerful an Excel array formula can be. When working with hundreds and thousands of rows of data, we can save time by entering one array formula in a single cell.

We can use Array Formulas to perform complex tasks, such as:

- Quickly create sample datasets.

- Count the number of characters contained in a range of cells.
- Sum only numbers that meet certain conditions, such as the lowest values in a range, or numbers that fall between an upper and lower boundary.
- Sum every Nth value in a range of values.

How to enter array formula in Excel (Ctrl + Shift + Enter)

The combination of the 3 keys CTRL + SHIFT + ENTER is a magic touch that turns a regular formula into an array formula. When entering an array formula in Excel, there are 4 important things to keep in mind:

i) Once we have finished typing the formula and simultaneously pressed the keys CTRL SHIFT ENTER, Excel automatically encloses the formula between {curly braces}. When we select such a cell(s), we can see the braces in the formula bar, which gives a clue that an array formula is in there.

ii) Manually typing the braces around a formula won't work. We must press the Ctrl+Shift+Enter shortcut to complete an array formula.

iii) Every time we edit an array formula, the braces disappear and we must press Ctrl+Shift+Enter again to save the changes.

iv) If we forget to press Ctrl+Shift+Enter, the formula will behave like a usual Excel formula and process only the first value(s) in the specified array(s).

Use the F9 key to evaluate portions of an Array Formula

When working with array formulas in Excel, we can observe how they calculate and store their items (internal arrays) to display the final result we see in a cell. To do this, select one or several arguments within a function's parentheses, and then press the F9 key. To exit the formula evaluation mode, press the Esc key.

In the above example, to see the sub-totals of all products, we select C2:C11*D2:D11 as shown below:

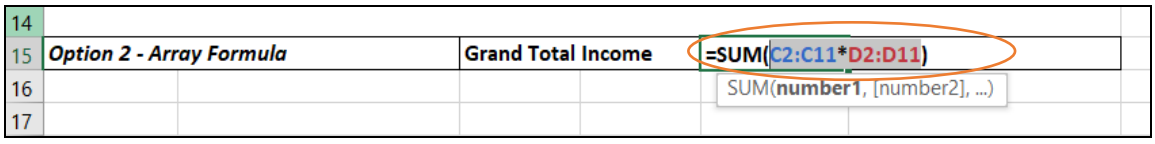
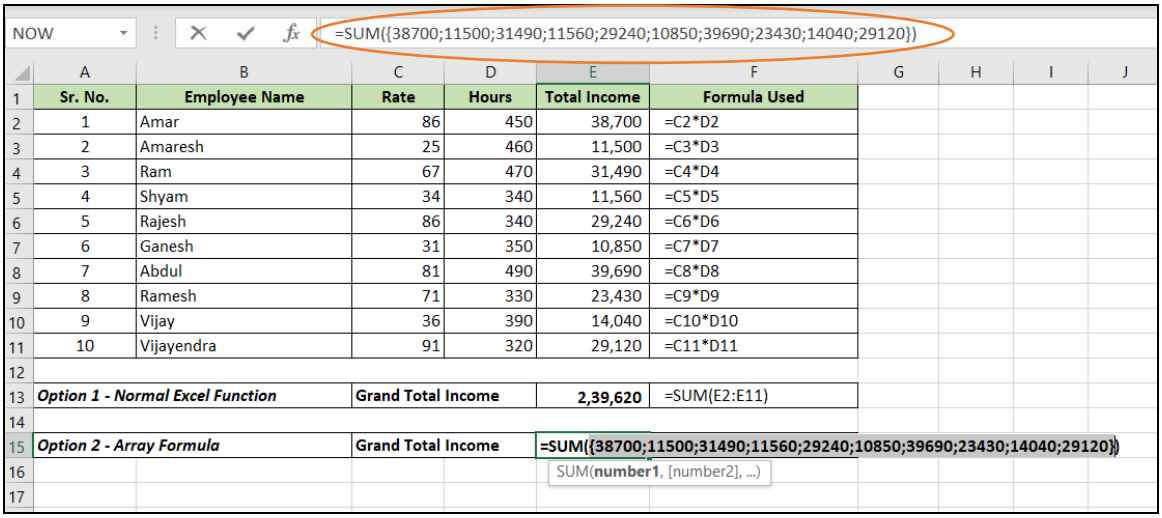


Figure 5.108 – Array Formula selection for evaluation

then press F9 and get the following result. To exit the formula evaluation mode, press the Esc key.



	A	B	C	D	E	F
1	**Sr. No.**	**Employee Name**	**Rate**	**Hours**	**Total Income**	**Formula Used**
2	1	Amar	86	450	38,700	=C2*D2
3	2	Amaresh	25	460	11,500	=C3*D3
4	3	Ram	67	470	31,490	=C4*D4
5	4	Shyam	34	340	11,560	=C5*D5
6	5	Rajesh	86	340	29,240	=C6*D6
7	6	Ganesh	31	350	10,850	=C7*D7
8	7	Abdul	81	490	39,690	=C8*D8
9	8	Ramesh	71	330	23,430	=C9*D9
10	9	Vijay	36	390	14,040	=C10*D10
11	10	Vijayendra	91	320	29,120	=C11*D11
12						
13	***Option 1 - Normal Excel Function***		**Grand Total Income**		**2,39,620**	=SUM(E2:E11)
14						
15	***Option 2 - Array Formula***		**Grand Total Income**		**=SUM({38700;11500;31490;11560;29240;10850;39690;23430;14040;29120})**	
16						
17						

Figure 5.109 – Array Formula evaluation

Note: Please pay attention that you must select some part of the formula prior to pressing F9, otherwise the F9 key will simply replace your Excel array formula with the calculated value(s).

Single-cell and multi-cell array formulas in Excel

Excel array formulas can return a result in a single cell or in multiple cells. An array formula entered in a range of cells is called a multi-cell formula. An array formula residing in a single cell is called a single-cell formula.

There exist a few Excel array functions that are designed to return multi-cell arrays, for example TRANSPOSE, TREND, FREQUENCY, LINEST, etc.

Other functions, such as SUM, AVERAGE, AGGREGATE, MAX, MIN, can calculate array expressions when entered into a single cell by using Ctrl + Shift + Enter.

The following examples demonstrate how to use a single-cell and multi-cell array formula in Excel.

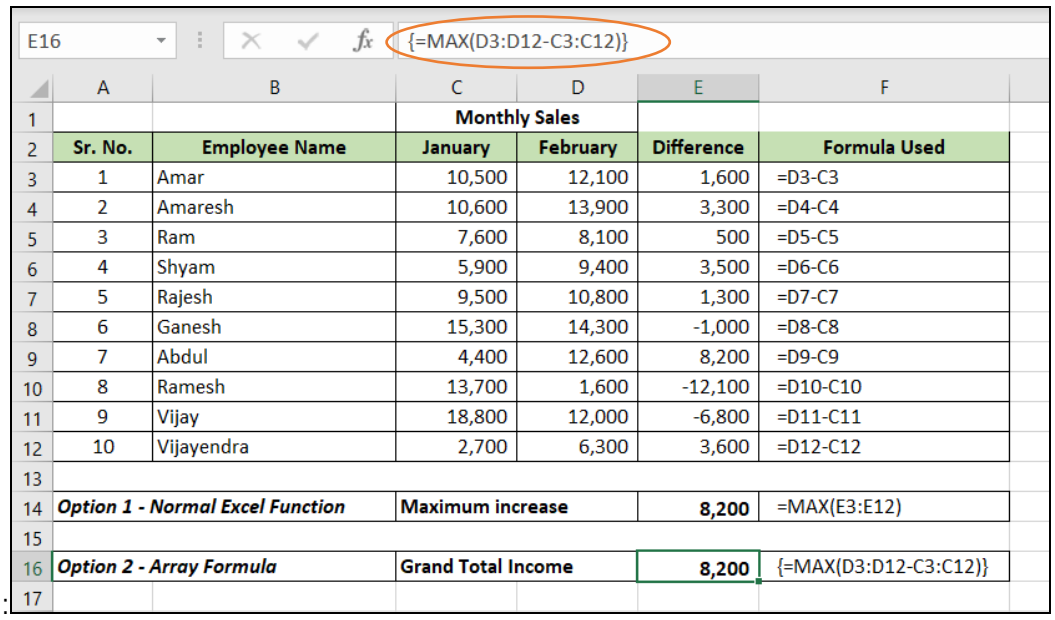
Example 1: A single-cell array formula

Suppose we have two columns listing the number of items sold in 2 different months, say columns C and D, and we want to find the maximum sales increase by an employee.

Normally, we would add an additional column, say column E, that calculates the sales change for each product using a formula like =D3-C3, and then find the maximum value in that additional column =MAX(E3:E12).

An Excel array formula does not need an additional column since it perfectly stores intermediate results in memory. So, you just enter the following formula and press Ctrl + Shift + Enter:

=MAX(D3:D12-C3:C12)



E16 {=MAX(D3:D12-C3:C12)}

	A	B	C	D	E	F
1			Monthly Sales			
2	Sr. No.	Employee Name	January	February	Difference	Formula Used
3	1	Amar	10,500	12,100	1,600	=D3-C3
4	2	Amaresh	10,600	13,900	3,300	=D4-C4
5	3	Ram	7,600	8,100	500	=D5-C5
6	4	Shyam	5,900	9,400	3,500	=D6-C6
7	5	Rajesh	9,500	10,800	1,300	=D7-C7
8	6	Ganesh	15,300	14,300	-1,000	=D8-C8
9	7	Abdul	4,400	12,600	8,200	=D9-C9
10	8	Ramesh	13,700	1,600	-12,100	=D10-C10
11	9	Vijay	18,800	12,000	-6,800	=D11-C11
12	10	Vijayendra	2,700	6,300	3,600	=D12-C12
13						
14	*Option 1 - Normal Excel Function*		Maximum increase		8,200	=MAX(E3:E12)
15						
16	*Option 2 - Array Formula*		Grand Total Income		8,200	{=MAX(D3:D12-C3:C12)}
17						

Figure 5.110 – Array Formula in Max Function

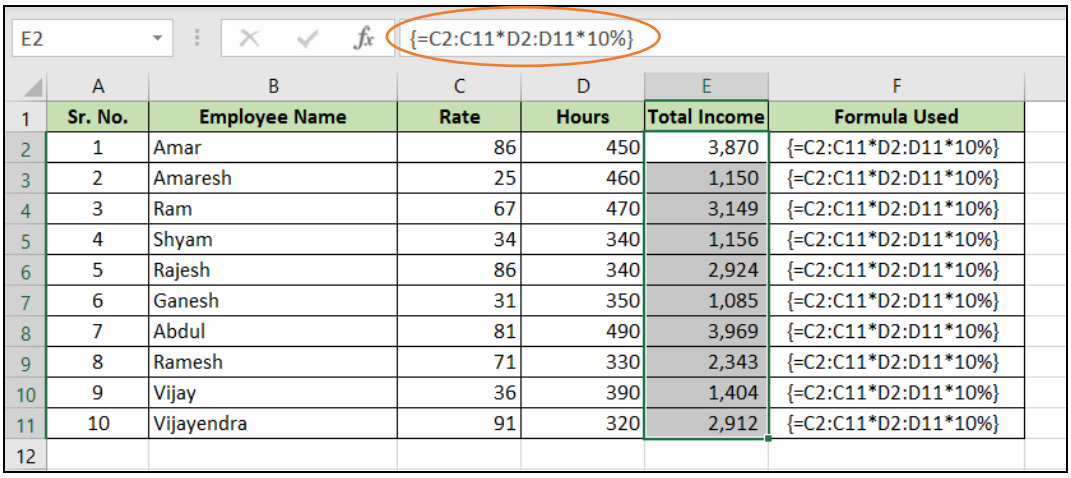
Example 2: A multi-cell array formula in Excel

In the previous SUM example, suppose we have to deduct 10% TDS from each employee's Total Income and we want to calculate the Total Tax amount for each employee with one formula.

Select the range of cell in some blank column, say E2:E11, and enter the following formula in the formula bar:

=C2:C11 * D2:D11 * 10%

Once we press Ctrl + Shift + Enter, Excel will place an instance of array formula in each cell of the selected range, and we will get the following result:



E2 {=C2:C11*D2:D11*10%}

	A	B	C	D	E	F
1	Sr. No.	Employee Name	Rate	Hours	Total Income	Formula Used
2	1	Amar	86	450	3,870	{=C2:C11*D2:D11*10%}
3	2	Amaresh	25	460	1,150	{=C2:C11*D2:D11*10%}
4	3	Ram	67	470	3,149	{=C2:C11*D2:D11*10%}
5	4	Shyam	34	340	1,156	{=C2:C11*D2:D11*10%}
6	5	Rajesh	86	340	2,924	{=C2:C11*D2:D11*10%}
7	6	Ganesh	31	350	1,085	{=C2:C11*D2:D11*10%}
8	7	Abdul	81	490	3,969	{=C2:C11*D2:D11*10%}
9	8	Ramesh	71	330	2,343	{=C2:C11*D2:D11*10%}
10	9	Vijay	36	390	1,404	{=C2:C11*D2:D11*10%}
11	10	Vijayendra	91	320	2,912	{=C2:C11*D2:D11*10%}
12						

Figure 5.111 – Multi-cell Array Formula

Example 3: Using an Excel array function to return a multi-cell array (TRANSPOSE Function)

Microsoft Excel provides a few so called "array functions" that are specially designed to work with multi-cell arrays. Excel TRANSPOSE is one of such functions and we are going to utilize it to transpose the above table, i.e. convert rows to columns.

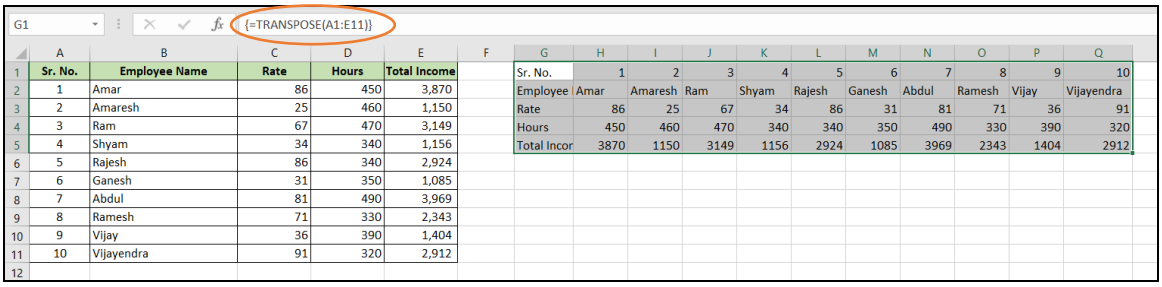
i) Select an empty range of cells where we want to output the transposed table. Since we are converting rows to columns, be sure to select the same number of rows and columns as our source table has columns and rows, respectively. In this example, we are selecting 11 columns and 5 rows.

ii) Press F2 to enter the edit mode.

iii) Enter the array function =TRANSPOSE(array) and press Ctrl + Shift + Enter.

In our example, the formula is "=TRANSPOSE($A$1:$E$11)".

The result is going to look similar to this:

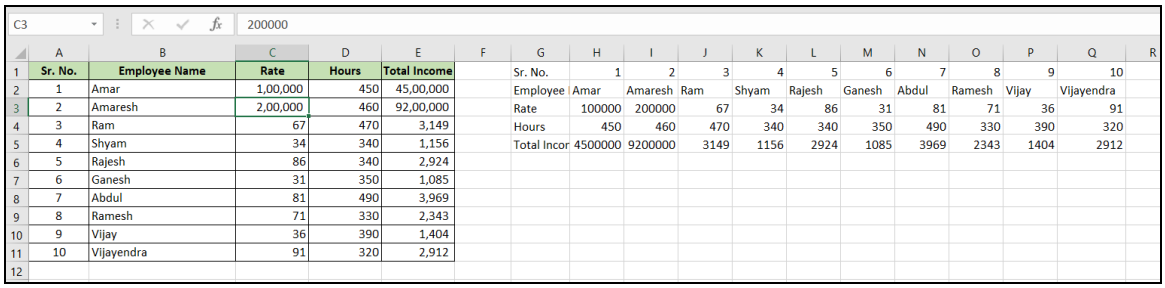


Sr. No.	Employee Name	Rate	Hours	Total Income
1	Amar	86	450	3,870
2	Amaresh	25	460	1,150
3	Ram	67	470	3,149
4	Shyam	34	340	1,156
5	Rajesh	86	340	2,924
6	Ganesh	31	350	1,085
7	Abdul	81	490	3,969
8	Ramesh	71	330	2,343
9	Vijay	36	390	1,404
10	Vijayendra	91	320	2,912

Sr. No.	1	2	3	4	5	6	7	8	9	10
Employee	Amar	Amaresh	Ram	Shyam	Rajesh	Ganesh	Abdul	Ramesh	Vijay	Vijayendra
Rate	86	25	67	34	86	31	81	71	36	91
Hours	450	460	470	340	340	350	490	330	390	320
Total Incor	3870	1150	3149	1156	2924	1085	3969	2343	1404	2912

Figure 5.112 – Array Formula in Transpose Function

Any change in Original table, will be reflected in Transposed table:

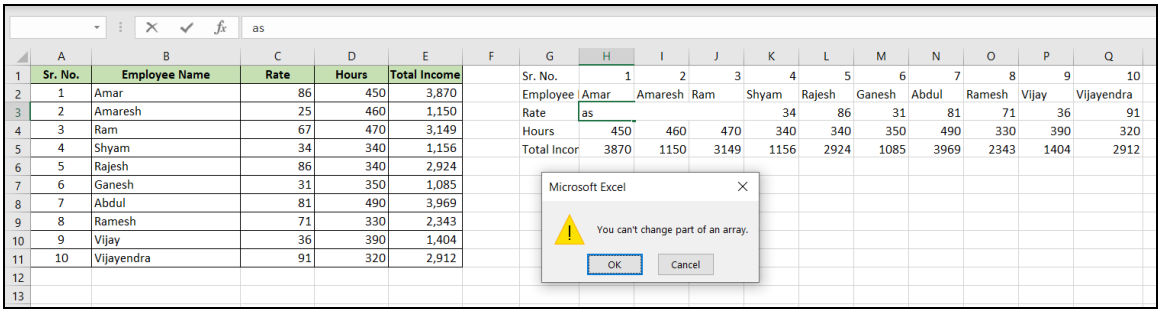


Sr. No.	Employee Name	Rate	Hours	Total Income
1	Amar	1,00,000	450	45,00,000
2	Amaresh	2,00,000	460	92,00,000
3	Ram	67	470	3,149
4	Shyam	34	340	1,156
5	Rajesh	86	340	2,924
6	Ganesh	31	350	1,085
7	Abdul	81	490	3,969
8	Ramesh	71	330	2,343
9	Vijay	36	390	1,404
10	Vijayendra	91	320	2,912

Sr. No.	1	2	3	4	5	6	7	8	9	10
Employee	Amar	Amaresh	Ram	Shyam	Rajesh	Ganesh	Abdul	Ramesh	Vijay	Vijayendra
Rate	100000	200000	67	34	86	31	81	71	36	91
Hours	450	460	470	340	340	350	490	330	390	320
Total Incor	4500000	9200000	3149	1156	2924	1085	3969	2343	1404	2912

Figure 5.113 – Array Formula in Transpose Function – data change effect

Error when we try to change any cell value in Transposed table:



Sr. No.	Employee Name	Rate	Hours	Total Income
1	Amar	86	450	3,870
2	Amaresh	25	460	1,150
3	Ram	67	470	3,149
4	Shyam	34	340	1,156
5	Rajesh	86	340	2,924
6	Ganesh	31	350	1,085
7	Abdul	81	490	3,969
8	Ramesh	71	330	2,343
9	Vijay	36	390	1,404
10	Vijayendra	91	320	2,912

Sr. No.	1	2	3	4	5	6	7	8	9	10
Employee	Amar	Amaresh	Ram	Shyam	Rajesh	Ganesh	Abdul	Ramesh	Vijay	Vijayendra
Rate	as			34	86	31	81	71	36	91
Hours	450	460	470	340	340	350	490	330	390	320
Total Incor	3870	1150	3149	1156	2924	1085	3969	2343	1404	2912

Figure 5.114 – Array Formula in Transpose Function – error

Advantages of TRANSPOSE function:

The main benefit of using the TRANSPOSE function is that the rotated table retains the connection to the source table and whenever we change the source data, the transposed table will change accordingly.

Weaknesses of TRANSPOSE function:

- The original table formatting is not saved in the converted table, as we see in the screenshot above.
- If there are any empty cells in the original table, the transposed cells will contain 0 instead. To fix this, use TRANSPOSE in combination with the IF function.
- We cannot edit any cells in the rotated table because it is very much dependent on the source data. If we try to change some cell value, we will end up with the "You cannot change part of an array" error.

Whatever good and easy-to-use the TRANSPOSE function is, it certainly lacks flexibility and therefore may not be the best way to go in many situations.

How to work with multi-cell array formulas

When working with multi-cell array formulas in Excel, we need to follow following rules to get the correct results:

i) Select the range of cells where we want to output the results before entering the formula.

ii) To delete a multi-cell array formula, either select all the cells containing it and press DELETE, or select the entire formula in the formula bar, press DELETE, and then press Ctrl + Shift + Enter.

iii) We cannot edit or move the contents of an individual cell in an Excel array formula, nor can we insert new cells into or delete existing cells from a multi-cell array formula. Whenever we try doing this, Microsoft Excel will throw the warning "You cannot change part of an array".

iv) To shrink an array formula, i.e. to apply it to fewer cells, we need to delete the existing formula first and then enter a new one.

v) To expand an array formula, i.e. apply it to more cells, select all cells containing the current formula plus empty cells where we want to have it, press F2 to switch to the edit mode, adjust the references in the formula and press Ctrl + Shift + Enter to update it.

vi) We cannot use multi-cell array formulas in Excel tables.

vii) We should enter a multi-cell array formula in a range of cells of the same size as the resulting array returned by the formula. If Excel array formula produces an array larger than the selected range, the excess values won't appear on the worksheet. If an array returned by the formula is smaller than the selected range, #N/A errors will appear in extra cells.

If our formula may return an array with a variable number of elements, enter it in a range equal to or larger than the maximum array returned by the formula and wrap the formula in the IFERROR function.

- **Array Constants**

In Microsoft Excel, an array constant is simply a set of static values. These values never change when you copy a formula to other cells or values. Above we have seen few examples of array constants *(Figure 5.106 – Array example)*.

In fact, every time we have evaluated a formula element by using the F9 key, we have seen an example of an array constant. For example in "*Figure 5.108 – Array Formula selection for evaluation"*, we highlighted the array operation C2:C11*D2:D11, pressed the F9 key, and revealed the resultant array {38700;11500;31490;11560;29240;10850;39690;23430;14040;29120} *(see Figure 5.109 above)*.

The objective of doing using F9 was to look at the resultant array to make sure the formula was calculating correctly. But immediately after that wee undid the evaluation by using Ctrl+Z, so that we did not hard code the values into the formula. If we had not used undo, we would have created an array constant that hard codes the values into the formula.

There are 3 types of array constants –

i) Column Array Constants (Vertical Array Constants),

ii) Row Array Constants (Horizontal Array Constants)

iii) Table Array Constants (Two-Way Array Constants)

Column Array Constants (Vertical Array Constants)

The below Figures 5.1115 and 5.116 illustrate that if we were to highlight a column of items and evaluate it by pressing F9, we would see the following:

— The array of items (here Employee Names) is housed in curly braces;

— the text is always shown in double quotes;

— the semicolon means to go down a row; and

— the column of items uses semicolons.

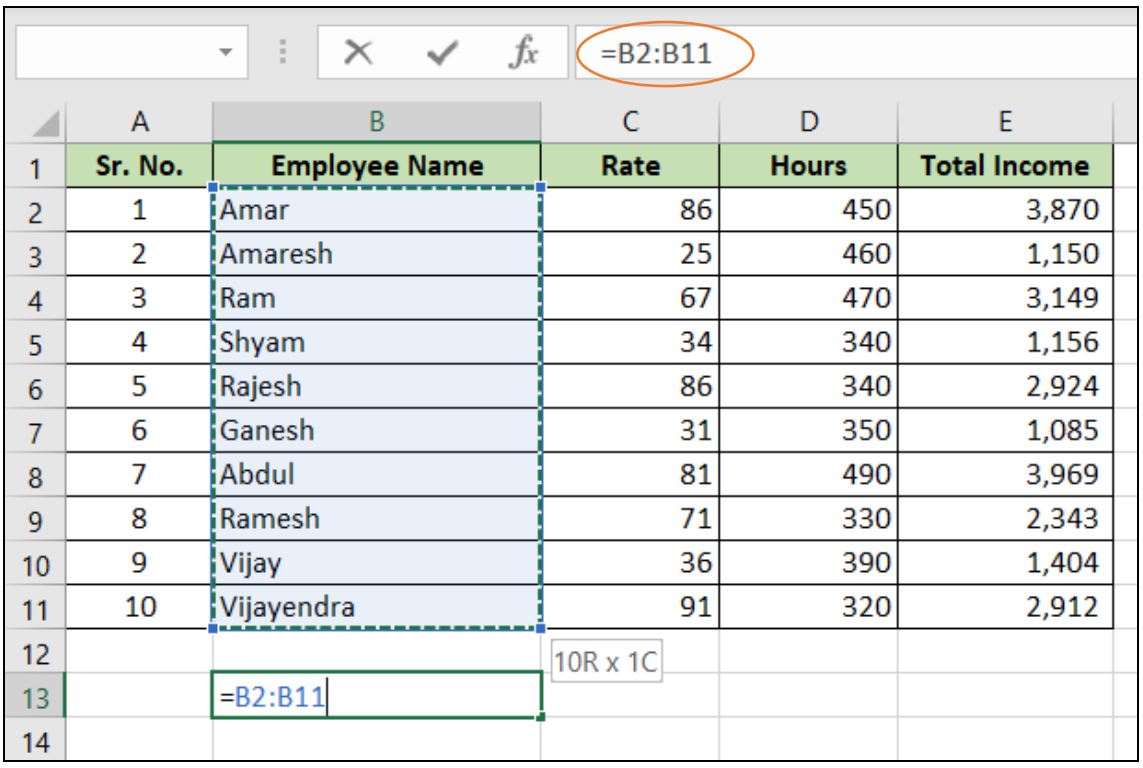


Figure 5.115 – Cell Range – A column filled with rows

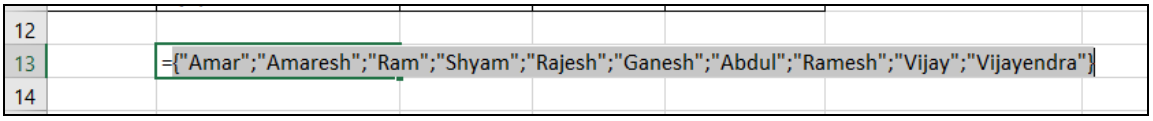


Figure 5.116 – Array Constant - A column filled with rows. Semicolon = rows

Row Array Constants (Horizontal Array Constants)

The Figures 5.117 and 5.118 illustrate that if we were to highlight a row of items and evaluate it (by pressing F9), we would see the following:

- The array of items is housed in curly braces;
- the text is shown in double quotes, and
- numbers are just numbers;
- a comma means to go over a column; and
- the row of items uses commas.

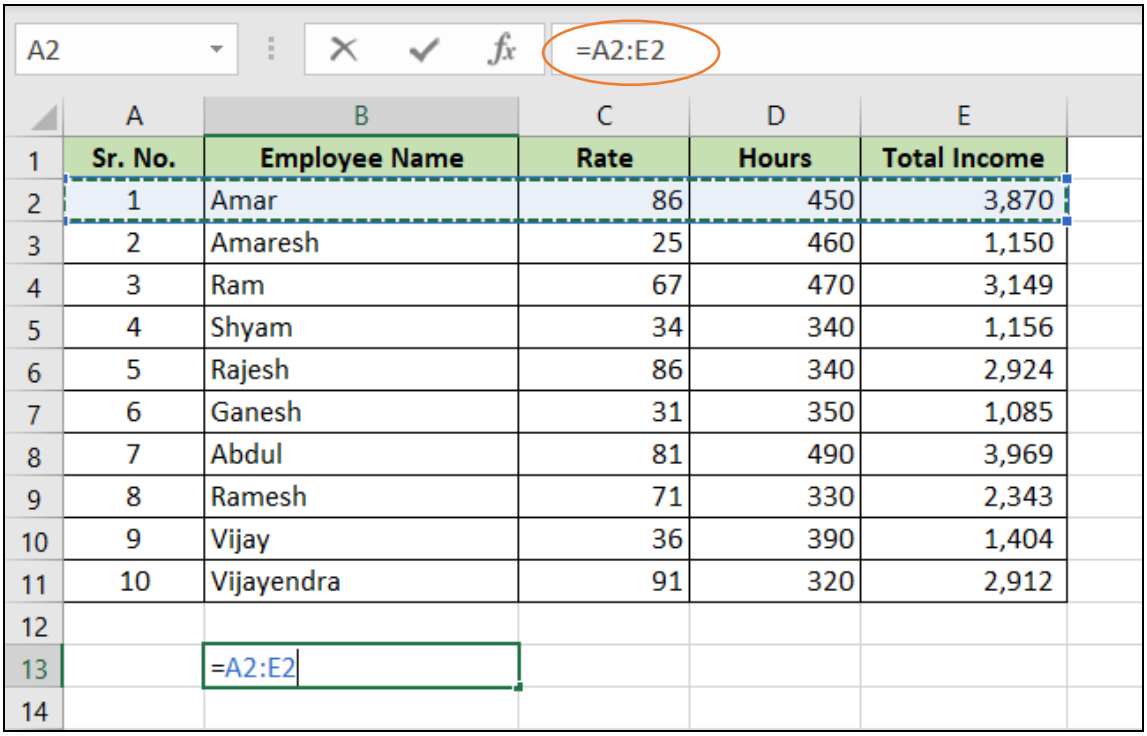


Figure 5.117 – Cell Range – A row filled with columns

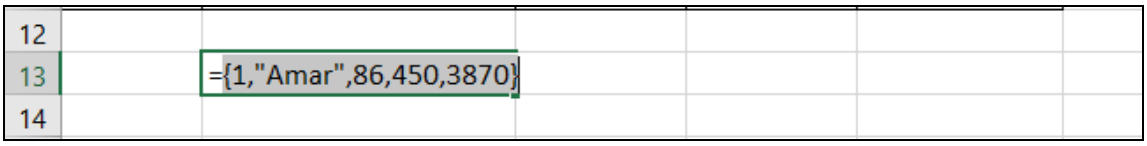
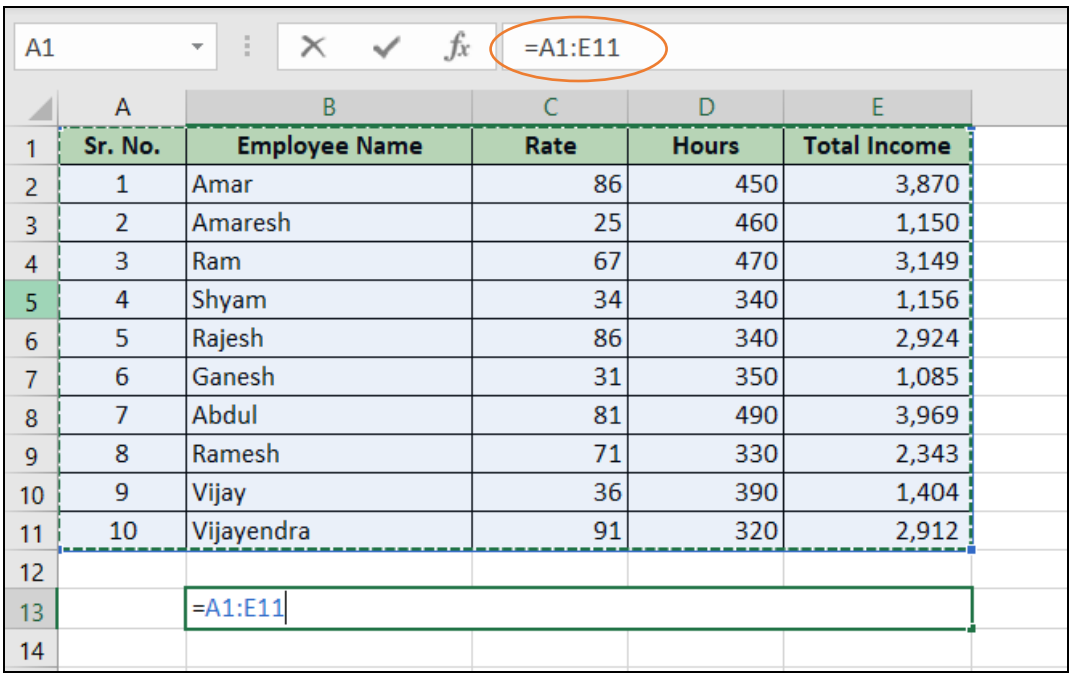


Figure 5.118 – Array Constant - A row filled with columns. Comma = rows

Table Array Constants (Two-Way Array Constants)

The Figures 5.119 and 5.120 illustrate that if we were to highlight a table (rows and columns) of items and evaluate it (by pressing F9), we would see the following:

- The array of items is housed in curly braces;
- the text is shown in double quotes, and
- numbers are just numbers; and
- the table of items uses semicolons to mean go down a row and
- commas to mean go over a column.



	A	B	C	D	E
1	Sr. No.	Employee Name	Rate	Hours	Total Income
2	1	Amar	86	450	3,870
3	2	Amaresh	25	460	1,150
4	3	Ram	67	470	3,149
5	4	Shyam	34	340	1,156
6	5	Rajesh	86	340	2,924
7	6	Ganesh	31	350	1,085
8	7	Abdul	81	490	3,969
9	8	Ramesh	71	330	2,343
10	9	Vijay	36	390	1,404
11	10	Vijayendra	91	320	2,912
12					
13		=A1:E11			
14					

Figure 5.119 – Cell Range – A table filled with rows and columns

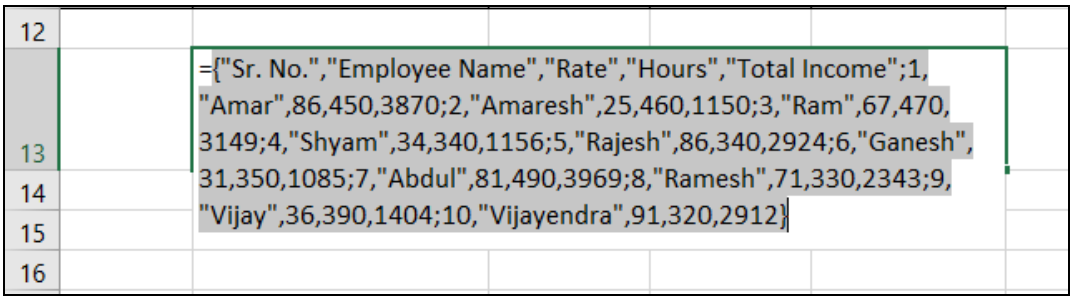


Figure 5.120 – Array Constant - A table filled with rows & columns. Comma = rows & Semicolon = column)

Array Syntax Rules

From the preceding text, we can glean the following array syntax rules:

- Curly braces house the array: one at the beginning and one at the end.
- A semicolon means to go down a row.
- A comma means to go over a column.
- A text item is contained in two double quotes.
- Numbers, logical values, and error values are not contained in two double quotes.
- The three types of array constants are column (vertical), row (horizontal), and table (two-way).

Why would we want to hard code an array constant into a formula, especially given Excel's Golden Rule:

- If formula input data can change, put it in a cell and refer to it with cell references.
- If data will not change, we can hard code it into formula.

Working with Excel array constants

Array constants are one of the cornerstones of Excel array formulas. The following information and tips might help to use them in the most efficient way.

i) **Elements of an array constant:** An Excel array constant can contain numbers, text values, Booleans (TRUE and FALSE) and error values, separated by commas or semicolons.

We can enter a numerical value as an integer, decimal, or in scientific notation. If we use text values, they should be surrounded in double quotes (") like in any Excel formula.

An array constant cannot include other arrays, cell references, ranges, dates, defined names, formulas, or Excel functions.

ii) **Naming array constants:** To make an array constant easier to use, give it a name:

— Switch to the ***Formulas tab*** > ***Defined Names*** group and click ***Define Name***. Alternatively, press "**Ctrl+F3**" and click **New**.

— Type the name in the Name

— In the Refers to box, enter the items of array constant surrounded in braces with the preceding equality sign (=). For example:

={"Sun", "Mon", "Tue", "Wed", "Thu", "Fri", "Sat"}

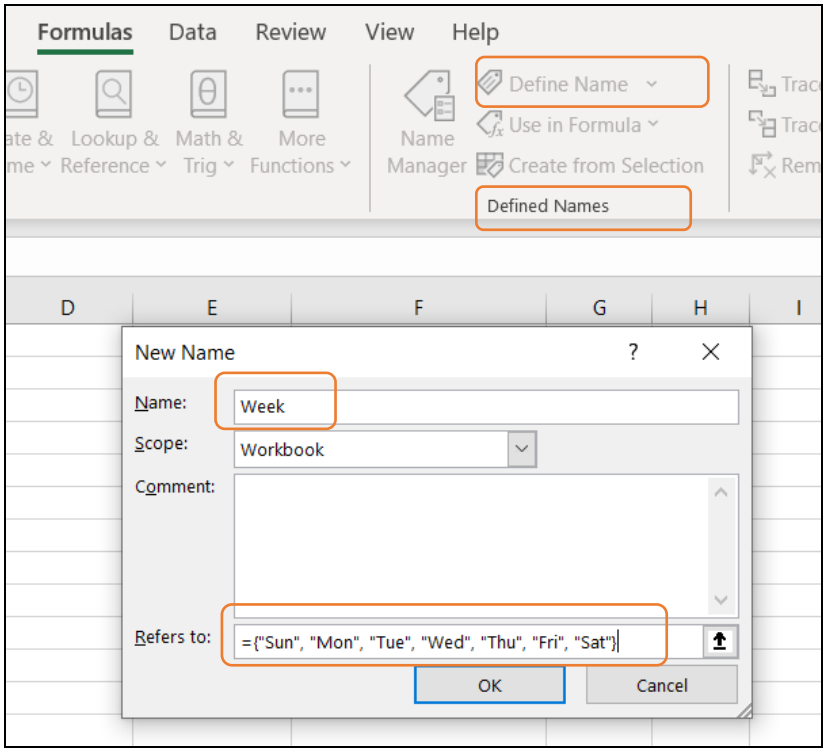


Figure 5.121 – Define Name

— Click OK to save your named array and close the window.

To enter the named array constant in a sheet, select as many cells in a row or column as there are items in your array, type the array's name in the formula bar preceded with the = sign and press Ctrl + Shift + Enter.

The result should resemble this:

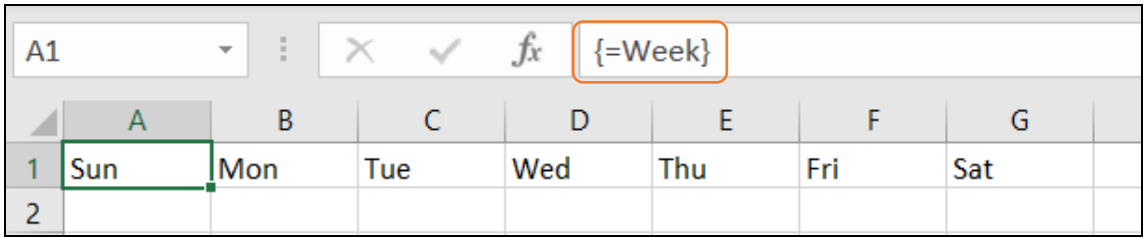


Figure 5.122 – Define Name - Output

iii) **Preventing errors:** If the array constant does not work correctly, check for the following problems:

— Delimit the elements of array constant with the proper character - comma in horizontal array constants and semicolon in vertical ones.

— Selected a range of cells that exactly matches the number of items in array constant. If we select more cells, each extra cell will have the #N/A error. If we select fewer cells, only a part of the array will be inserted.

Using array constants in Excel formulas

Now that we are familiar with the concept of array constants, let's see how to use arrays in Excel formulas to solve practical tasks.

Example 1: Sum N largest / smallest numbers in a range

We start by creating a vertical array constant containing as many numbers as we want to sum. For example, if we want to add up 3 smallest or largest numbers in a range, the array constant is {1,2,3}.

Then, we take either LARGE or SMALL function, specify entire range of cells in the first parameter and include the array constant in the second. Finally, embed it in the SUM function, like this:

Sum the largest 3 numbers: =SUM(LARGE(range, {1,2,3}))

Sum the smallest 3 numbers: =SUM(SMALL(range, {1,2,3}))

Don't forget to press Ctrl + Shift + Enter, since we are entering the Excel array formula, and you will get the following result:

C13 {=SUM(LARGE(E2:E11,{1,2,3}))}

	A	B	C	D	E
1	**Sr. No.**	**Employee Name**	**Rate**	**Hours**	**Total Income**
2	1	Amar	86	450	38,700
3	2	Amaresh	25	460	11,500
4	3	Ram	67	470	31,490
5	4	Shyam	34	340	11,560
6	5	Rajesh	86	340	29,240
7	6	Ganesh	31	350	10,850
8	7	Abdul	81	490	39,690
9	8	Ramesh	71	330	23,430
10	9	Vijay	36	390	14,040
11	10	Vijayendra	91	320	29,120
12					
13	**Sum of 3 Large value**				**1,09,880**
14		*Formula used*	{=SUM(LARGE(E2:E11,{1,2,3}))}		
15					
16	**Sum of 3 Small value**				**33,910**
17		*Formula used*	{=SUM(SMALL(E2:E11,{1,2,3}))}		
18					

Figure 5.123 – Array Formula – SUM & LARGE / SMALL Function

In a similar fashion, you can calculate the average of N smallest or largest values in a range:

Average of the top 3 numbers: =AVERAGE(LARGE(range, {1,2,3}))

Average of the bottom 3 numbers: =AVERAGE(SMALL(range, {1,2,3}))

Example 2: Array formula to count cells with multiple conditions

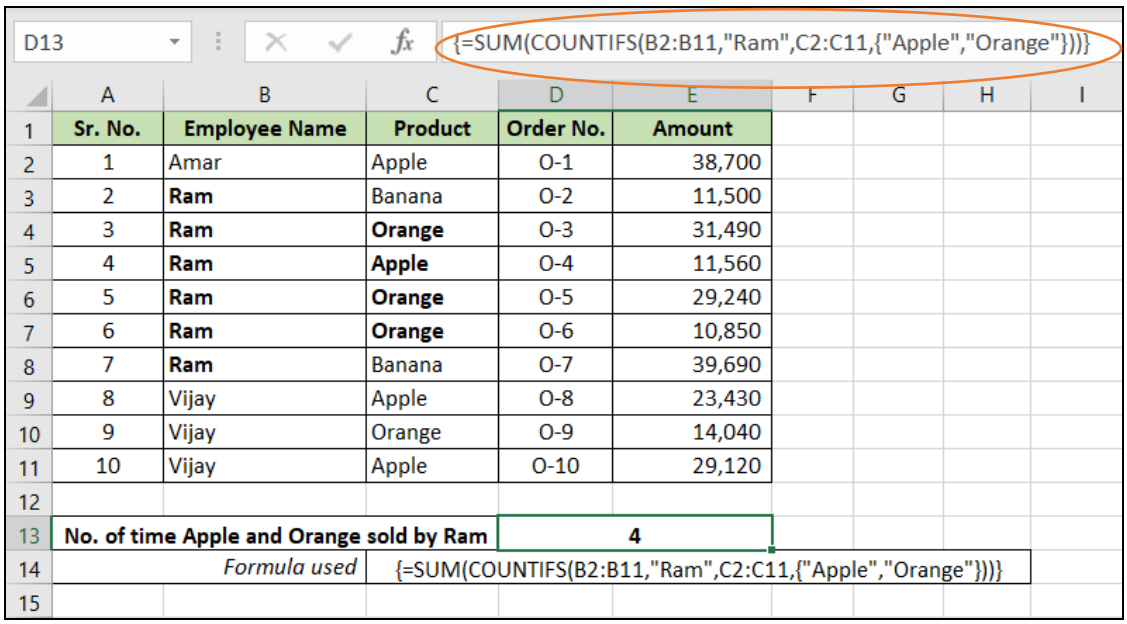
Suppose, we have a list of orders and we want to know how many times a given Employee has sold given products.

The easiest way would be using a COUNTIFS formula with multiple conditions. However, if we want to include many products, the COUNTIFS formula may grow too big in size. To make it more compact, we can use COUNTIFS together with SUM and include an array constant in one or several arguments, for example:

=SUM(COUNTIFS(range1, "criteria1", range2, {"criteria1", "criteria2"}))

The real formula may look as follows:

=SUM(COUNTIFS(B2:B9, "sally", C2:C9, {"apples", "lemons"}))



D13 {=SUM(COUNTIFS(B2:B11,"Ram",C2:C11,{"Apple","Orange"}))}

	A	B	C	D	E
1	**Sr. No.**	**Employee Name**	**Product**	**Order No.**	**Amount**
2	1	Amar	Apple	O-1	38,700
3	2	**Ram**	Banana	O-2	11,500
4	3	**Ram**	**Orange**	O-3	31,490
5	4	**Ram**	**Apple**	O-4	11,560
6	5	**Ram**	**Orange**	O-5	29,240
7	6	**Ram**	**Orange**	O-6	10,850
8	7	**Ram**	Banana	O-7	39,690
9	8	Vijay	Apple	O-8	23,430
10	9	Vijay	Orange	O-9	14,040
11	10	Vijay	Apple	O-10	29,120
12					
13	**No. of time Apple and Orange sold by Ram**			**4**	
14	*Formula used*		{=SUM(COUNTIFS(B2:B11,"Ram",C2:C11,{"Apple","Orange"}))}		
15					

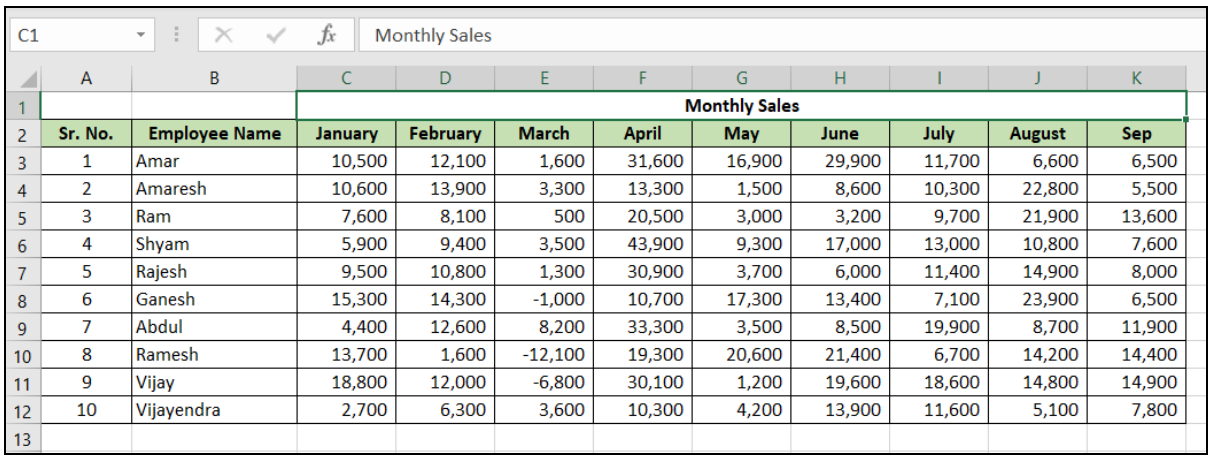
Figure 5.124 – Array Formula – SUM & COUNTIFS Function

Our sample array consists of only two elements since the goal is to demonstrate the approach. In real array formulas, we may include as many elements as the business logic requires, provided that the total length of the formula does not exceed 8,192 characters in Excel 2013, 2010 and 2007 (1,024 characters in Excel 2003 and lower) and the computer is powerful enough to process large arrays.

Excel VLOOKUP and SUM - find the sum of matching values

If we work with numerical data in Excel, quite often we have not just to extract associated values from another table but also sum numbers in several columns or rows. To do this, we can use a combination of the SUM and VLOOKUP functions as demonstrated below.

For example, we have an Employee list with sales figures for several months, a column per each month.



C1 Monthly Sales

	A	B	C	D	E	F	G	H	I	J	K
1			**Monthly Sales**								
2	**Sr. No.**	**Employee Name**	**January**	**February**	**March**	**April**	**May**	**June**	**July**	**August**	**Sep**
3	1	Amar	10,500	12,100	1,600	31,600	16,900	29,900	11,700	6,600	6,500
4	2	Amaresh	10,600	13,900	3,300	13,300	1,500	8,600	10,300	22,800	5,500
5	3	Ram	7,600	8,100	500	20,500	3,000	3,200	9,700	21,900	13,600
6	4	Shyam	5,900	9,400	3,500	43,900	9,300	17,000	13,000	10,800	7,600
7	5	Rajesh	9,500	10,800	1,300	30,900	3,700	6,000	11,400	14,900	8,000
8	6	Ganesh	15,300	14,300	-1,000	10,700	17,300	13,400	7,100	23,900	6,500
9	7	Abdul	4,400	12,600	8,200	33,300	3,500	8,500	19,900	8,700	11,900
10	8	Ramesh	13,700	1,600	-12,100	19,300	20,600	21,400	6,700	14,200	14,400
11	9	Vijay	18,800	12,000	-6,800	30,100	1,200	19,600	18,600	14,800	14,900
12	10	Vijayendra	2,700	6,300	3,600	10,300	4,200	13,900	11,600	5,100	7,800
13											

Figure 5.125 – Array Formula – Sales Table

Now, we want to make a summary table with the total sales for each Employee.

The solution is to use an array in the 3rd parameter (col_index_num) of the Excel VLOOKUP function. Here is a sample VLOOKUP formula:

=SUM(VLOOKUP(lookup value, lookup range, {2,3,4}, FALSE))

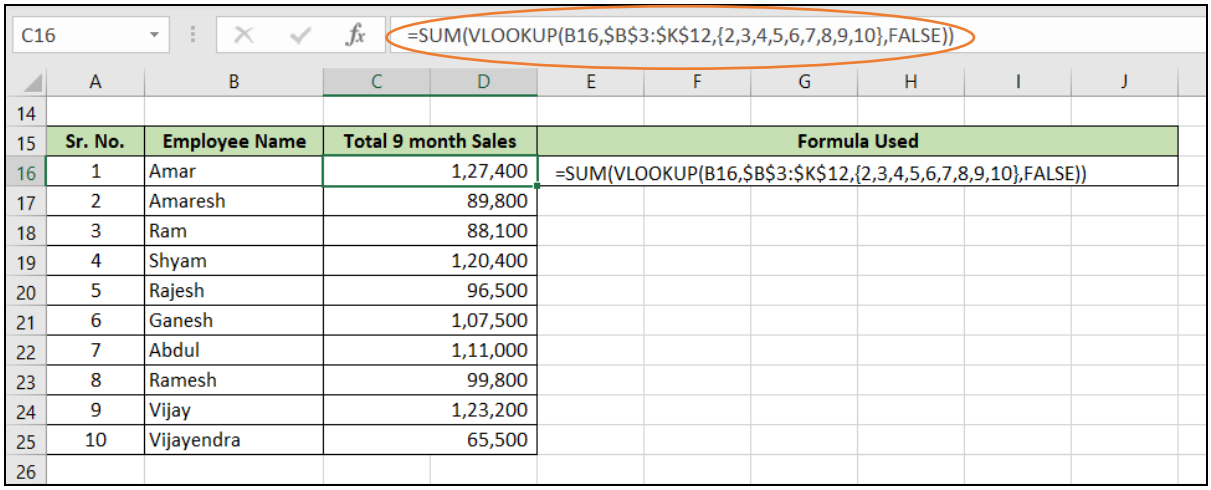
As you see, we use an array {2,3,4} in the third argument to perform several lookups within the same VLOOKUP formula in order to get the sum of values in columns 2,3 and 4.

And now, let's adjust this combination of VLOOKUP and SUM functions for our data to find the total of sales in columns C - K in the above table:

=SUM(VLOOKUP(B16, $B$3:$K$12, {2,3,4,5,6,7,8,9,10}, FALSE))

Since we are building an array formula, be sure to hit Ctrl + Shift + Enter instead of a simple Enter keystroke when finished typing. When we do this, Microsoft Excel encloses the formula in curly braces like this:

{=SUM(VLOOKUP(B16, $B$3:$K$12, {2,3,4,5,6,7,8,9,10}, FALSE))}



Sr. No.	Employee Name	Total 9 month Sales	Formula Used
1	Amar	1,27,400	=SUM(VLOOKUP(B16,$B$3:$K$12,{2,3,4,5,6,7,8,9,10},FALSE))
2	Amaresh	89,800	
3	Ram	88,100	
4	Shyam	1,20,400	
5	Rajesh	96,500	
6	Ganesh	1,07,500	
7	Abdul	1,11,000	
8	Ramesh	99,800	
9	Vijay	1,23,200	
10	Vijayendra	65,500	

Figure 5.126 – Array Formula – SUM & VLOOKUP Function

If we press the Enter key as usual, only the first value in the array will get processed, which will produce incorrect results.

As we see, using the VLOOKUP and SUM functions in Excel is easy. However, this is not the ideal solution, especially if we are working with big tables. The point is that using array formulas may adversely affect the workbook's performance since each value in the array makes a separate call of the VLOOKUP function. So, the more values we have in the array and the more array formulas we have in the workbook, the slower Excel works.

AND and OR operators in Excel array formulas

An array operator tells the formula how we want to process the arrays - using AND or OR logic.

— AND operator is the asterisk (*) which is the multiplication symbol. It instructs Excel to return TRUE if ALL of the conditions evaluate to TRUE.

— OR operator is the plus sign (+). It returns TRUE if ANY of the conditions in a given expression evaluates to TRUE.

Array formula with the AND operator

In this example, we find the sum of sales where the employee is Ram AND the product is Orange:

=SUM((B2:B11="RAM") * (C2:C11="ORANGE") * (E2:E11))

Or

=SUM(IF(((B2:B11="RAM") * (C2:C11= "orange")), (E2:E11)))

D13 =SUM((B2:B11="RAM") * (C2:C11 = "ORANGE") * (E2:E11))

	A	B	C	D	E	F	G	H
1	**Sr. No.**	**Employee Name**	**Product**	**Order No.**	**Amount**			
2	1	Amar	Apple	O-1	38,700			
3	2	**Ram**	Banana	O-2	11,500			
4	3	**Ram**	**Orange**	O-3	31,490			
5	4	**Ram**	Apple	O-4	11,560			
6	5	**Ram**	**Orange**	O-5	29,240			
7	6	**Ram**	**Orange**	O-6	10,850			
8	7	**Ram**	Banana	O-7	39,690			
9	8	Vijay	Apple	O-8	23,430			
10	9	Vijay	Orange	O-9	14,040			
11	10	Vijay	Apple	O-10	29,120			
12								
13	**Total Sales Amount of Ram for Orange**				**71,580**	Option 1		
14	*Formula used*	=SUM((B2:B11="RAM") * (C2:C11 = "ORANGE") * (E2:E11))						
15								
16	**Total Sales Amount of Ram for Orange**				**71,580**	Option 2		
17	*Formula used*	=SUM(IF(((B2:B11 = "RAM") * (C2:C11 = "orange")), E2:E11))						
18								

Figure 5.127 – Array Formula – SUM & AND / OR Function

Technically, this formula multiplies the elements of the three arrays in the same positions. The first two arrays are represented by TRUE and FALSE values which are the results of comparing B2:B11 to “RAM" and C2:C11 to "ORANGE". The third array contains the sales amount from the range E2:E11.

Like any math operation, multiplication converts TRUE and FALSE to 1 and 0, respectively. And because multiplying by 0 always gives zero, the resulting array has 0 when either or both conditions are not met. If both conditions are met, the corresponding element from the third array gets into the final array (e.g. 1*1*E4 = 31490). So, the result of multiplication is this array: {0;0;31490;0;29240;10850;0;0;0;0}. Finally, the SUM function adds up the array's elements and return a result of 71580.

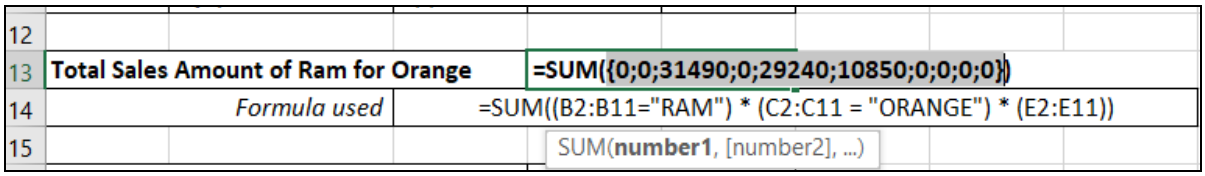
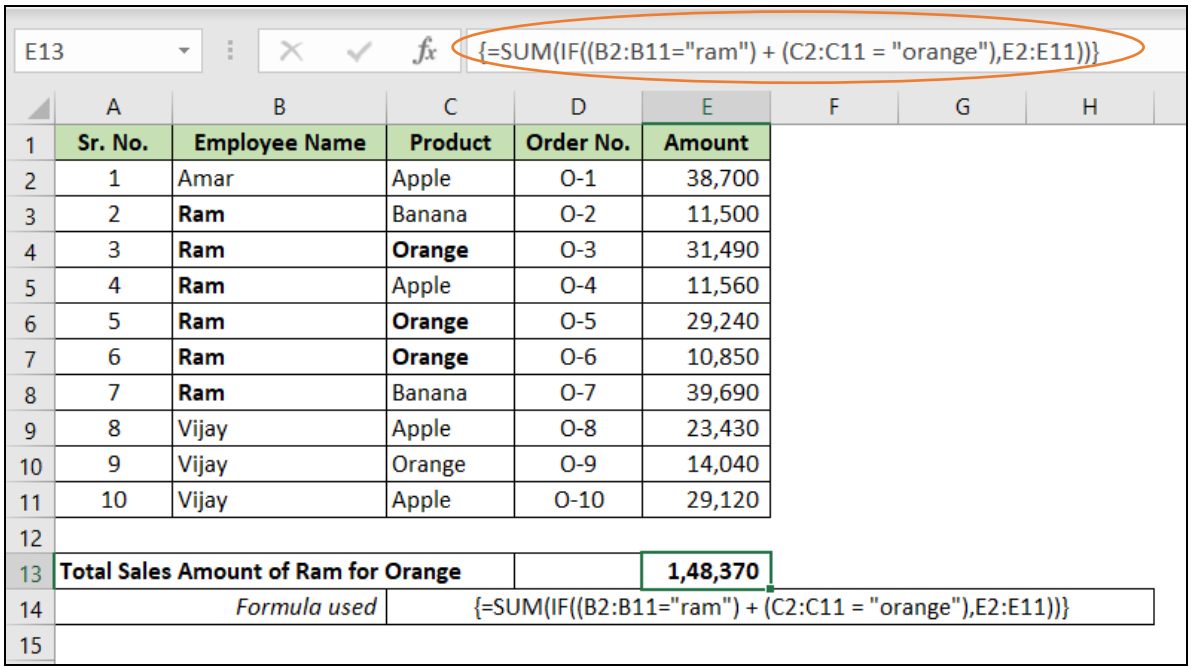


Figure 5.128 – Array Formula – SUM & AND / OR Function – use of F9

Excel array formula with the OR operator

The following array formula with the OR operator (+) adds up all sales where the employee is RAM OR product is ORANGE:

=SUM(IF((B2:B11="ram") + (C2:C11 = "orange"),E2:E11))



	A	B	C	D	E	F	G	H
1	**Sr. No.**	**Employee Name**	**Product**	**Order No.**	**Amount**			
2	1	Amar	Apple	O-1	38,700			
3	2	**Ram**	Banana	O-2	11,500			
4	3	**Ram**	**Orange**	O-3	31,490			
5	4	**Ram**	Apple	O-4	11,560			
6	5	**Ram**	**Orange**	O-5	29,240			
7	6	**Ram**	**Orange**	O-6	10,850			
8	7	**Ram**	Banana	O-7	39,690			
9	8	Vijay	Apple	O-8	23,430			
10	9	Vijay	Orange	O-9	14,040			
11	10	Vijay	Apple	O-10	29,120			
12								
13	**Total Sales Amount of Ram for Orange**				**1,48,370**			
14	*Formula used*	{=SUM(IF((B2:B11="ram") + (C2:C11 = "orange"),E2:E11))}						
15								

Figure 5.129 – Array Formula – SUM & IF Function

In this formula, we add up the elements of the first two arrays (which are the conditions we want to test), and get TRUE (>0) if at least one condition evaluates to TRUE; FALSE (0) when all the conditions evaluates to FALSE. Then, IF checks if the result of addition is greater than 0, and if it is, SUM adds up a corresponding element of the third array (E2:E11).

In Excel versions, there is no need to use an array formula for this kind of tasks - a simple SUMIFS formula handles them perfectly. Nevertheless, the AND and OR operators in array formulas may prove helpful in more complex scenarios.

Double unary operator in Excel array formulas

If we have ever worked with array formulas in Excel, chances are we came across a few ones containing a double dash (--) and we may have wondered what it was used for.

A double dash, which is technically called the double unary operator, is used to convert non-numeric Boolean values (TRUE / FALSE) returned by some expressions into 1 and 0 that an Excel array function can understand.

The following example will hopefully make things easier to understand. Suppose we have a list of dates in column A and we want to know how many dates occur in January, regardless of the year.

The following formula will work a treat:

=SUM(--(MONTH(A2:A11)=1))

Since this is an Excel array formula, remember to press Ctrl + Shift + Enter to complete it.

If we are interested in some other month, replace 1 with a corresponding number. For example, 2 stands for February, 3 means March, and so on. To make the formula more flexible, you can specify the month number in some cell, like demonstrated in the screenshot below:

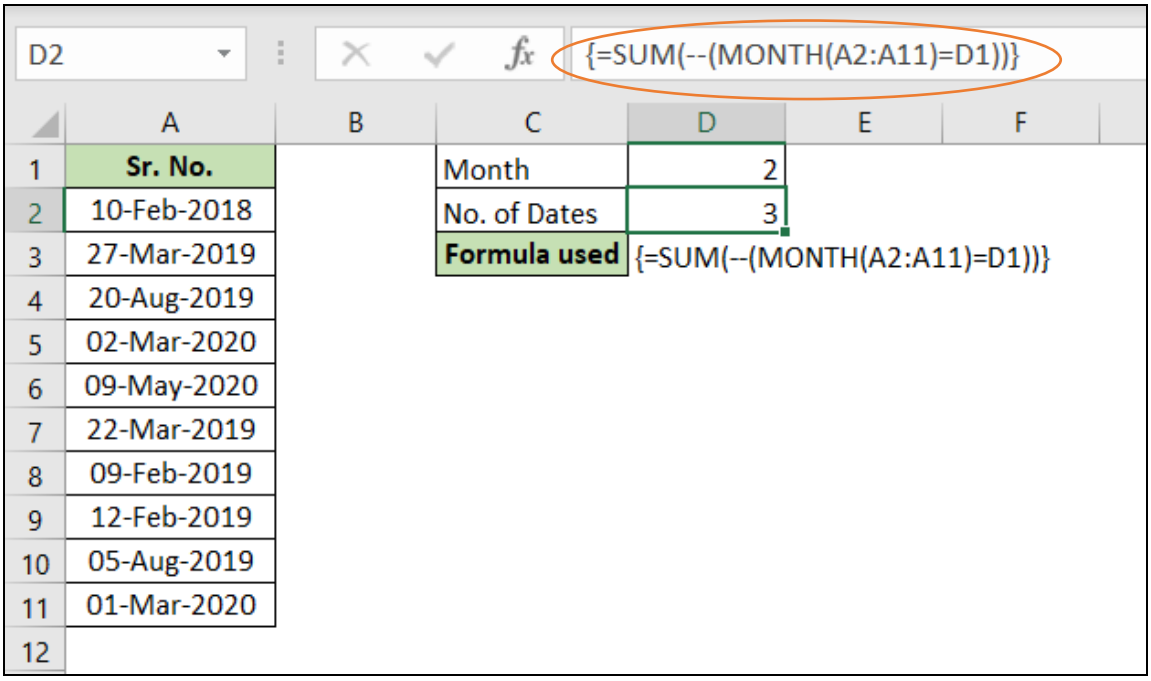


Figure 5.130 – Double Unary Operator - SUM & MONTH Function

And now, let's analyse how this Excel array formula works. The MONTH function returns the month of each date in cells A2 through A11 represented by a serial number, which producing the array {2;3;8;3;5;3;2;2;8;3}.

After that, each element of the array is compared to the value in cell D1, which is number 2 in this example. The result of this comparison is an array of Boolean values TRUE and FALSE. As you remember, we can select a certain portion of an array formula and press F9 to see what that part equates to:

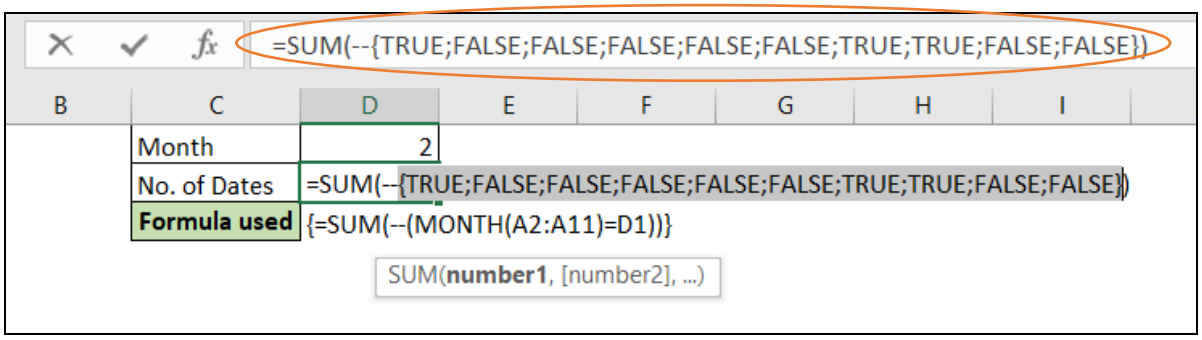


Figure 5.131 – Double Unary Operator - SUM & MONTH Function – use of F9

Finally, we have to convert these Boolean values to 1's and 0's that the SUM function can understand. And this is what the double unary operator is needed for. The first unary coerces TRUE/FALSE to -1/0, respectively. The second unary negates the values, i.e. reverses the sign, turning them into +1 and 0, which most of Excel functions can understand and work with. If we remove the double unary from the above formula, it won't work.

5.15 ADVANCE ARRAY FORMULAS

- **OFFSET:** The OFFSET function returns a reference to a range that is a given number of rows and columns from a given reference. It gets the address of the cell that is offset from another cell by a certain number of rows and/or columns. For example, cell E4 is offset from cell B4 by three columns because it is three columns to the right. OFFSET takes up to five arguments. The general syntax is as follows:

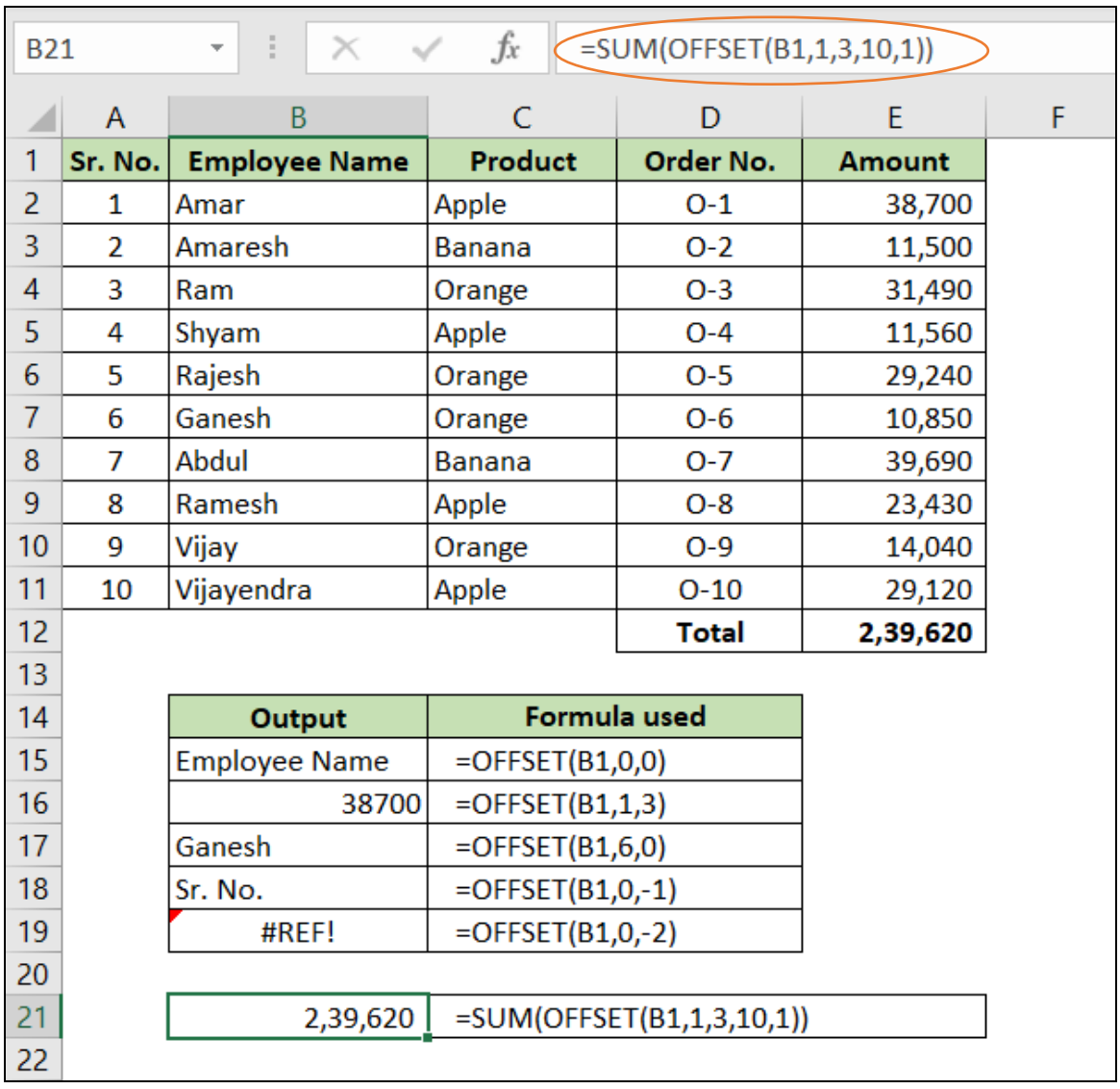
  =OFFSET(reference, rows, cols, [height], [width])

  The first three are required:

  - **A cell address or a range address:** Named ranges are not allowed.
  - **The number of rows to offset:** This can be a positive or negative number. Use 0 for no row offset.
  - **The number of columns to offset:** This can be a positive or negative number. Use 0 for no column offset.
  - **The number of rows in the returned range:** The default is the number of rows in the reference range (the first argument).
  - **The number of columns to return:** The default is the number of columns in the reference range.

  If we omit the last two arguments, OFFSET returns a reference to a single cell. If you include a value greater than 1 for either or both, the function's return references a range of the specified size with the top-left cell at the specified offset.

  Below Figure 5.132 shows some examples of using OFFSET. Columns A through E contain list of Employees and Sales Amount. The Output Table shows how OFFSET has returned different values from cells that are offset from cell B1.



B21 | =SUM(OFFSET(B1,1,3,10,1))

	A	B	C	D	E	F
1	**Sr. No.**	**Employee Name**	**Product**	**Order No.**	**Amount**	
2	1	Amar	Apple	O-1	38,700	
3	2	Amaresh	Banana	O-2	11,500	
4	3	Ram	Orange	O-3	31,490	
5	4	Shyam	Apple	O-4	11,560	
6	5	Rajesh	Orange	O-5	29,240	
7	6	Ganesh	Orange	O-6	10,850	
8	7	Abdul	Banana	O-7	39,690	
9	8	Ramesh	Apple	O-8	23,430	
10	9	Vijay	Orange	O-9	14,040	
11	10	Vijayendra	Apple	O-10	29,120	
12				**Total**	**2,39,620**	
13						
14		**Output**	**Formula used**			
15		Employee Name	=OFFSET(B1,0,0)			
16		38700	=OFFSET(B1,1,3)			
17		Ganesh	=OFFSET(B1,6,0)			
18		Sr. No.	=OFFSET(B1,0,-1)			
19		#REF!	=OFFSET(B1,0,-2)			
20						
21		2,39,620	=SUM(OFFSET(B1,1,3,10,1))			
22						

Figure 5.132 – OFFSET Function

Some highlights follow:

- Cell B15 returns the value of cell B1 because both the row and column offset is set to 0: =OFFSET(B1,0,0).
- Cell B16 returns the value we find in cell E3 i.e. 38700. This is because from cell B1, the row offset is 1 and column offset is 3.
- Cell B17 returns the value we find in cell B7 i.e. Ganesh. This is because from cell B1, the row offset is 6 and column offset is 0.

- Cell B18 returns the value we find in cell A1 i.e. Sr. No. This is because the row offset is 0 and column offset is -1. From the perspective of B1, minus 1 column: =OFFSET(B1,0,-1).
- Cell B19 displays an error because OFFSET is attempting to reference a column that is less than the first column: =OFFSET(B1,0,-2).
- Cell B21 makes use of the two optional OFFSET arguments to tell the SUM function to calculate the sum of the range E2:E11 =SUM(OFFSET(B1,1,3,10,1)).

Here's how to use the OFFSET function:

i. Click a cell where you want the result to appear.

ii. Type =OFFSET(to start the function.

iii. Enter a cell address or click a cell to get its address.

iv. Type a comma (,).

v. Enter the number of rows we want to offset where the function looks for a value. This number can be a positive number, a negative number, or 0 for no offset.

vi. Type a comma (,).

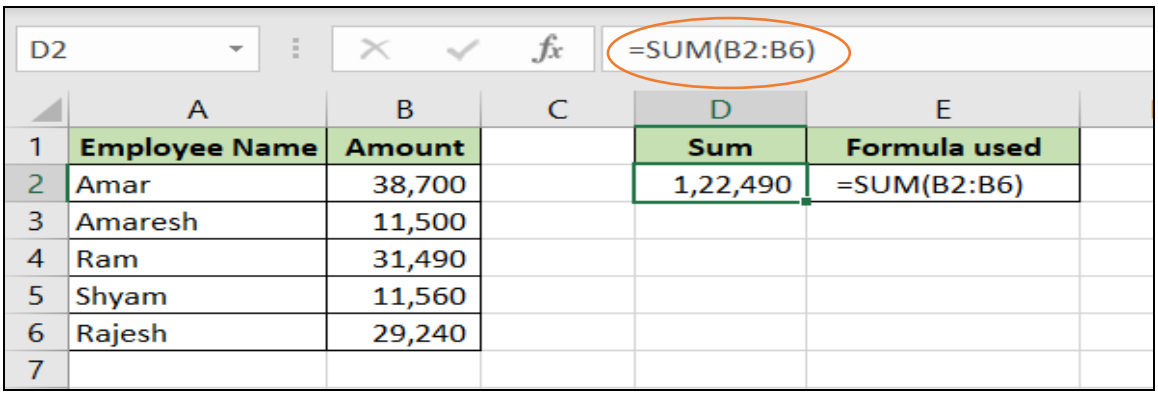
vii. Enter the number of columns we want to offset where the function looks for a value. This can be a positive number, a negative number, or 0 for no offset.

viii. Type a) and press Enter.

OFFSET is another of those functions that can be used alone but is usually used as part of a more complex formula.

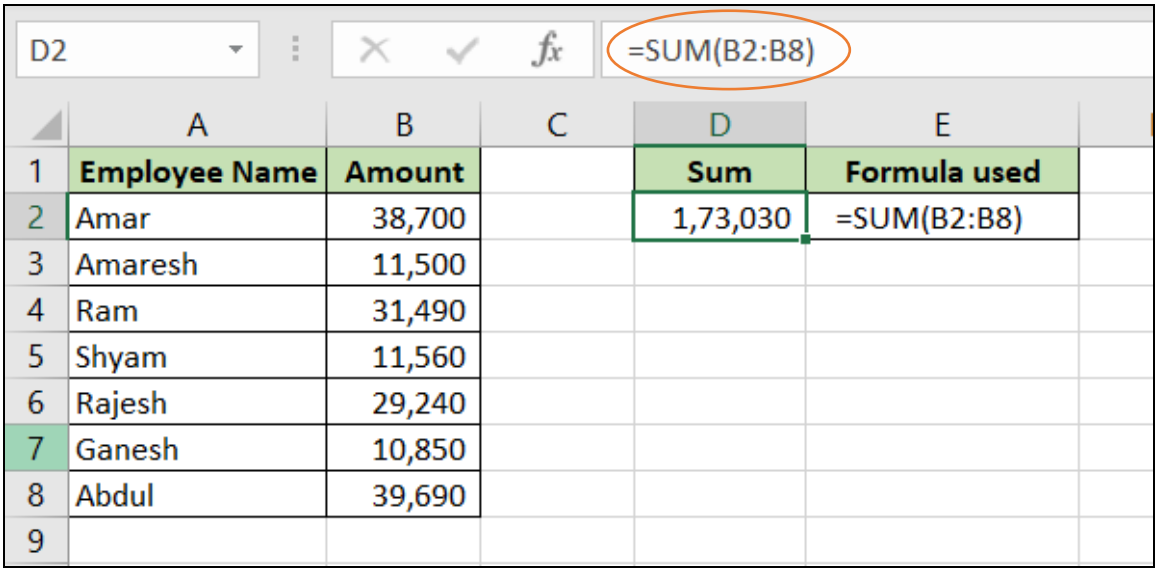
- **Use of OFFSET for Dynamic Ranges:** An array is a collection of two or more items. A range is one example of an array. A dynamic range is a range, which we can use in formulas, charts, PivotTables, and other locations that expect a range that can expand or contract when we add or delete source data or change based on some criteria as the formula that contains the range is copied down a column.

  Figures 5.133 and 5.134 illustrate an expanding source data set (just an example). If we use a dynamic range inside the SUM function, the range B2:B6 will automatically expand to B2:B8 if a 2 new number is added to the source data set.



	A	B	C	D	E
1	**Employee Name**	**Amount**		**Sum**	**Formula used**
2	Amar	38,700		1,22,490	=SUM(B2:B6)
3	Amaresh	11,500			
4	Ram	31,490			
5	Shyam	11,560			
6	Rajesh	29,240			
7					

Figure 5.133 – SUM is looking in range of B2:B6



	A	B	C	D	E
1	**Employee Name**	**Amount**		**Sum**	**Formula used**
2	Amar	38,700		1,73,030	=SUM(B2:B8)
3	Amaresh	11,500			
4	Ram	31,490			
5	Shyam	11,560			
6	Rajesh	29,240			
7	Ganesh	10,850			
8	Abdul	39,690			
9					

Figure 5.134 – Inside SUM function, a dynamic range will automatically expand from B2:B6 to B2:B8 if new data is added

Figure 5.135 illustrates a second good use for dynamic ranges in formulas. We need the formula range inside the COUNTIF function to look at the count of Amount for each Employee. In cell D5, we need the COUNTIF to look at the range B2:B5, but when we copy the formula down to D8, we need COUNTIF to be looking at the range B6:B8.

D5 =COUNTIF(B2:B5, ">"&D$2)

	A	B	C	D	E
1	**Employee Name**	**Amount**		**Average Amount**	
2	Amar	38,700		23,500	
3		14,500			
4		31,490		**Count of Amount**	**Formula used**
5		11,560		2	=COUNTIF(B2:B5, ">"&D$2)
6	Rajesh	18,000			
7		10,850		**Count of Amount**	**Formula used**
8		39,690		1	=COUNTIF(B6:B8, ">"&D$2)
9					

Figure 5.135 – As we copy the formula down the column, the range must change to e pointing only at amount for the given employee

Creating Dynamic Ranges with Formulas

As we have seen some examples of dynamic arrays for understanding, now we need to understand how do create dynamic ranges? We can use either the INDEX function or the OFFSET function to create dynamic ranges. It is advisable to use the INDEX function more often than OFFSET because OFFSET is a volatile function.

A Volatile functions recalculate every time Excel recalculates the spreadsheet, even if the precedents (formula inputs) have not changed. Recalculation can be triggered by simple actions such as entering an item into a cell or inserting a new row. Therefore, Volatile functions increase formula calculation time.

These are some of the actions that can trigger recalculation:

— Entering new data

— Deleting or inserting a row or column

— Performing certain AutoFilter actions

— Double-clicking a row or column divider

— Renaming a worksheet

— Changing the position of a worksheet in relation to other worksheets

The following are also volatile functions: CELL, INDIRECT, INFO, NOW, OFFSET, RAND, and TODAY.

If recalculation speed is not an issue (because we have small data sets, not many formulas, or don't care), the volatility of OFFSET is not a problem.

Using the Excel Table Feature to Create Dynamic Ranges

If we convert a proper data set to an Excel table, the ranges in the table are dynamic. This amazing feature was added in Excel 2003 and is great if we are going to have an expanding table of data. In Excel

2003, we can convert a proper data set using the keyboard Ctrl + L and in Excel 2007 or later you can use Ctrl + T. A number of limitations to this feature make dynamic ranges created with formulas useful.

These are some of the limitations of Excel tables:

— Dynamic ranges like the one illustrated in Figure 5.135 cannot be created using an Excel table.

— Some formulas do not work in Excel tables (for example, =ROWS(A$1:A1)).

— We can't use the Subtotal feature in an Excel table.

— We can't use custom views in an Excel table.

This means it can be useful to learn how to create dynamic ranges with formulas.

Using OFFSET to Create Dynamic Ranges: The OFFSET function can define a dynamic range. This function has five arguments:

=OFFSET(reference, rows, cols, height, width)

Using the OFFSET function is quite straightforward. We tell OFFSET the starting point of the dynamic range (reference), how many cells we want to move away from the start position (rows and cols), and the height and width (height and width). Then OFFSET defines the range.

Figure 5.136 shows an OFFSET example which assumes that only numbers will be entered. There will be no empty cells before the last number, and we will not enter numbers past cell B12.

D2 {=OFFSET($B$2,,,COUNT($B$2:$B$12))}

	A	B	C	D	E
1	**Employee Name**	**Amount**		**Dynamic Range**	**Formula used**
2	Amar	38,700		38700	{=OFFSET($B$2,,,COUNT($B$2:$B$12))}
3	Amaresh	11,500			
4	Ram	31,490			
5	Shyam	11,560			
6					
7					

Figure 5.136 – Dynamic Range using OFFSET Function

The dynamic range starts in cell $B$2 (the reference argument), it moves no rows up or down (the rows argument is empty), it moves no columns left or right (the cols argument is empty), the height is 4 (the height argument), and the width is 1 because the width argument is left empty and so it assumes the width of the reference.

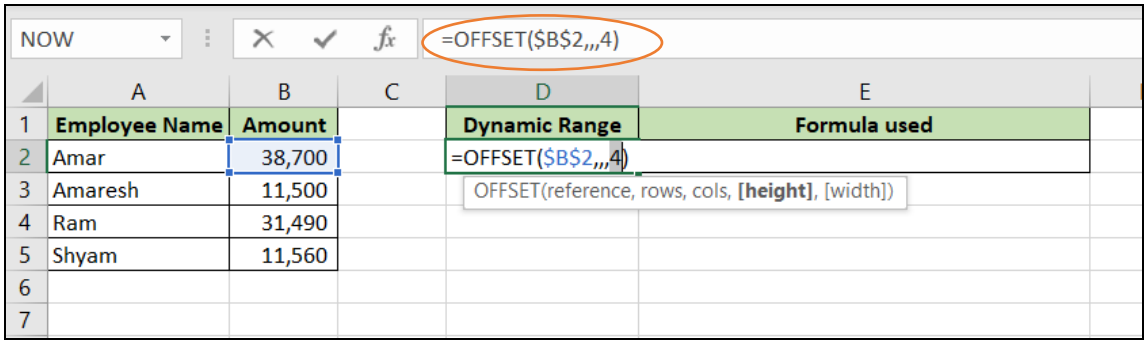


Figure 5.137 – OFFSET defines the range as 4 i.e. $B$2:$B$5 instead of $B$2:$B$12

Using a Dynamic Range to Define a Table Within a Table: OFFSET or INDEX

As shown in Figure 5.138, we have a data set with Products in the Second column. Some products are repeated because we have more than one Employee Name against each product. We would like to create a dynamic range of Employee names based on the product name in cell B13. If we have Apple in cell B13, we need the formula to define the range $C$2:$C$5. If we have Banana in cell B13, we need the formula to define the range $C$6:$C$7. This sort of formula could be used for a data validation drop-down list where we want the list of names to be based on a product name.

Option 1: Dynamic range using OFFSET function

B2 | Apple

	A	B	C	D
1	Sr. No.	Product	Employee Name	Amount
2	1	Apple	Amar	38,700
3	2	Apple	Shyam	11,560
4	3	Apple	Ramesh	23,430
5	4	Apple	Vijayendra	29,120
6	5	Banana	Amaresh	11,500
7	6	Banana	Abdul	39,690
8	7	Orange	Ram	31,490
9	8	Orange	Rajesh	29,240
10	9	Orange	Ganesh	10,850
11	10	Orange	Vijay	14,040
12				

Figure 5.138 – Table for Dynamic Range

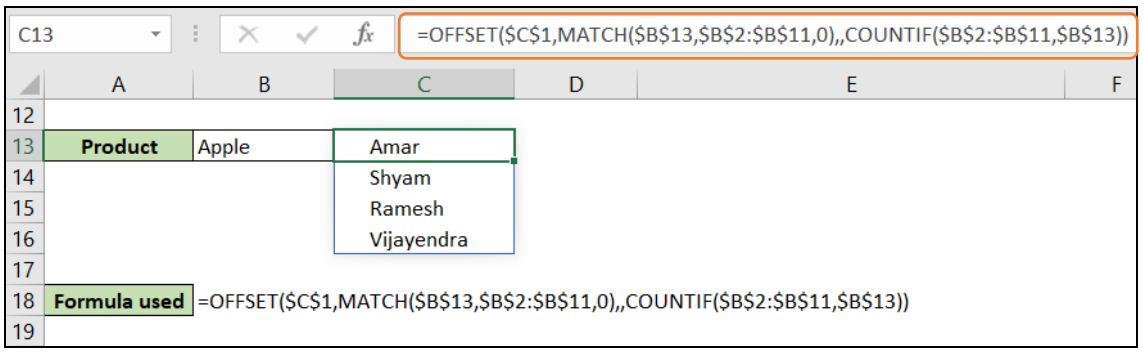


Figure 5.139 – OFFSET, MATCH & COUNTIF define the range as $C$2:$C$5

We change product name to Banana using Data Validation as shown in Figure 5.140 below:

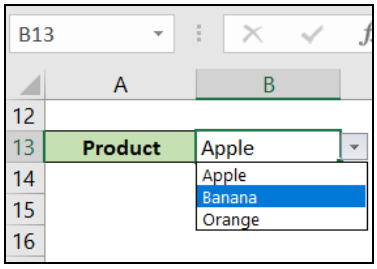


Figure 5.140 – Data Validation for Product name

The revised output, due to dynamic range:

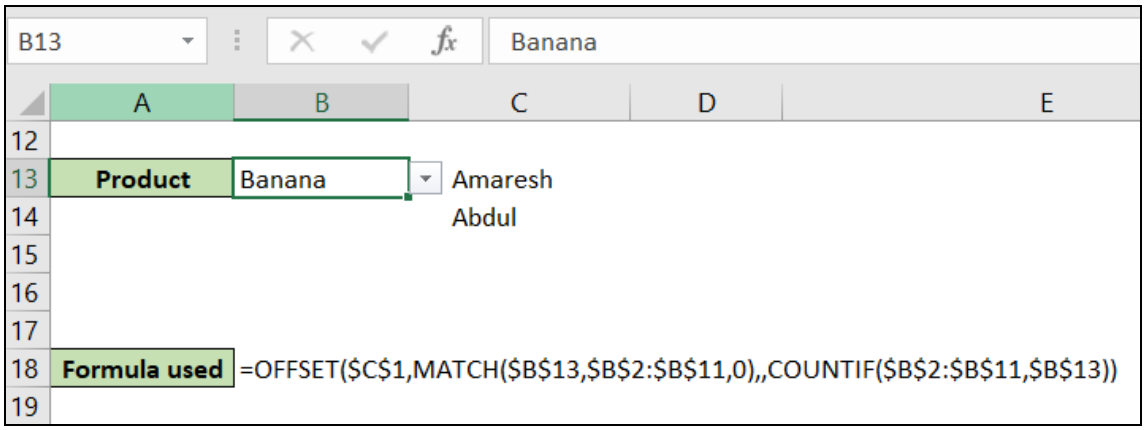


Figure 5.141 – OFFSET, MATCH & COUNTIF define the range for Banana as $C$6:$C$7

Option 2: Use of INDEX, MATCH & COUNTIF function to create dynamic range:

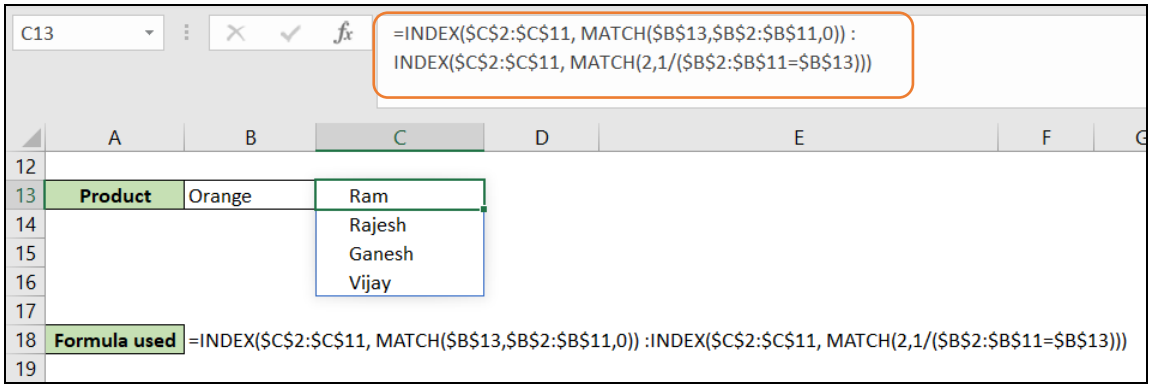


Figure 5.142 – 2 INDEX function, 2 MATCH function & COUNTIF define the range for Orange as $C$8:$C$11

Option 3: Use of INDEX, MATCH function and array operation to create dynamic range:

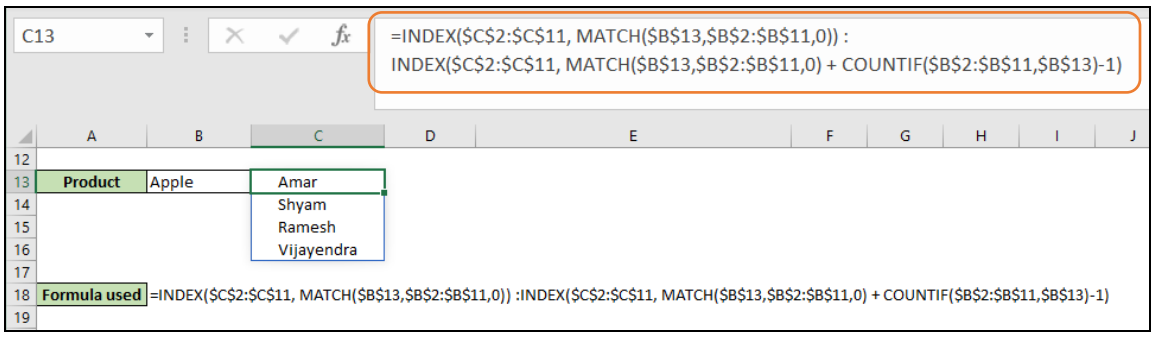


Figure 5.143 – 2 INDEX functions, 2 MATCH functions, and an array operation define range for Apple as $C$2:$C$5

Comparing OFFSET and INDEX for Creating Dynamic Ranges

OFFSET is a volatile function, and INDEX is not. If volatility is not important in a particular case, which one you use really is a matter of preference. Whereas the OFFSET function creates the dynamic range with inputs that define the starting position and the size of the range, INDEX creates the dynamic range by looking up cell references or a row and column reference.

- **FREQUENCY Array Function:**

The FREQUENCY array function is very simple, yet very powerful and versatile. In this way it is sort of like the SUMPRODUCT function because it has its base use, and then there are many other ways people apply it to solve a myriad of problems.

The base use for FREQUENCY is to count how many numbers fall into a set of categories. As shown in Figure 5.144, the goal in this example is to count numbers given in the range A4:A21 fall into the categories shown in the range C4:C7.

We enter the FREQUENCY function into the range E4:E8 using Ctrl+Shift+Enter. (In Office 365, we enter

FREQUENCY function in cell E13 and press Enter key. It spills the function in all the range till E17). It delivers a vertical array of numbers that reports the count of numbers that fall within each category.

For example, we can see that 5 numbers are <= 25 and only 4 numbers >= 90. If we change the Upper Limit or the data numbers, the FREQUENCY will automatically update.

There are two arguments for FREQUENCY: data_array contains the numbers to be counted in the range A4:A21, and bins_array contains the upper limits for each category in the range C4:C7. Notice that FREQUENCY delivers one more result than there are bins_array: there are 5 frequency numbers delivered by the FREQUENCY function, but only 4 upper limit numbers entered into the bins_array argument.

This occurs because this function adds an extra category in case we do not provide enough upper limits to count all the numbers in the data_array. As a user, we simply provide a set of upper limits, and an all-inclusive set of categories are automatically created.

Notice the brilliance of the categories created:

- The first category counts all the values less than or equal the first upper limit.
- The middle categories count between a lower limit and an upper limit. The lower limit is not included in the category. The upper limit is included in the category. For example, in 2nd upper limit, 25 will not be counted in the category "25 < Number <= 50".
- The last category catches all the values that are greater than the last upper limit.

E4 {=FREQUENCY(A4:A21,C4:C7)}

	A	B	C	D	E	F
1	data_array		bins_array		Array Funciton	
3	Numbers		Upper Limit	Category automatically created by FREQUENCY	Frequency	Formula used
4	66		25	Number <= 25	6	{=FREQUENCY(A4:A21,C4:C7)}
5	5		50	25 < Number <= 50	4	
6	86		75	50 < Number <= 75	3	
7	9		90	75 < Number <= 90	1	
8	49			90 < Number	4	
9	16					
10	28					
11	96				Office 365	
12	17		Upper Limit	Category automatically created by FREQUENCY	Frequency	Formula used
13	94		25	Number <= 25	6	=FREQUENCY(A4:A21,C13:C16)
14	20		50	25 < Number <= 50	4	
15	98		75	50 < Number <= 75	3	
16	65		90	75 < Number <= 90	1	
17	23			90 < Number	4	
18	44					
19	71					
20	100					
21	43					
22						

Figure 5.144 – FREQUENCY counts numbers, given a set of category upper limits

The fact that we can type in any set of numbers into the range A4:A21 and have all the numbers accounted for in the frequency count means the function is robust.

Details of How the FREQUENCY Array Function Works

- FREQUENCY counts how many numbers are in each category.
- The bins_array argument contains the upper values for the categories—numbers only.
- The data_array argument contains the values to count—numbers only.
- Keep in mind the following about categories:
  - Categories are automatically created. There is no visual indication of how the categories are organized.
  - The first category counts all the values less than or equal to the first upper limit.
  - The middle categories count between a lower limit and an upper limit. The lower limit is not included in the category. The upper limit is included in the category.
  - The last category catches all the values that are greater than the last upper limit.
  - There is always one more category than there are bins.
- FREQUENCY delivers a vertical array. If we need a horizontal array, use the TRANSPOSE function to convert a vertical array to a horizontal array.
- Because this is an array function, we must select the destination range before creating the formula and enter the formula with Ctrl+Shift+Enter. If we have **n** values in the bins_array argument, the selected destination range should contain **n+1** cells.
- FREQUENCY ignores empty cells and text.
- If there are duplicate bins, the duplicates get a count of zero.
- After the FREQUENCY function has been entered into the destination range with Ctrl+Shift+Enter, the resultant array of values is considered a single unit and individual cells cannot be deleted. We can delete all the values, but not just one.
- FREQUENCY can be used in larger array formulas and will deliver a vertical array.

Use TRANSPOSE Wrapped Around FREQUENCY for Horizontal array

Figure 5.145 shows that if we enter the FREQUENCY function into a horizontal range, the cells display the first value in the resultant array in all the cells.

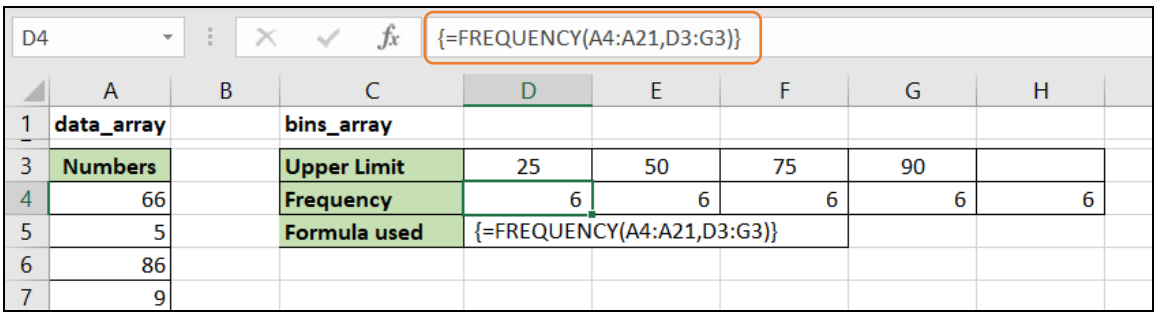


Figure 5.145 – FREQUENCY Function without TRANSPOSE Function

Figure 5.146 shows the evaluated results of the FREQUENCY function. Notice the array syntax semicolons (semicolons mean go down a row). We cannot display a vertical array horizontally unless we convert the semicolons to commas (commas mean go to the next column), which then converts the vertical array to a horizontal array.

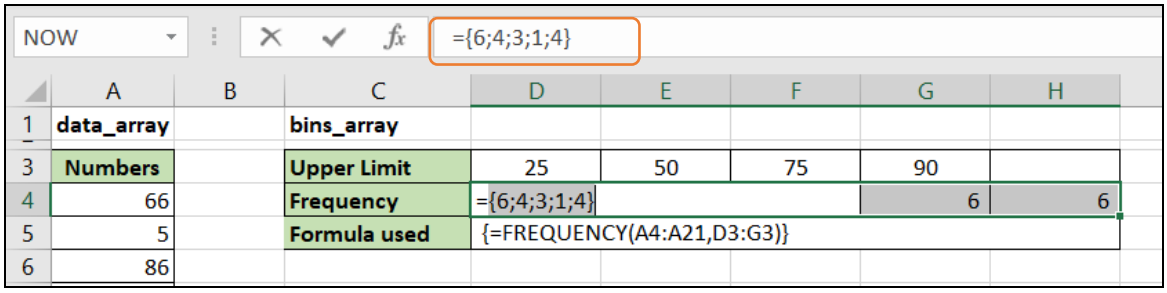


Figure 5.146 – FREQUENCY evaluates vertical resultant array by semicolon

Figure 5.147 shows how we can wrap the TRANSPOSE array function around FREQUENCY and convert the semicolons to commas. (Evaluate it by pressing F9 to see the commas.)

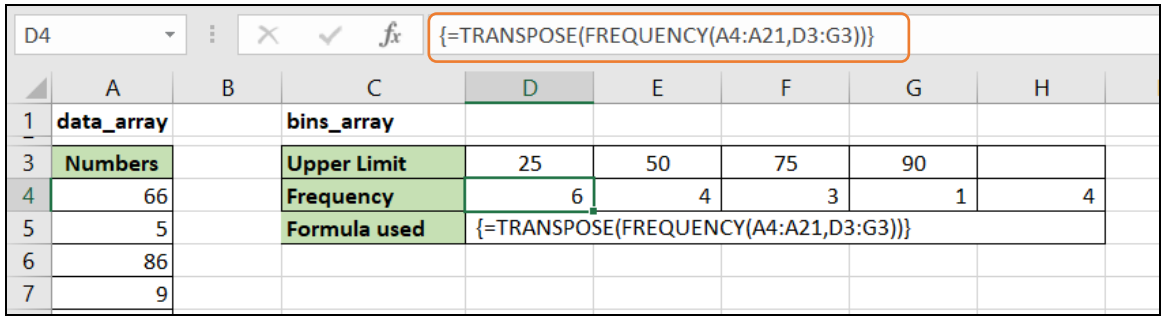


Figure 5.147 – FREQUENCY Function with TRANSPOSE Function

D4 =TRANSPOSE(FREQUENCY(A4:A21,D3:G3))

	A	B	C	D	E	F	G	H
1	data_array		bins_array					
3	Numbers		Upper Limit	25	50	75	90	
4	66		Frequency	6	4	3	1	4
5	5		Formula used	=TRANSPOSE(FREQUENCY(A4:A21,D3:G3))				
6	86							
7	9							

Figure 5.148 – FREQUENCY Function with TRANSPOSE Function in Office 365

Empty Cells, Text, and Duplicate Bin Values

As shown in Figure 5.149, the FREQUENCY function ignores empty cells and text. The FREQUENCY function entered into the range E4:E9 ignores the empty cell A8 and the Apple text in cell A12. It also shows that duplicate bins_array values will get a count of zero (25 < Number <= 50). In formulas for counting or extracting unique values, the fact that the duplicate bin values get a zero will be particularly useful.

E4 {=FREQUENCY(A4:A21,C4:C8)}

	A	B	C	D	E	F
1	data_array		bins_array			
3	Numbers		Upper Limit	Category automatically created by FREQUENCY	Frequency	Formula used
4	66		25	Number <= 25	5	{=FREQUENCY(A4:A21,C4:C8)}
5	5		50	25 < Number <= 50	3	
6	86		50	25 < Number <= 50	0	
7	9		75	50 < Number <= 75	3	
8			90	75 < Number <= 90	1	
9	16			90 < Number	4	
10	28					
11	96					
12	Apple					
13	94					
14	20					
15	98					
16	65					
17	23					
18	44					
19	71					
20	100					
21	43					
22						

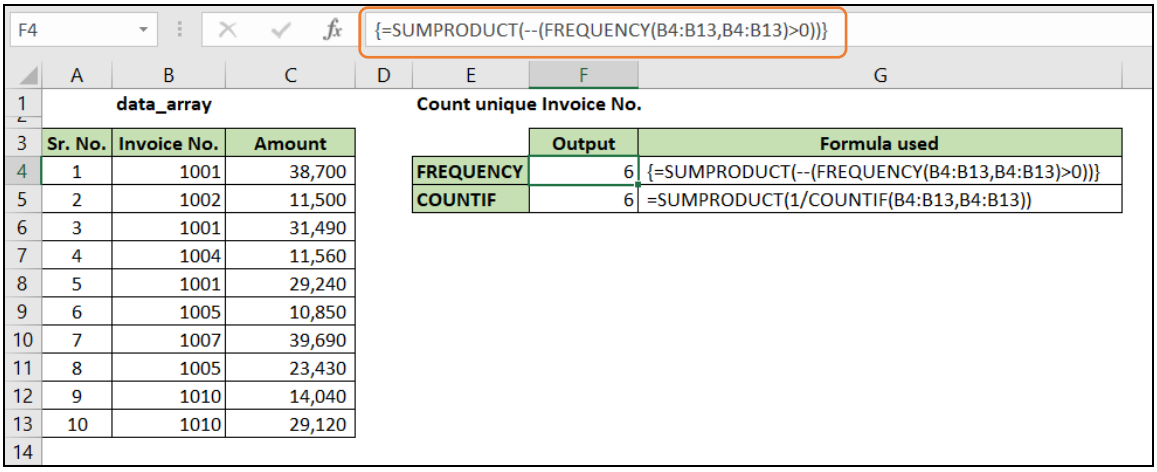
Figure 5.149 – FREQUENCY Function ignores empty cells and Text. It also gives 0 for duplicate bins

Using a Single-Cell Formula to Count Unique Numbers: FREQUENCY or COUNTIF

Figure 5.150 shows an sales data set that lists duplicate invoice numbers. The goal is to count the number of unique invoice numbers from the range B4:B13.

Figure 5.150 shows two formulas that both correctly calculate a count of 6. This section examines each of these formulas in turn. The key unique counting concept in the FREQUENCY formula is that we use the same range of values in both the data_array (the numbers to count) and bins_array (the upper limits for categories) arguments.

The key unique counting concept in the COUNTIF formula is that we use the same range of values in both the range (all the items to consider in the count) and criteria (the criteria for counting) arguments.



F4 {=SUMPRODUCT(--(FREQUENCY(B4:B13,B4:B13)>0))}

	A	B	C	D	E	F	G
1		data_array			Count unique Invoice No.		
3	Sr. No.	Invoice No.	Amount			Output	Formula used
4	1	1001	38,700		FREQUENCY	6	{=SUMPRODUCT(--(FREQUENCY(B4:B13,B4:B13)>0))}
5	2	1002	11,500		COUNTIF	6	=SUMPRODUCT(1/COUNTIF(B4:B13,B4:B13))
6	3	1001	31,490				
7	4	1004	11,560				
8	5	1001	29,240				
9	6	1005	10,850				
10	7	1007	39,690				
11	8	1005	23,430				
12	9	1010	14,040				
13	10	1010	29,120				
14							

Figure 5.150 – Using FREQUENCY Function to count unique Invoice numbers

5.16 SUMMARY

In this Chapter, we learned about Cell References. It can be Absolute Cell Reference, Relative Cell Reference and Mixed Cell Reference, which needs to be used depending upon the requirement.

There are several Text Functions in Excel – LEFT Function, RIGHT Function, MID function, LEN Function, FIND Function, SEARCH Function, SUBSTITUTE Function, REPLACE Function, TRIM Function and CONCATENATE. LEFT and RIGHT Function extracts a given number of characters from the left and right side of the selected text string respectively. But MID Function extracts the characters but not from left or right but from middle of a text string, given a starting position and number of characters. The LEN function returns the number for characters in a text string. The FIND Function returns the starting position of one text string within another text string. The SEARCH Function takes the same arguments as FIND Function. But SEARCH Function is not case sensitive. TRIM function removes all spaces from a text string except for single space between words. The CONCATENATE Function joins several text strings into one text string. Similarly "&" (ampersand) can also be used to join several text strings into one text string.

Flash Fill is only available in Excel 2013 and later versions of Excel. It is part of AutoFill feature. It automatically fills the data when it senses a pattern in the data. It does not matter whether you are extracting data in next column or several columns away as long as it is withing the same database range.

ROUND Function rounds the decimal places of the number by a specified number. There are three Round Functions i.e. ROUND, ROUNDUP and ROUNDDOWN.

In Excel, dates and times are not treated as text entries. Any entry with a format that looks like one the date and time number format being used by the program is converted into a serial number. We can use DATE Function to create a complete date from separate year, month and day information. DATE provides some extra flexibility with the month number. Negative month numbers are subtracted from the specified year. EDATE function returns a serial number of the date that is the indicated number of months before or after the start date in the function. The EOMONTH function returns the serial number of the last day of the month before or after a specified number of months. The DATEDIF function calculates the number of days, months, or years between two dates. This is an undocumented function; that is, we won't see it in the Insert Function dialog box, and we cannot find it in the Excel Help system. The only thing we can do is to remember how to enter it. The TEXT Function converts a value to text in a specific number format. It is useful in extraction of Day, Month from a give date. This function returns the reminder after a number is divided by a divisor. Normally, if we divide a number by some divisor and if the dividend is not completely divisible by the divisor, the quotient is expressed in decimal form. However, sometimes we are interested in obtaining the remainder separately. In such a case, we can make use of Mod function.

There are 4 main logical functions in Excel i.e. IF, AND, OR and NOT. However, there are other sub-logical functions derived from these 4 main logical functions, which are XOR, IFERROR, IFNA, NESTEDIF and ISBLANK. The IF Function is one of the most important Logical Excel Function. It helps to check a condition and give the output either TRUE or FALSE according to the condition. In Nested IF Function, we use multiple IF within IF based on conditions. The AND Function is used in a scenario where we need to check more than one condition. If all the logical arguments are valid, it returns value as "True" else it returns value as "False". The OR Function is used in a scenario where we need to check more than one condition. Even if anyone logical argument is valid, it returns value as "True". It will give FALSE with all the arguments are invalid. XOR Function will return TRUE only if majority condition is FALSE. XOR returns TRUE or FALSE depending on the number of conditions. The NOT Function returns the opposite of a given value. It means this function changes FALSE to TRUE or TRUE to FALSE. IFERROR Function will return "value if error" in case output value is expected to be an error, otherwise, it will return the value, whereas IFNA Function returns the value we specify if the expression resolves to #N/A, otherwise it returns the result of the expression.

The SUMIF Function adds the cells specified by a given criteria or condition, whereas SUMIFS Function sums a range of cells based on more than one condition or criteria. The COUNTIFS Function counts the number of cells in a range that meets all the given criteria. This function also works on AND concept. It means it will count only if all the criteria are met. SUMPRODUCT function returns the sum of the products of corresponding ranges or arrays.

VLOOKUP Function looks for a value in left most column of a table and then returns a value in the same row from a column specified by us. We learned how to use VLOOKUP function under various scenarios i.e. extracting single column value, extracting all values with one master formula, using VLOOKUP alongwith MATCH function. Similar to VLOOKUP function, HLOOKUP Function looks for a value in Top row of a table or array of values and then returns a value in the same column from a row specified by us. The MATCH Function returns the relative position of an item in an array that matches a specified value in a specified order. The INDEX Function returns a value or reference of the cell at the intersection of a particular row and column in a given range. The CHOOSE Function chooses a value or action to perform from a list of values based on an index number. The CHOOSE function is ideal for converting a value to a literal.

In Excel, errors are returned when the user is inserting wrong data or asking the Excel to do something that cannot be done. #NAME? error appears when Excel can't find a named range. #DIV/0 error appears when a mathematical operation attempts to divide by zero or an empty cell. Usually, this occurs because a COUNT or SUM results in zero and another cell is operating on its result. #N/A error is returned when a function like MATCH or VLOOKUP cannot find the value in the given range.

In Formula Auditing Toolbar, we learned about formula auditing techniques such as Trace Precedents, Trace Dependents, Evaluate Formulas, Error Checking option and use of Watch Window.

We also learned about Database functions such as DGET, DSUM, DCOUNT, DAVERAGE. To use the database functions, we need to put data into a structured format. To make the best of the database functions, we need to get data into a contiguous area of rows and columns. Each row is a record, and each column is a field. The top row contains labels that identify the fields. DCOUNT function counts the cells containing numbers in the field (column) of the records in the database that matches the condition we specify. DSUM function adds the numbers in the field (column) of records in the database that match the conditions we specify. In database function, we need to learn criteria writing options. In case, we right "Ram", then it will return all the names which begin with name "Ram". In case, we want exact match, then we have to write "=Ram" in criteria option. The DGET Function is a unique database function. It does not perform a calculation but checks for duplicate entries. The function returns one of three values i.e. (1) If one record matches the criterion, DGET returns the criterion. (2) If more than one record matches the criterion, DGET returns the #NUM! error. (3) If no records match the criterion, DGET returns the #VALUE! error.

We also learned about Array formulas and Array constants. An array in Excel is a collection of two or more items. The items can be text or numbers and they can reside in a single row or column, or in multiple rows and columns. An array formula is a formula that contains an operation (math, comparative, join, or function argument) on an array of items rather than on single items, and, the operation delivers a resultant array of items rather than a single item. Dynamic array formulas, whether they are using existing functions or the dynamic array functions, only need to be input into a single cell, then confirmed by pressing Enter. Earlier, legacy array formulas require first selecting the entire output range, then confirming the formula with Ctrl+Shift+Enter. They are commonly referred to as CSE formulas. We also learned about single-cell array formula, multi-cell array formula and multi-cell array formula (TRANSPOSE function used). An array constant is simply a set of static values. These values never change when you copy a formula to other cells or values. Above we have seen few examples of array constants. There are 3 types of array constants – (1) Column Array Constants (Vertical Array Constants), (2) Row Array Constants (Horizontal Array Constants) and (3) Table Array Constants (Two-Way Array Constants). We also saw the examples for using Array Constant in different scenarios - Sum N largest / smallest numbers in a range and Array formula to count cells with multiple conditions. Also learned about usage of VLOOKUP, SUM, AND, OR, Double Unary Operator in Excel array formulas. A double dash (--), which is technically called the double unary operator, is used to convert non-numeric Boolean values (TRUE / FALSE) returned by some expressions into 1 and 0 that an Excel array function can understand. We learned about advance array formulas such as OFFSET, TRANSPOSE and FREQUENCY functions. OFFSET function returns a reference to a range that is a given number of rows and columns from a given reference. It gets the address of the cell that is offset from another cell by a certain number of rows and/or columns. OFFSET can be used in dynamic ranges as well. The FREQUENCY array function is very simple, yet very powerful and versatile. The base use for FREQUENCY is to count how many numbers fall into a set of categories.

5.17 REFERENCES

[1] support.microsoft.com

[2] A Definitive Guide to Microsoft Excel 2019 by Derrick Richard

[3] Microsoft Excel Formulas & Functions for Dummies by Ken Bluttman

[4] Excel 2019 – Advanced Topics - Leverage More Powerful Tools to Enhance Your Productivity by Nathan George

[5] EXCEL 2019 - A Comprehensive Beginners Guide to Learn Excel 2019 Step by Step from A – Z by Alexander Cane

[6] Excel XLOOKUP and Other Lookup Functions by Nathan George

[7] Ctrl+Shift+Enter_ Mastering Excel Array Formula by Mike Girvin

[8] Ablebits.com (www.ablebits.com/office-addins-blog/2015/02/25/array-formulas-functions-excel/)

CHAPTER

6 ANALYSING AND VISUALISING DATA IN EXCEL

LEARNING OBJECTIVES

- Understanding Conditional formatting in Excel
- Understand the visualisation with Excel charts
- Understanding the use and benefits of PivotTables
- Understand use of Pivot Charts
- Understanding How to Build a Dynamic Dashboard
- Understand the power of Power Pivots

6.1 INTRODUCTION

Often when analysing data you want to highlight data that meets certain conditions, that is exactly what Excel conditional formatting does, It is a really powerful feature that applies different formats to data that meets certain conditions. It helps you highlight the critical information in your spreadsheets and identify variances of cells' values with a quick glance. It gives a visual feel to anomalies in data.

Though could be intimidating at first but is straightforward and easy to use once you understand it.

In several situations, you need to analyse data from a multi-dimensional perspective. You might need to slice & dice data. Like out of a Sales report of sales made in different periods, in different departments, of different products, by different salesmen, of different amounts; you need an overview of sales period wise, salesmen wise, product wise, with different subgroups; for that matter, you might have a hundred data points to track. All this is possible using a Star feature of Excel called PivotTables.

The Pivot Tables tool is one of the most powerful yet intimidating features in Excel. Pivot tables allow you to turn your data inside out, upside down, sideways, and backward, quickly summarize and analyze large amounts of data in lists and tables--independent of the original data layout in your spreadsheet--by dragging and dropping columns to different rows, columns, or summary positions.

Various features of Excel like PivotTables, Charts, and Slicers combine to help you build a Dynamic Dash Board. Power Pivot overcome the deficiencies in Pivot tables and can be used to analyse a huge amount of data

6.1.1 DATA VISUALISATION

These days data are coming at such a speed and volume that comprehending it and taking decisions based on that are becoming increasingly difficult. The solution is Data Visualisation which makes understanding data easy through visual elements.

Data visualisation is a graphical representation of information and data. It involves images that can communicate the relationship between data which could be useful to viewers, for this you can use visual elements like charts, graphs, and maps.

Your eyes are easily drawn to patterns and colours; therefore, Data visualisation provides an easy way to see and understand trends, outliers, and patterns in data.

There are many data visualization tools like Excel, Tableau, Power BI, D3.js, etc. Excel also offers many Data Visualisation Capabilities.

6.2 CONDITIONAL FORMATTING

To make your data easy to understand and highlight based on certain parameters you can use Conditional formatting. It allows you to automatically apply formatting—such as **colouring the cells**, **finding the duplicate values**, highlighting a cell or row, inserting icons on **data bars, and much more**.

As the name implies, it enables you to apply cell formatting automatically, based on certain condition i.e. the contents of the cells. You can apply conditional formatting in such a way that when you enter or change a value in the range, Excel automatically examines the value and checks the conditional formatting rules for the cell and work accordingly the conditional formatting set for the range.

For example, all values below 10 in a range will have a yellow background color and other data shall be not filled by any colour.

It is an easy way to quickly identify cell entries or cells of a particular type that might be of use to you.

There are different choices available for you under conditional formatting such as:

- **Highlighting Cell**: It includes highlighting cells that are greater, lower, or between the particular value, set by the rules. It may also highlight value containing a specific text string or a specific date.
- **Top Bottom values**: It highlights the top and bottom items as specified in the rule by the user. It is also useful in highlighting above or below average items.
- **Data Bars**: It helps in applying graphic data bars directly in the cells, proportional to the cell's value.
- **Colour Scales**: It helps in applying background colour, proportional to the value of the cell as defined in the rules.
- **Icon Sets**: It helps in Displaying icons directly in the cells, based on the rules set by the user based on cell value.
- **Creating, Managing, and Clearing Rule**: It enables you to create a new rule, edit an existing rule, or clearing the rule no longer needed, including rules based on a logical formula.

6.2.1 HOW TO USE

Conditional formatting is available in **Home > Styles**

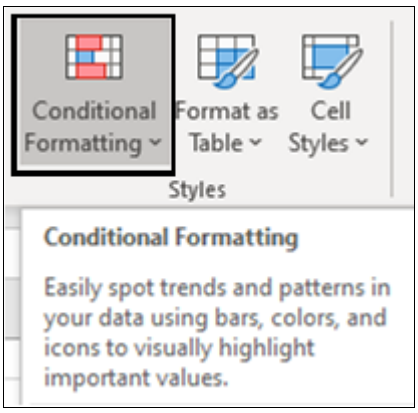


Figure 6.2.1 Conditional Formatting

6.2.2 HIGHLIGHT CELL

- You want to highlight all the negative revenue growth that is wherever differences are less than 0.

Select the cells you need to work upon for conditional formatting. You can select n numbers of cells, rows, or columns based on your requirement to apply such conditional formatting upon. Since you need to apply conditional formatting only to the numbers less than 0, you choose Highlight Cells Rules > Less Than.

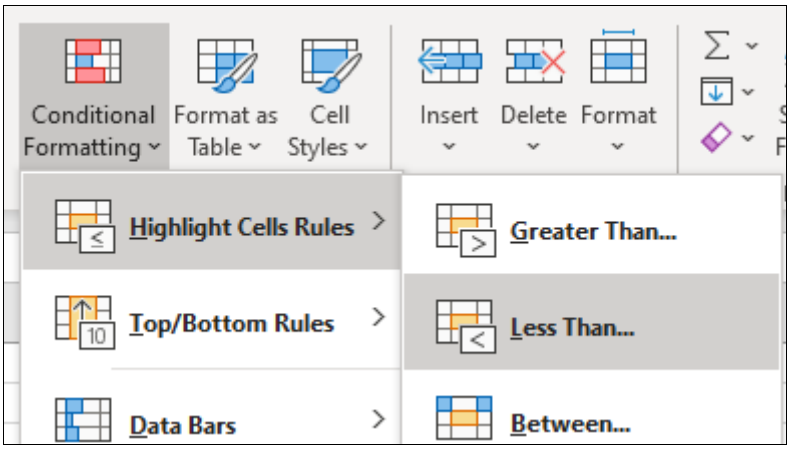


Figure 6.2.2 Highlight Cell Rules

Enter the value in the box in the right-hand part of the window under "**Format cells that are LESS THAN",** in your case you type 0. Excel will highlight cells meeting your condition and tells you to decide the formatting, it will bring up some readymade formatting like "**Light Red Fill With Dark Red Text"** but you have the option to change or decide your own custom format.

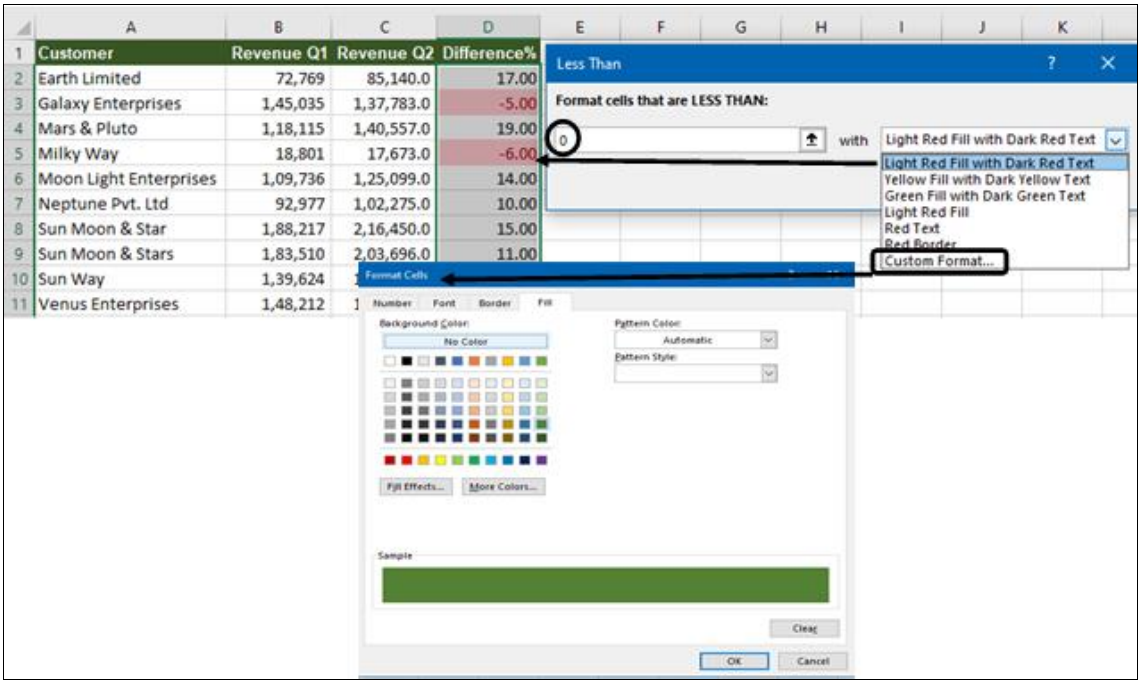


Figure 6.2.3 Format Cells

6.2.3 TOP BOTTOM RULES

You want to highlight the top 5 Borrowers in a bank so that you can go for in-depth auditing in those cases, you can use Conditional formatting for highlighting the top 5 borrowers. Go to **Conditional Formatting> Top/Bottom Rules**, Select 5, select formatting.

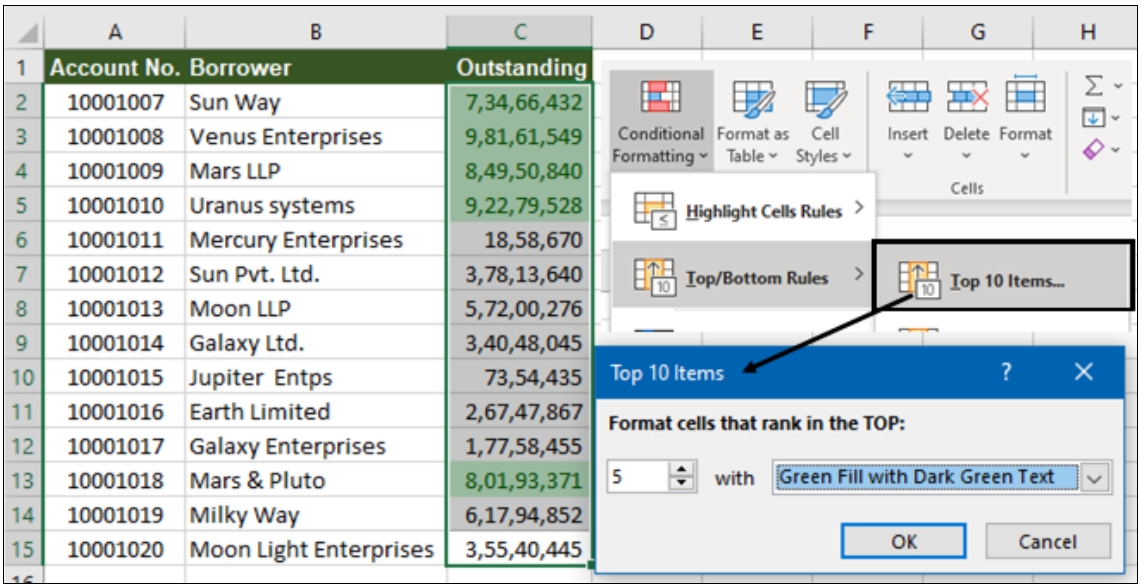


Figure 6.2.4 Top Bottom Rules

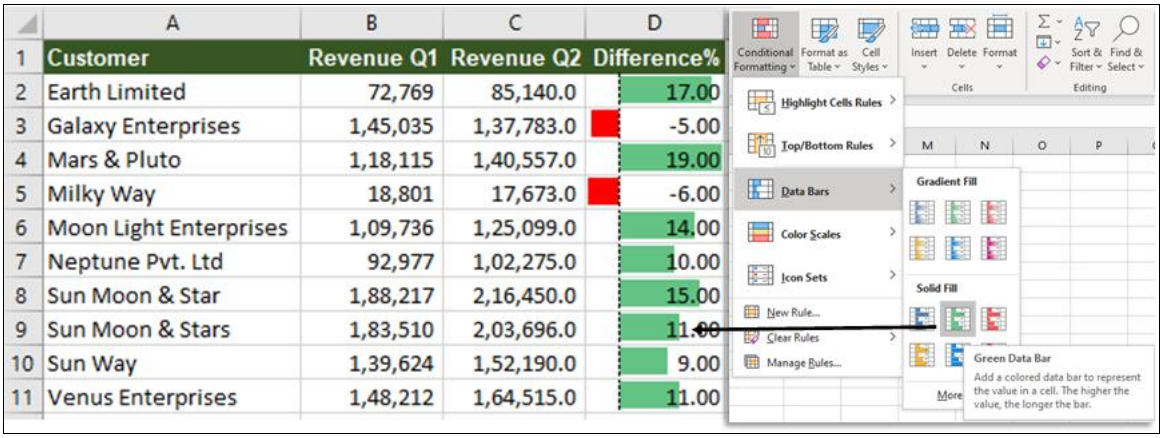
6.2.4 DATA BARS

A data bar helps you see the value of a cell relative to other cells. The length of the data bar represents the value in the cell. A longer bar represents a higher value, and a shorter bar represents a lower value. If you are not looking for specific values or text in your data, then data bars are perfect for showing you the top and bottom values, performers, etc. of your dataset.

Taking the earlier example, you want to review all the differences at one go. Go to **Conditional Formatting > Data Bars** select Gradient fill or Solid Fill:

- **Gradient Fill** means that the colour of each bar fades from the colour you chose to a lighter version of that colour.
- **Solid Fill** simply means that the entire bar is the same colour.

Let you say, if you select green then green bars will appear, in each of the cells, the length of the bar represents a high or low value and negative bars are represented by green value.



	A	B	C	D
1	Customer	Revenue Q1	Revenue Q2	Difference%
2	Earth Limited	72,769	85,140.0	17.00
3	Galaxy Enterprises	1,45,035	1,37,783.0	-5.00
4	Mars & Pluto	1,18,115	1,40,557.0	19.00
5	Milky Way	18,801	17,673.0	-6.00
6	Moon Light Enterprises	1,09,736	1,25,099.0	14.00
7	Neptune Pvt. Ltd	92,977	1,02,275.0	10.00
8	Sun Moon & Star	1,88,217	2,16,450.0	15.00
9	Sun Moon & Stars	1,83,510	2,03,696.0	11.00
10	Sun Way	1,39,624	1,52,190.0	9.00
11	Venus Enterprises	1,48,212	1,64,515.0	11.00

Figure 6.2.5 Data bars

6.2.5 ICON SETS

You can use an icon set to mark and classify data into three to five categories separated by some value limits. Each icon represents a range of values. For example, in the 3 Arrows icon set, the green up arrow represents higher values, the yellow sideways arrow represents middle values, and the red down arrow represents lower values.

Taking the above example, you can highlight differences by using icons that way you can easily see which cells have lower growth by just looking at the direction, and later on, you can also filter by the criteria of an icon . Go To **Conditional Formatting > Icon sets** select an appropriate icon set, as an example 3 arrows, now all your negative values are highlighted with **red down arrow.**

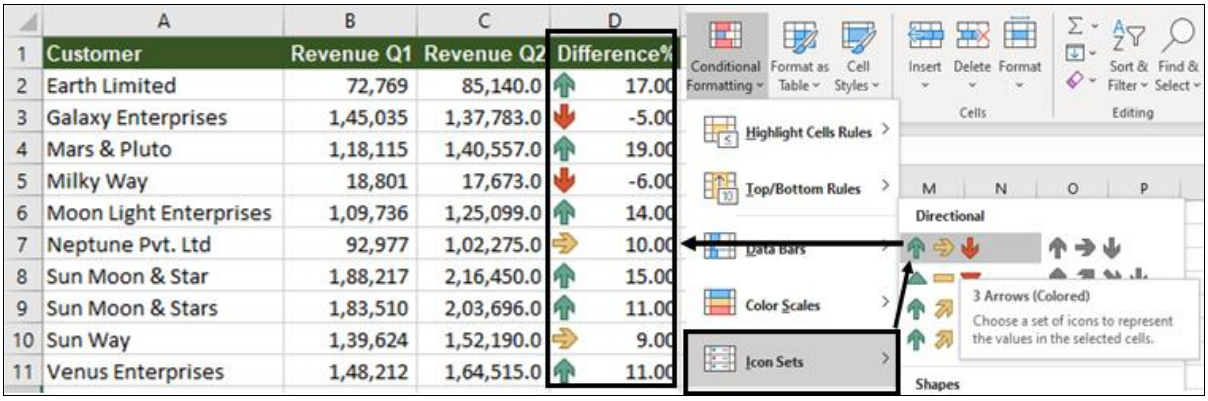


Figure 6.2.6 Icon Sets

6.2.6 COLOR SCALES

You can use Color scales sometimes also referred to as heat maps, which are visual guides that help you understand data distribution and variation. A three-color scale helps you compare a range of cells by using a gradation of three colours. The shade of the colour represents higher, middle, or lower values. For example, in a green, yellow, and red colour scale, you can specify that higher value cells have a green colour, middle-value cells have a yellow colour, and lower value cells have a red colour.

You have expenditure variations from the budget for four quarters of different sections of a company and you want to review which sections in which quarters variations were less and in which sections variations were high, a convenient tool you can use is color scales. Go to **Conditional Formatting > Color Scales,** select the colour scale as an example Red yellow-green colour scale. All the high variances will be marked red and the intensity of red colour will depend on how high value is, and all the low variances will have a colour green.

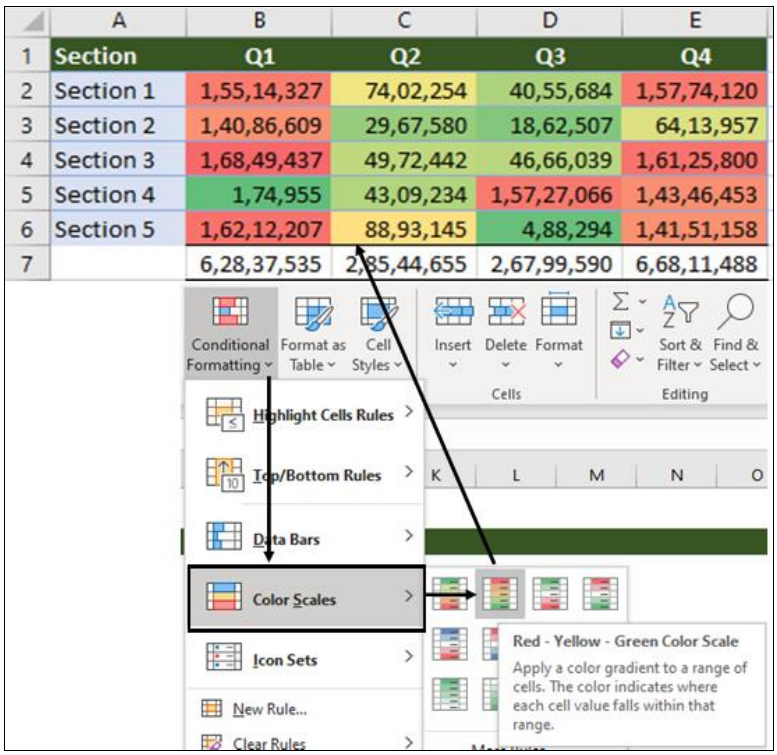


Figure 6.2.7 Color Scales

***CASE STUDY:** In an audit, your principal CA Chandiwala gives you a Sales register as shown in Figure and wants you to mark all the duplicate invoices which could have been used by the Sales team to pad up sales so that they can claim commission. You are a smart auditor, is there a smart way to mark duplicate invoices?*

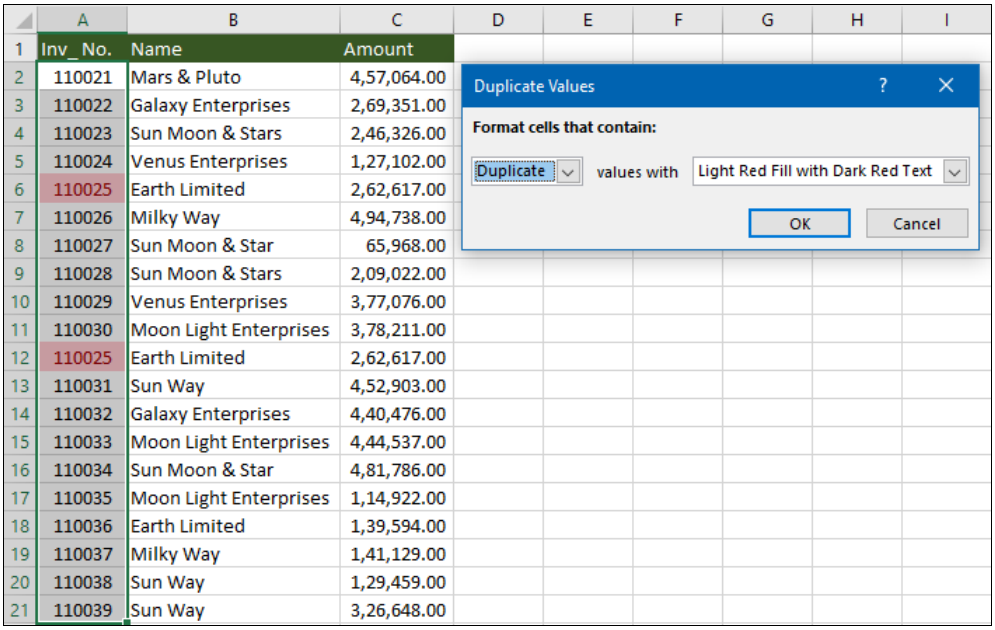
	A	B	C
1	Inv_No.	Name	Amount
2	110021	Mars & Pluto	4,57,064.00
3	110022	Galaxy Enterprises	2,69,351.00
4	110023	Sun Moon & Stars	2,46,326.00
5	110024	Venus Enterprises	1,27,102.00
6	110025	Earth Limited	2,62,617.00
7	110026	Milky Way	4,94,738.00
8	110027	Sun Moon & Star	65,968.00
9	110028	Sun Moon & Stars	2,09,022.00
10	110029	Venus Enterprises	3,77,076.00
11	110030	Moon Light Enterprises	3,78,211.00
12	110025	Earth Limited	2,62,617.00
13	110031	Sun Way	4,52,903.00
14	110032	Galaxy Enterprises	4,40,476.00
15	110033	Moon Light Enterprises	4,44,537.00

Figure 6.2.8 sales data

Strategy:

You can use the feature of **Conditional formatting** highlighting the option to mark duplicates.

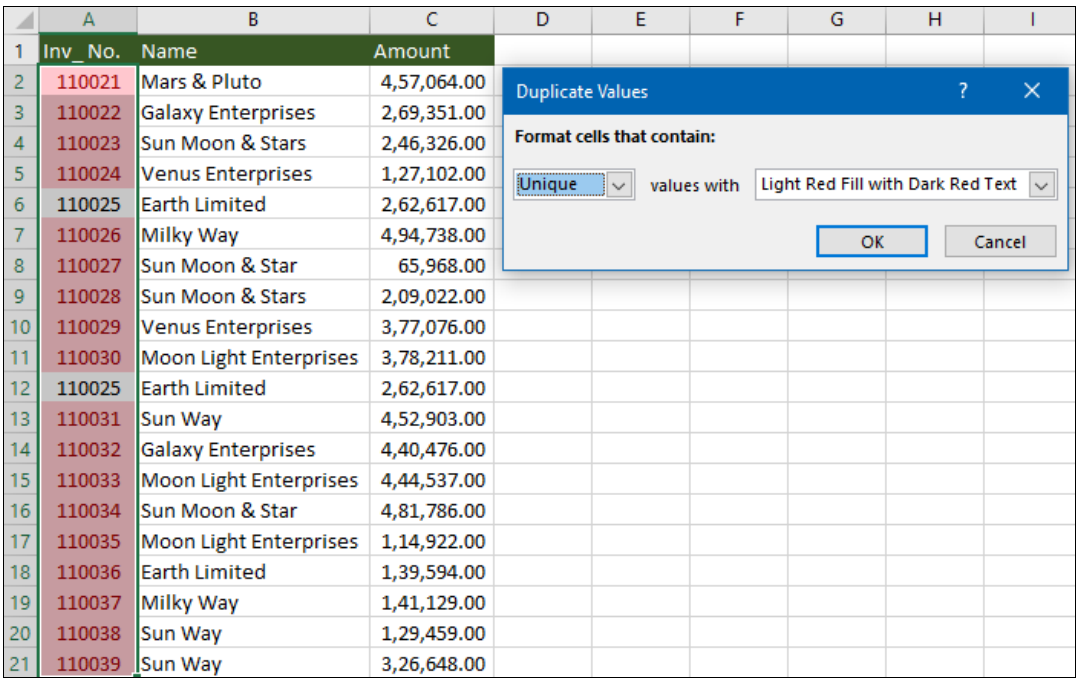
1. Select the data A2:A21
2. Go to Home > Conditional formatting > Highlight Cell Rules
3. Select Duplicate
4. Select the formatting colour



	A	B	C
1	Inv_ No.	Name	Amount
2	110021	Mars & Pluto	4,57,064.00
3	110022	Galaxy Enterprises	2,69,351.00
4	110023	Sun Moon & Stars	2,46,326.00
5	110024	Venus Enterprises	1,27,102.00
6	110025	Earth Limited	2,62,617.00
7	110026	Milky Way	4,94,738.00
8	110027	Sun Moon & Star	65,968.00
9	110028	Sun Moon & Stars	2,09,022.00
10	110029	Venus Enterprises	3,77,076.00
11	110030	Moon Light Enterprises	3,78,211.00
12	110025	Earth Limited	2,62,617.00
13	110031	Sun Way	4,52,903.00
14	110032	Galaxy Enterprises	4,40,476.00
15	110033	Moon Light Enterprises	4,44,537.00
16	110034	Sun Moon & Star	4,81,786.00
17	110035	Moon Light Enterprises	1,14,922.00
18	110036	Earth Limited	1,39,594.00
19	110037	Milky Way	1,41,129.00
20	110038	Sun Way	1,29,459.00
21	110039	Sun Way	3,26,648.00

Figure 6.2.9 Duplicates

5. All the duplicate invoice numbers will be highlighted.
6. There is also an option of highlighting unique values, just select **Únique'** in the dropdown as shown in Figure



	A	B	C
1	Inv_No.	Name	Amount
2	110021	Mars & Pluto	4,57,064.00
3	110022	Galaxy Enterprises	2,69,351.00
4	110023	Sun Moon & Stars	2,46,326.00
5	110024	Venus Enterprises	1,27,102.00
6	110025	Earth Limited	2,62,617.00
7	110026	Milky Way	4,94,738.00
8	110027	Sun Moon & Star	65,968.00
9	110028	Sun Moon & Stars	2,09,022.00
10	110029	Venus Enterprises	3,77,076.00
11	110030	Moon Light Enterprises	3,78,211.00
12	110025	Earth Limited	2,62,617.00
13	110031	Sun Way	4,52,903.00
14	110032	Galaxy Enterprises	4,40,476.00
15	110033	Moon Light Enterprises	4,44,537.00
16	110034	Sun Moon & Star	4,81,786.00
17	110035	Moon Light Enterprises	1,14,922.00
18	110036	Earth Limited	1,39,594.00
19	110037	Milky Way	1,41,129.00
20	110038	Sun Way	1,29,459.00
21	110039	Sun Way	3,26,648.00

Figure 6.2.10 Unique values

Gist: You have highlighted duplicate values using conditional formatting.

Commands learnt: **Home > Conditional formatting > Highlight Cell Rules** select duplicate values.

6.2.7 CREATING CONDITIONAL FORMATTING RULE

You can set up a conditional formatting rule on your own and can make a rule based on a formula also. An important thing to remember is that if the result of the formula is true Conditional formatting will give the desired format.

***CASE STUDY**: In a forensic investigation your principal CA Chandiwala gives you a sheet as shown in the figure below which contains some amounts relating to some customers it is suspected that some of these figures are in fact text that is why not reflected in the total amount. You are required to find which of these amounts is not a number. You are a smart auditor, is there a smart way to locate Numbers stored as text so that they are not included in the total?*

	A	B
1	Customer	Amount
2	Earth Limited	72769
3	Galaxy Enterprises	145035
4	Mars & Pluto	128536
5	Milky Way	18801
6	Moon Light Enterprises	109736
7	Neptune Pvt. Ltd	92977
8	Sun Moon & Star	188217
9	Sun Moon & Stars	183510
10	Sun Way	139624
11	Venus Enterprises	148212
12		1098881

Figure 6.2.11 data to be tested for text

Strategy:

You can use the feature of Custom **Conditional formatting** based on a formula. You will build conditional formatting from scratch:

1. Select the data **B2:B11**
2. Go to **Home > Conditional formatting > New rule**

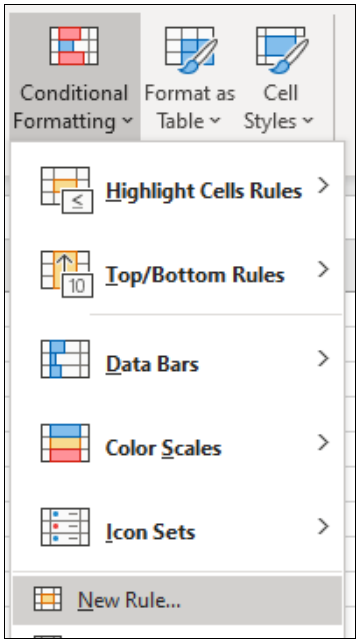


Figure 6.2.12 New Rule

3. A new dialog box appears,
4. Under '**Select a rule type'** select **'Use a formula to determine which cells to format'**
5. In '**Format values where this formula is true'** you can define a formula whose result if true the cell would be highlighted.
6. There is a function in excel ISTEXT(value) which examines if Value refers to text and gives results in true or false, If the value is true it means that is a text value
7. You will enter **'=ISTEXT(B2)'**

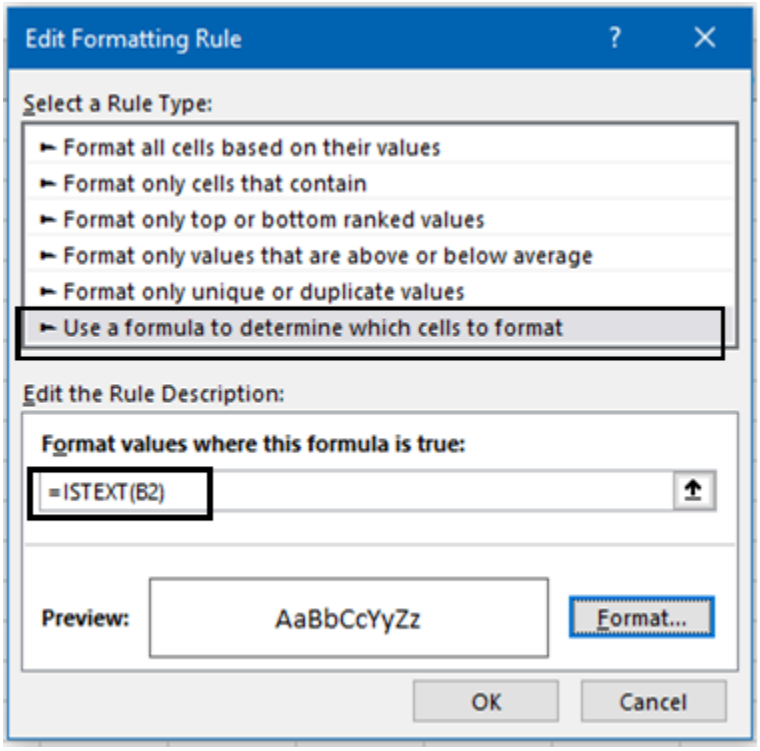


Figure 6.2.13 Rule-based on Formula

8. Click on Format, and you can select the Font, Border, and Fill tabs of your choice, further you can also switch between different options such as font style, pattern colour and fill effects to set up the format that works best for you. When done, click the OK button.

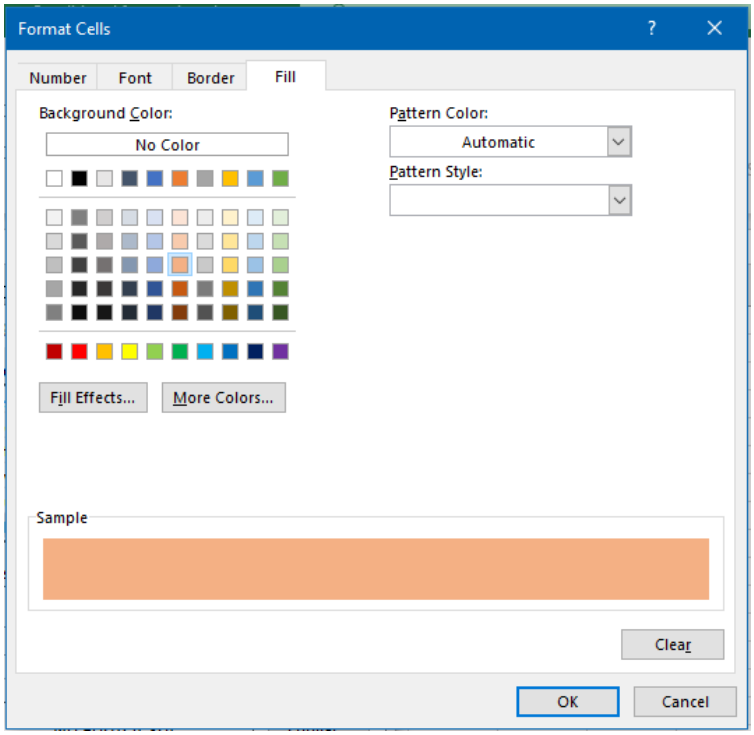


Figure 6.2.14 Select a Format

9. The values which are text would be highlighted

	A	B
1	Customer	Amount
2	Earth Limited	72769
3	Galaxy Enterprises	145035
4	Mars & Pluto	128536
5	Milky Way	18801
6	Moon Light Enterprises	109736
7	Neptune Pvt. Ltd	92977
8	Sun Moon & Star	188217
9	Sun Moon & Stars	183510
10	Sun Way	139624
11	Venus Enterprises	148212
12		1098881

Figure 6.2.15 Format applied

Gist: You have highlighted Numbers stored as text using conditional formatting and learnt how to use formulas in conditional formatting.

Commands learnt: **Home > Conditional formatting >New Rules** select using **"Use a formula to determine which cells to format."**

6.2.8 MANAGE CONDITIONAL FORMATTING RULES

You can manage a rule by selecting the range of cells, click Conditional Formatting >manage Rules it will open a dialog box 'Conditional Manager Rules Manager', select the rule and click the Edit Rule button.

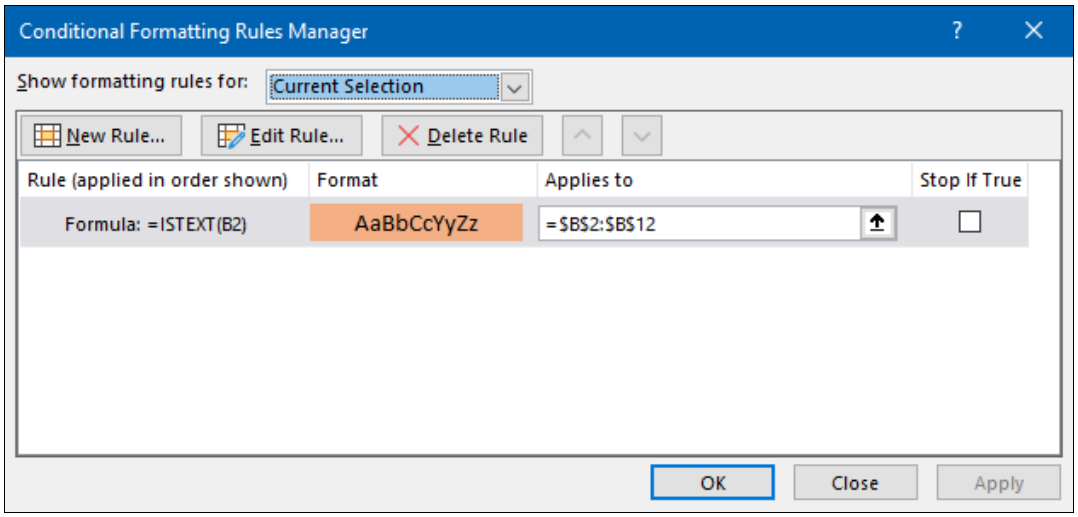


Figure 6.2.16 Manage Conditional formatting Rules

6.2.9 DELETE CONDITIONAL FORMATTING RULES

You can delete a rule

Method 1: Open the Conditional Manager Rules Manager, select the rule and click the **Delete** Rule button.

Method 2: Select the range of cells, click **Conditional Formatting > Clear Rules** and choose one of the available options.

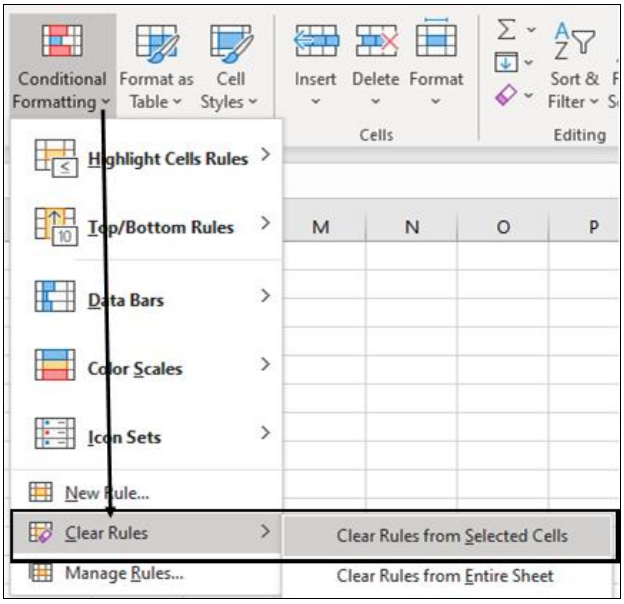


Figure 6.2.17 Delete a Conditional Formatting Rule

6.3 EXCEL CHARTS

Patterns and relationships between numbers are easier to interpret when they are displayed in visual chart form. Excel chart, sometimes also referred to as graph, is a graphical representation of numeric data where the data is represented by symbols such as Column, Pie Bar, lines, etc. Graphs in Excel give a visual feel and make it possible to understand large amounts of data or the relationship between different data subsets.

You can create a chart in Excel, based on the numeric data in a worksheet, Your data needs to be organised in rows or columns, and Microsoft Excel will automatically determine the best way to plot the data in your graph, You can of course change it later.

For a Chart

- Normally the column headings, alternately data in the first column are used in the chart legend.
- The data in the first column (or columns headings) is used as labels along the **X axis** of your chart.
- The numerical data in other columns are used to create the labels for the **Y axis**.

Step 1

- Select all the data you want to include in your Excel chart. Select the column headings, they will appear either in the chart legend or axis labels.
- You can select only one cell and Excel will automatically include all contiguous cells that contain data.
- If you want to create a graph based on the data in non-adjacent cells, select the first cell or a range of cells, hold down the CTRL key and select other cells or ranges.

Step 2

- You can add a Chart on the worksheet, go to the **Insert tab > Charts group**,
- Click on a chart type you would like to create.
- You can click the Recommended Charts button to view a gallery of pre-configured graphs that best match the selected data.

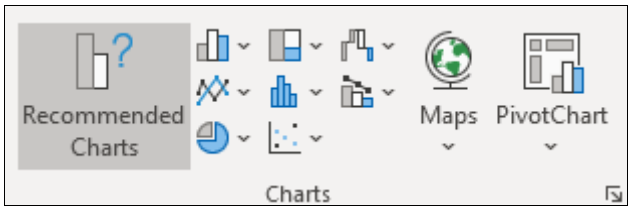


Figure 6.3.1 Insert Charts

Creating a Chart is easy.

CASE STUDY: Your principal CA Chandiwala has given you data relating to Unit wise Expenditure and Budget. It is suspected that in some Units expenditure is much beyond the Budget allocated, He wants to investigate those units. You have to make a presentation on Expenditure vs Budget so that defaulting units could be investigated.

	A	B	C
1	**Unit**	**Expenditure**	**Budget**
2	Brake Unit	10,85,191	12,00,000
3	Chassis Unit	14,19,365	12,00,000
4	Dashboard Unit	16,19,833	17,00,000
5	Electrical Unit	31,12,034	33,00,000
6	Engine Unit	32,99,367	35,00,000
7	Headlight Unit	42,72,218	44,00,000
8	Paintshop Unit	27,33,346	27,00,000
9	Seats Unit	47,47,303	43,00,000
10	Steering Unit	16,68,663	18,00,000

Figure 6.3.2 Unit wise Expenditure

Strategy:

You can achieve your objective with a Chart, which will help you visualise the expenses vis-à-vis Budget.

1. You can go for a plain vanilla Column chart
2. Select any cell in the range, go to **Insert > Chart.**
3. In the recommended chart, Excel shows a clustered Column chart as shown in Figure

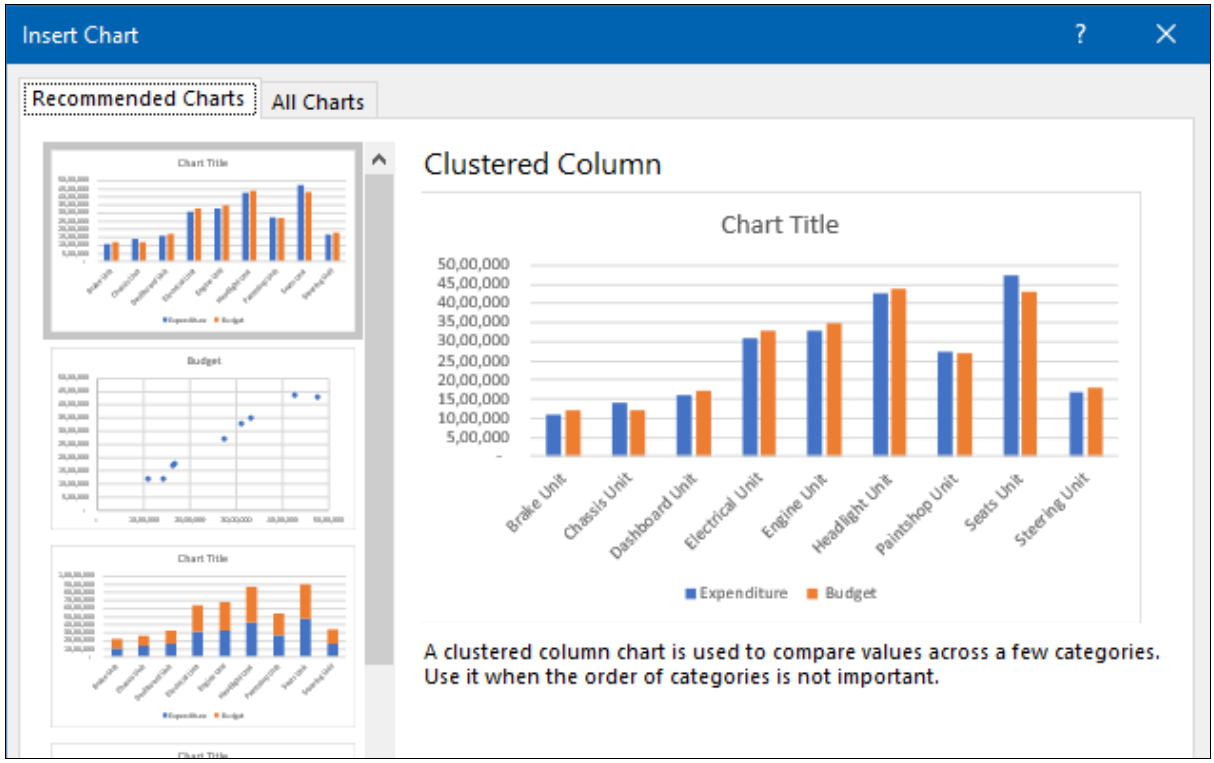
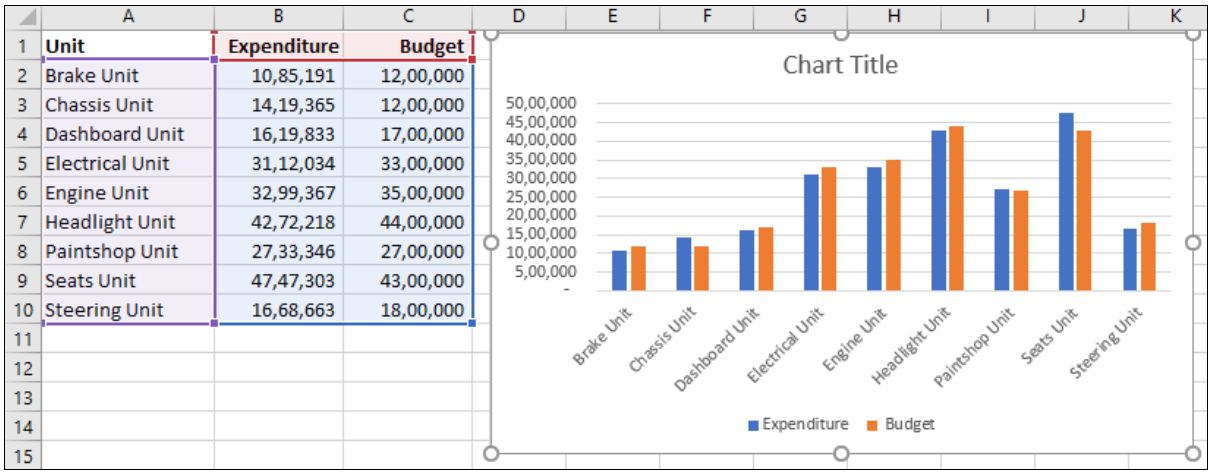


Figure 6.3.3 Insert Recommended Chart

4. Click on Ok.

5. A chart will be added as shown.



	Unit	Expenditure	Budget
2	Brake Unit	10,85,191	12,00,000
3	Chassis Unit	14,19,365	12,00,000
4	Dashboard Unit	16,19,833	17,00,000
5	Electrical Unit	31,12,034	33,00,000
6	Engine Unit	32,99,367	35,00,000
7	Headlight Unit	42,72,218	44,00,000
8	Paintshop Unit	27,33,346	27,00,000
9	Seats Unit	47,47,303	43,00,000
10	Steering Unit	16,68,663	18,00,000

Figure 6.3.4 Clustered Column Chart

6. Compared to Numbers with this visualisation it is easy to interpret data.

7. But still, it is difficult to comprehend.

8. For Data analysis it becomes easy, if each of the data points can be represented by a different chart, for this you can use Combination Charts.

9. You can use **Column Chart** for Budget Figures and Expenditure with **Line chart**, that way if the **Line is not intersecting Column, it is over budget**.

10. Go to **Chart Design Tab > Change Chart type**.

11. Change Chart Type Dialog Box appears.

12. Click on **Combo** as shown in Figure

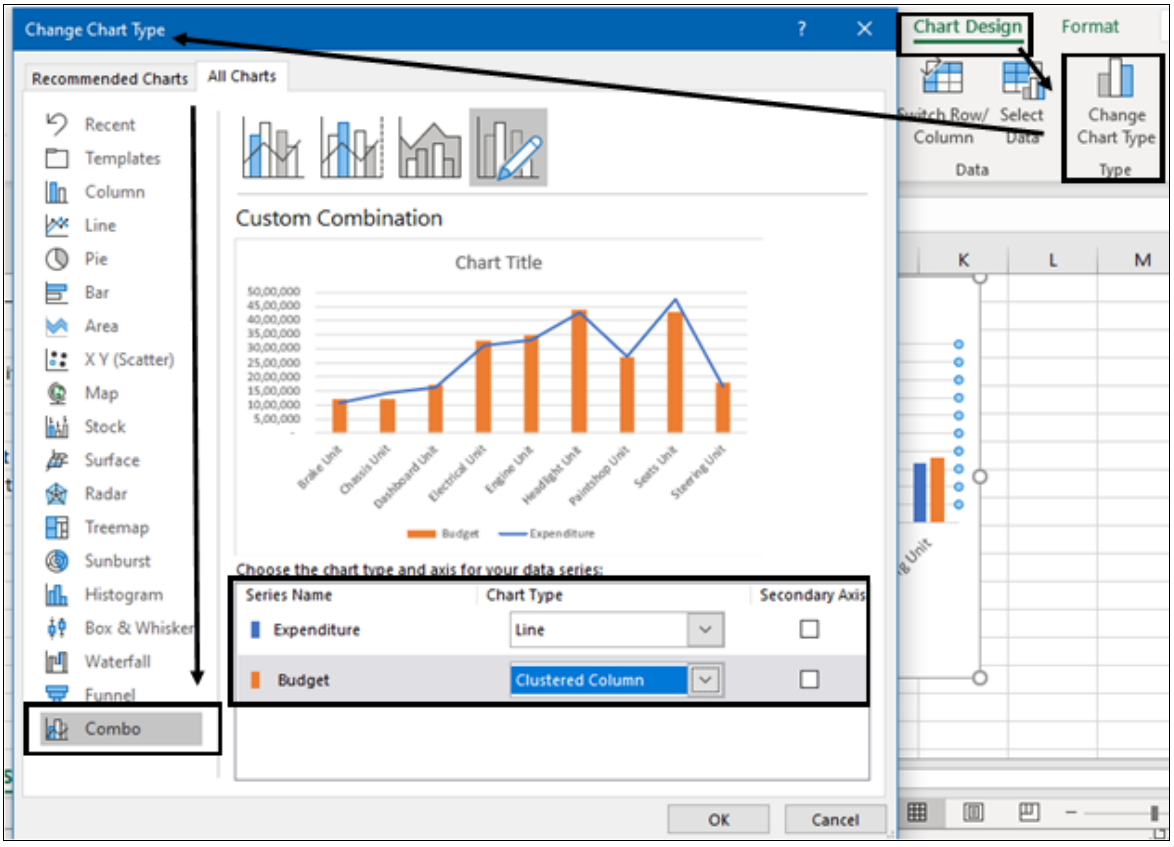
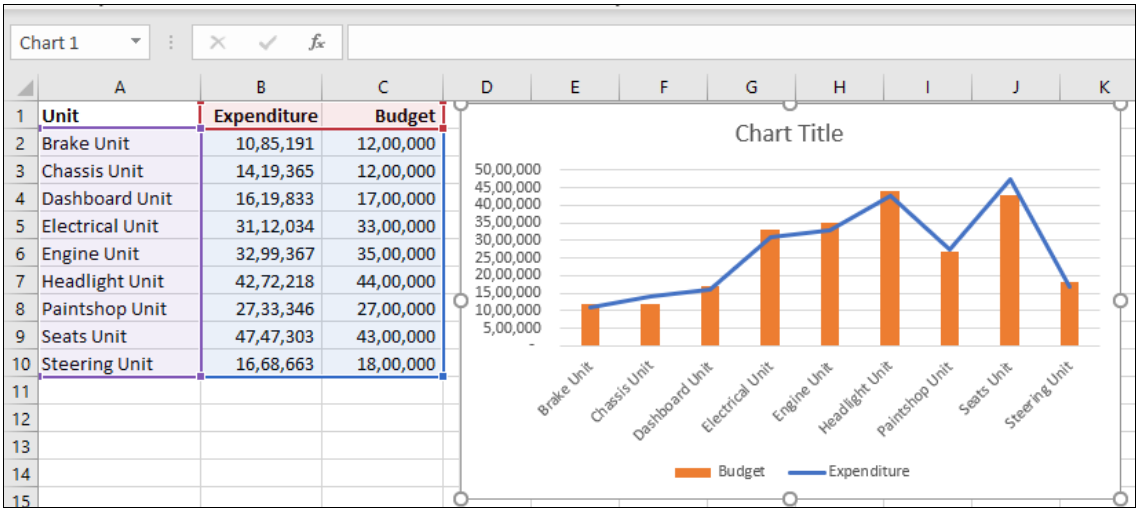


Figure 6.3.5 Chart Design-Change Chart Type-Combo

13. Change the Chart Type for the series Expenditure to Line with Markers. The preview appears under Custom Combination.

14. Click OK.



	A	B	C
1	Unit	Expenditure	Budget
2	Brake Unit	10,85,191	12,00,000
3	Chassis Unit	14,19,365	12,00,000
4	Dashboard Unit	16,19,833	17,00,000
5	Electrical Unit	31,12,034	33,00,000
6	Engine Unit	32,99,367	35,00,000
7	Headlight Unit	42,72,218	44,00,000
8	Paintshop Unit	27,33,346	27,00,000
9	Seats Unit	47,47,303	43,00,000
10	Steering Unit	16,68,663	18,00,000

Figure 6.3.6 Combo of Column and Line

15. You will observe in the chart, the Budget values are in Columns and the Expenditure values are marked along the line.
16. The data visualization has become better as it also shows you the trend of your results.
17. Wherever the line is not intersecting column Expenditure has exceeded the budget and needs to be investigated.

Gist: You have created a Combo chart to analyse data for Expenditure vis-à-vis Budgets

Commands learnt: **Insert >Charts**

6.4 PIVOT TABLES & CHARTS

Excel pivot tables are a very useful and powerful feature of MS Excel. They can be used to summarize, analyze, explore, and present your data. Source data could be:

- An Excel worksheet database/list or any range that has labeled columns
- A collection of ranges to be consolidated. The ranges must contain both labeled rows and columns.
- A database file was created in an external application.

The data in a PivotTable cannot be changed as it is a summarized view of other data. Any change if needed has to be done in source data.

You often use a PivotTable report when you want to analyze related totals, especially when you have a long list of figures to sum and you want to compare several facts about each figure.

Here are some example uses of pivot tables:

- Summarizing data like finding the average sales for each region for each product from a product sales data table.

- Listing unique values in any column of a table
- Creating a pivot report with sub-totals and custom formats
- Making a dynamic pivot chart
- Filtering, sorting, drilling-down data in the reports without writing one formula or macro.
- Transposing data – i.e. moving rows to columns or columns to rows. [learn more]
- Linking data sources outside excel and be able to make pivot reports out of such data.

A PivotTable report is an interactive way to quickly summarize large amounts of data. You can use a PivotTable report to analyze numerical data in detail and to answer unanticipated questions about your data. A PivotTable report is specially designed for:

- Querying large amounts of data in many user-friendly ways.
- Subtotaling and aggregating numeric data, summarizing data by categories and subcategories and creating custom calculations and formulas.
- Expanding and collapsing levels of data to focus your results and drilling down to details from the summary data for areas of interest to you.
- Moving rows to columns or columns to rows (or "pivoting") to see different summaries of the source data.
- Filtering, sorting, grouping, and conditionally formatting the most useful and interesting subset of data to enable you to focus on the information that you want.
- Presenting concise, attractive, and annotated online or printed reports.

6.4.1 CREATING PIVOT TABLES

Some prerequisites when creating a PivotTable

- The Excel data to be used as the source for a pivot table must be organized in rows and columns,
- Each column in the source data must contain a heading. If you try to create a pivot table from data that has blank heading cells, you see an error message.
- Each column in the source data should contain one type of data.
- Don't create multiple columns to store the same type of information.
- Each row in the source data should contain the details for one record
- The source data should not have any blank rows within it and cannot include any completely blank columns.
- You set the data up as Table.

To create a PivotTable you select a cell in an Excel table or the entire Excel table. As you start creating a pivot table, that Excel table is shown as the default source range for the pivot table. If you want to use a different table or range, you can type an Excel table name or range address in the Table/Range box.

You can create a Pivot in the same worksheet or an altogether different Worksheet.

You see the Column Headings in the upper area of the PivotTable Field Pane and at the bottom of the PivotTable Field List pane is the four areas of the pivot table:

- Report Filter,
- Column Labels,
- Row Labels, and
- Values.

You can simply drag & drop the fields into these areas, and they'll appear in the matching area of the pivot table layout on the worksheet.

CASE STUDY: As a part of the investigation your principal CA Chandiwala has given you data relating to power consumption in different units of a company for various dates as shown in Fig 6.4.1. He wants you to analyse the power consumption on Unit wise & Week day wise.

	A	B	C	D
1	**Date**	**Unit**	**Power**	**Day**
2	01-01-2019	Brake	0.872	Tue
3	01-01-2019	Chassis	0.656	Tue
4	01-01-2019	Dashboard	2.354	Tue
5	01-01-2019	Electrical	0.396	Tue
6	01-01-2019	Engine	2.304	Tue
7	01-01-2019	Headlight	0.830	Tue
8	01-01-2019	Paintshop	0.388	Tue
9	01-01-2019	Seats	4.248	Tue
10	01-01-2019	Steering	0.340	Tue
11	01-01-2019	Suspension	1.404	Tue
12	01-01-2019	Transmission	0.800	Tue

Figure 6.4.1 Datewise data

Strategy:

You can achieve your objective with a pivot table; you can create summary tables just by choosing the columns you want to compare.

And once you've built your summary, you don't need to stick with it—instead, you can Spin the Data ie pivot it to have a Multi Dimensional View of data

You can even drill-down into the details, apply filter criteria, and apply advanced subtotaling calculations, etc.

1. Before creating a PivotTable, you must ensure that you have headings in the first row of your data in this case headers are already there. You will **select the data** range from which you want to make the pivot table.

2. Go to **Insert** ribbon and click on **Pivot table** option as shown in Fig 6.4.2

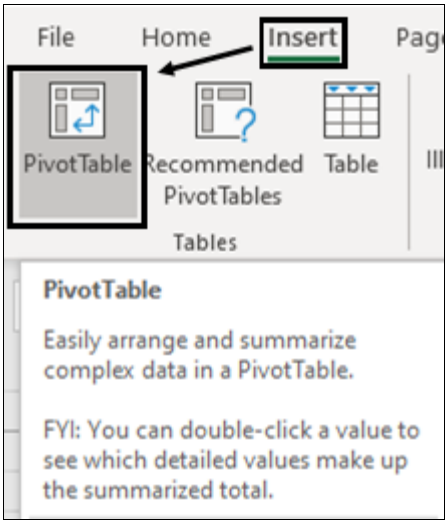


Figure 6.4.2 Pivot table on Insert Ribbon

3. Select the **target cell** where you want to place the pivot table if you want pivot table in the same worksheet but you could also select **pivot table to be placed in a new worksheet** by choosing the "New worksheet" option in **Create PivotTable** Dialog Box shown in Fig6.4.3

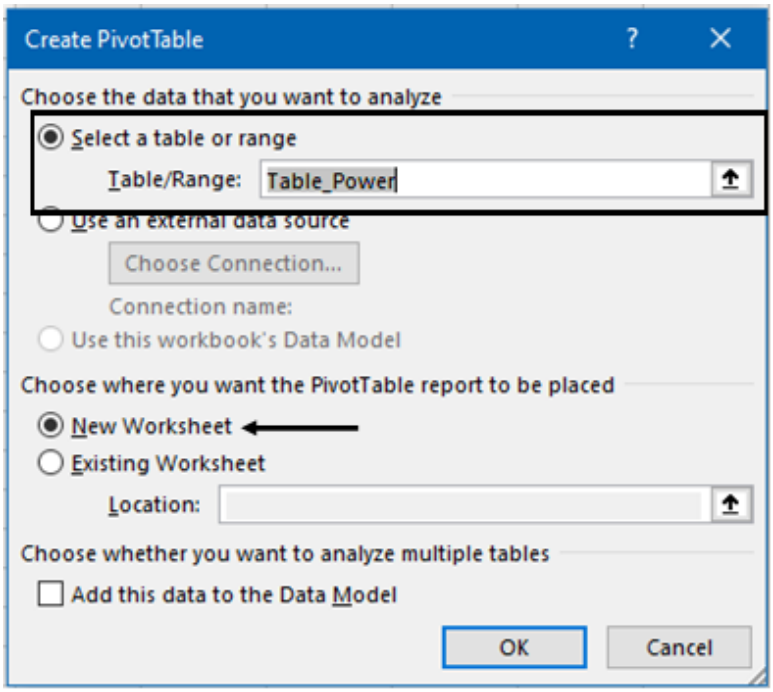


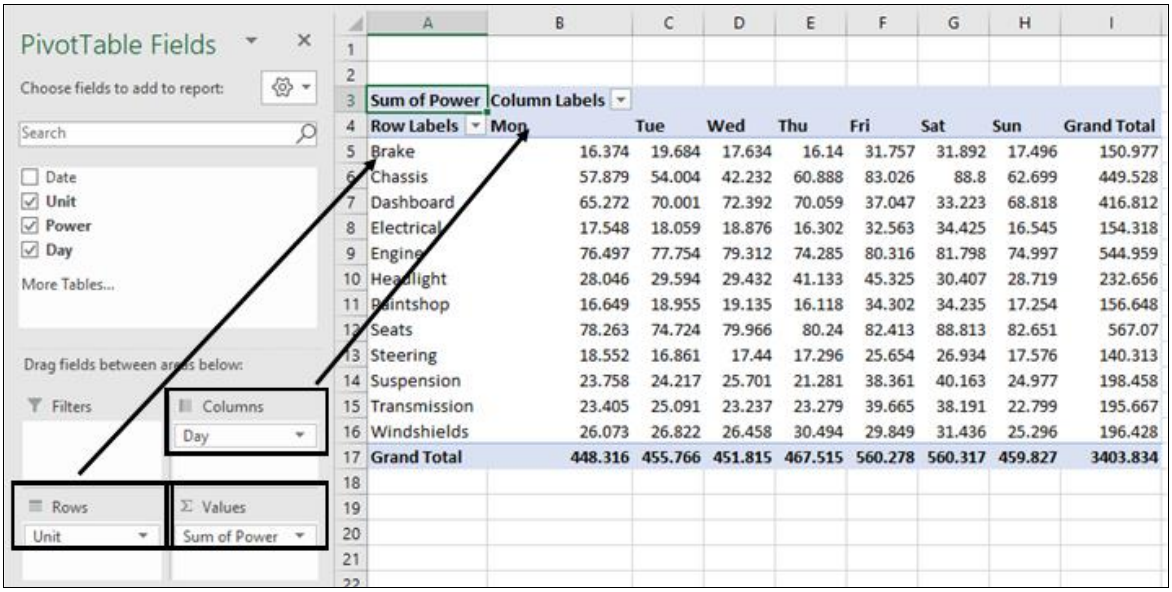
Figure 6.4.3 Create PivotTable Dialog Box

4. The **pivot report User Interface** is very intuitive and sandbox-like. To make a powerful analysis, all you have to do is **drag and drop fields** into the pivot table grid area.

Fill in the PivotTable Field List by dragging PivotTable headings, or fields, into the boxes, or zones. These zones are:

- **Row Labels:** Fields dragged here are listed on the left side of the table in the order in which they are added to the box, *In your case, you select "Unit"*
- **Column Labels:** Fields dragged here have their values listed across the top row of the PivotTable, *in your case you select "Day"*
- **Values Fields:** dragged here are summarized mathematically in the table, *In your case, you select "Power"*
- **Report Filter:** These fields use filtering to limit which data the pivot table displays**.** For fields dragged to the Report Filter area, you can easily pick any subset of the field values so that the PivotTable shows calculations based only on that subset.

5. You can also control this by using the "**Pivot table Field List**". What you see is as shown in Fig.6.4.4



Sum of Power	Column Labels							
Row Labels	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Grand Total
Brake	16.374	19.684	17.634	16.14	31.757	31.892	17.496	150.977
Chassis	57.879	54.004	42.232	60.888	83.026	88.8	62.699	449.528
Dashboard	65.272	70.001	72.392	70.059	37.047	33.223	68.818	416.812
Electrical	17.548	18.059	18.876	16.302	32.563	34.425	16.545	154.318
Engine	76.497	77.754	79.312	74.285	80.316	81.798	74.997	544.959
Headlight	28.046	29.594	29.432	41.133	45.325	30.407	28.719	232.656
Paintshop	16.649	18.955	19.135	16.118	34.302	34.235	17.254	156.648
Seats	78.263	74.724	79.966	80.24	82.413	88.813	82.651	567.07
Steering	18.552	16.861	17.44	17.296	25.654	26.934	17.576	140.313
Suspension	23.758	24.217	25.701	21.281	38.361	40.163	24.977	198.458
Transmission	23.405	25.091	23.237	23.279	39.665	38.191	22.799	195.667
Windshields	26.073	26.822	26.458	30.494	29.849	31.436	25.296	196.428
Grand Total	448.316	455.766	451.815	467.515	560.278	560.317	459.827	3403.834

Figure 6.4.4 Pivot Table created showing Unitwise Daywise Power consumption

Excel generates the pivot table, updating it dynamically as you add, rearrange, or remove columns.

- You can **dock** the PivotTable Field List to either side of the Excel window and horizontally **resize** it.
- You can also **undock** the PivotTable Field List and can resize it both vertically and horizontally.
- To see the PivotTable Field List, make sure that you **click anywhere in the PivotTable**.
- After closing the PivotTable Field List, you can display it again by **Right-click on the PivotTable**, and then click **Show Field List**. You can also click Field List on the Ribbon.as shown in Fig 6.4.5

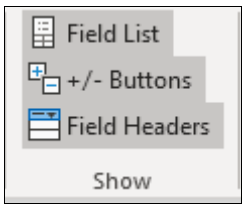


Figure 6.4.5 Show field list

Gist: You have created a PivotTable out of power consumption Data

Commands learnt: **Insert > PivotTables**

Food for thought: It is immaterial whether you group by column or rows, but what is sensible is to make the data readable. It makes more sense to make the field with Long names as rows which improves readability.

SWAPPING ROWS AND COLUMNS

PIVOT TABLES ARE CALLED PIVOT TABLES BECAUSE YOU CAN EASILY PIVOT FIELDS FROM A ROW TO A COLUMN AND VICE VERSA TO CREATE A DIFFERENT LAYOUT.

One key feature of PivotTables is flexibility. You can move fields, recalculate, summarise in different ways, innumerable times.

6.4.2 GROUPING FIELDS- CREATING MONTHLY REPORTS

Pivot gives you the facility to **Group & subgroup** fields up to any Level.by adding to Rows or Columns.

Further, you can **hide or show** different groups. This feature lets you see detailed information for just the part of the table that you're interested in while hiding the rest.

CASE STUDY: In the previous Case study, after you submit your results to your principal CA Chandiwala showing Unitwise and Daywise Power consumption, your principal wants to know the monthlyunitwise power consumption.

Strategy:

In a PivotTable, you can Drag "Unit", "Day" to **Rows** in that order as shown in Fig.6.4.6. to group the data by "Unit" and then "Day"

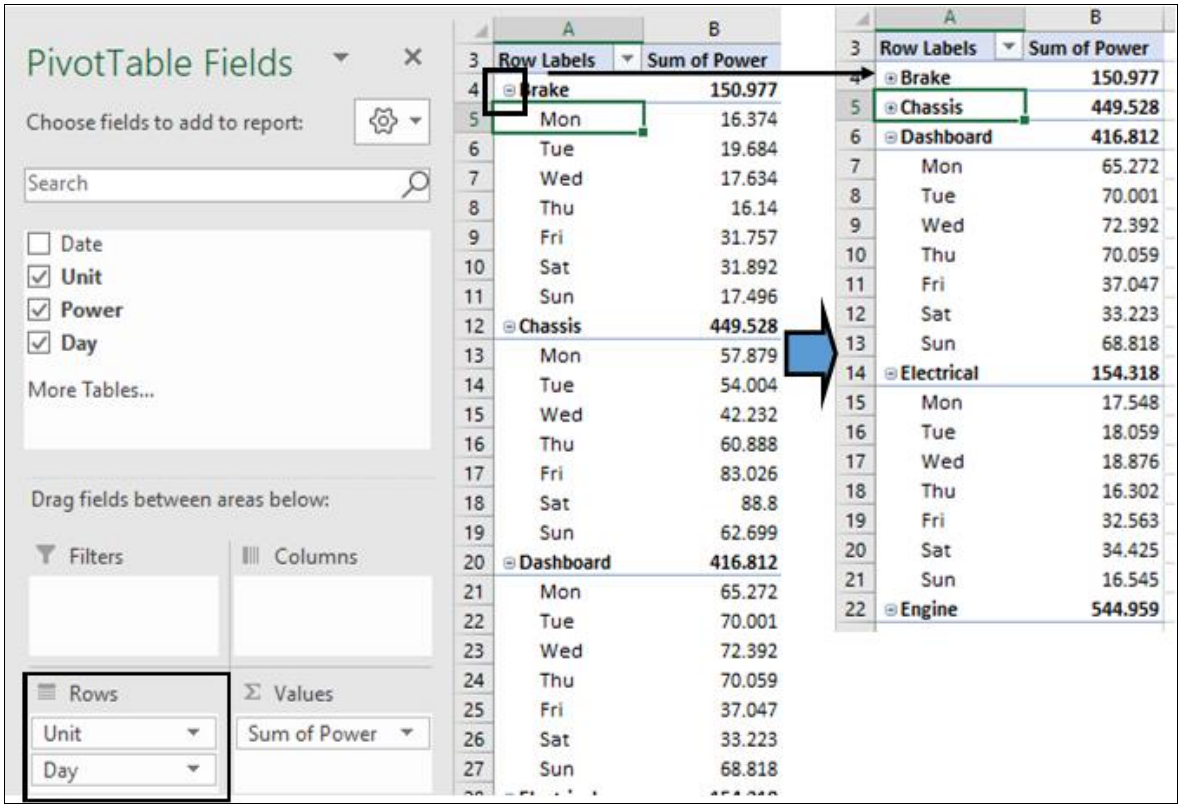


Figure 6.4.6 Grouping

You can **Expand/ Collapse** grouped Fields since the ability to expand and collapse (hide) fields is a great advantage in PivotTables. In Fig.6.4.6you see minus (-) signs by each "Unit", group. **Clicking the minus sign** collapses a field and changes the sign to a plus (+) sign. Clicking the plus sign expands the field as shown in Fig.6.4.6.

By clicking the (-) sign by in any cell in column A, the Unitwise, Day wise sales are contracted to one row and months are hidden, by clicking + sign detailed view comes back.

Alternate strategy;

You **click** on any cell in the **Row label** column which you want to expand or collapse

To **Expand** all your categories: Select **PivotTable tools** from Ribbon go to **Options> Active field> Expand entire Field**.

To **Collapse** all your categories: Select **PivotTable tools** from Ribbon go to **Options> Active field> Collapse entire Field**.

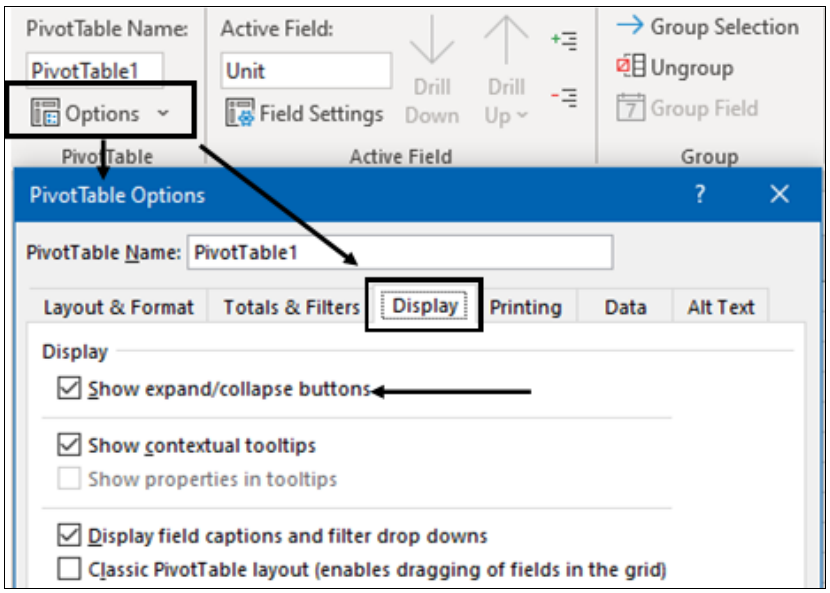
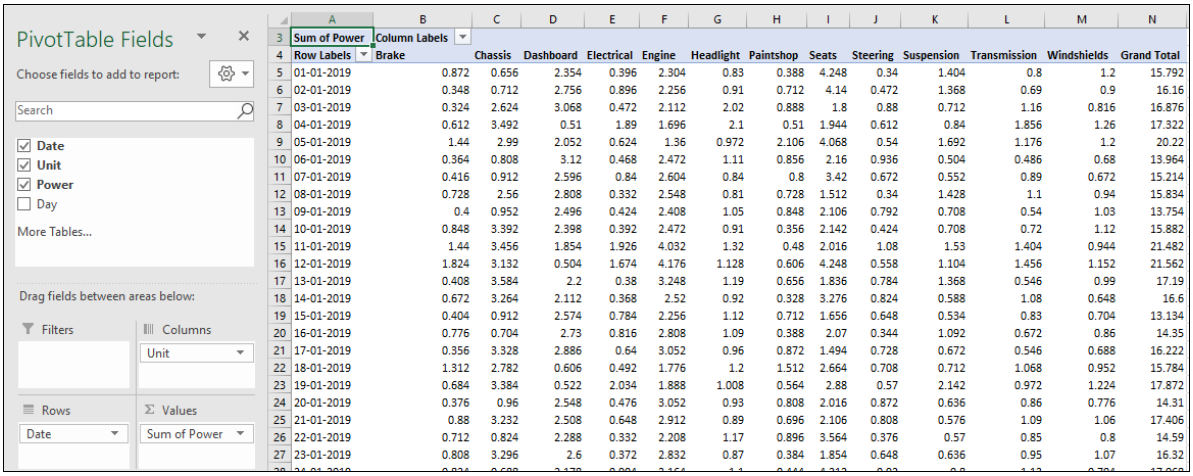


Figure 6.4.7 PivotTable Options

By simply dragging "Unit " to **column**& "Date" to rows Layout is changed as follows



Sum of Power	Column Labels												
Row Labels	Brake	Chassis	Dashboard	Electrical	Engine	Headlight	Paintshop	Seats	Steering	Suspension	Transmission	Windshields	Grand Total
01-01-2019	0.872	0.656	2.354	0.396	2.304	0.83	0.388	4.248	0.34	1.404	0.8	1.2	15.792
02-01-2019	0.348	0.712	2.756	0.896	2.256	0.91	0.712	4.14	0.472	1.368	0.69	0.9	16.16
03-01-2019	0.324	2.624	3.068	0.472	2.112	2.02	0.888	1.8	0.88	0.712	1.16	0.816	16.876
04-01-2019	0.612	3.492	0.51	1.89	1.696	2.1	0.51	1.944	0.612	0.84	1.856	1.26	17.322
05-01-2019	1.44	2.99	2.052	0.624	1.36	0.972	2.106	4.068	0.54	1.692	1.176	1.2	20.22
06-01-2019	0.364	0.808	3.12	0.468	2.472	1.11	0.856	2.16	0.936	0.504	0.486	0.68	13.964
07-01-2019	0.416	0.912	2.596	0.84	2.604	0.84	0.8	3.42	0.672	0.552	0.89	0.672	15.214
08-01-2019	0.728	2.56	2.808	0.332	2.548	0.81	0.728	1.512	0.34	1.428	1.1	0.94	15.834
09-01-2019	0.4	0.952	2.496	0.424	2.408	1.05	0.848	2.106	0.792	0.708	0.54	1.03	13.754
10-01-2019	0.848	3.392	2.398	0.392	2.472	0.91	0.356	2.142	0.424	0.708	0.72	1.12	15.882
11-01-2019	1.44	3.456	1.854	1.926	4.032	1.32	0.48	2.016	1.08	1.53	1.404	0.944	21.482
12-01-2019	1.824	3.132	0.504	1.674	4.176	1.128	0.606	4.248	0.558	1.104	1.456	1.152	21.562
13-01-2019	0.408	3.584	2.2	0.38	3.248	1.19	0.656	1.836	0.784	1.368	0.546	0.99	17.19
14-01-2019	0.672	3.264	2.112	0.368	2.52	0.92	0.328	3.276	0.824	0.588	1.08	0.648	16.6
15-01-2019	0.404	0.912	2.574	0.784	2.256	1.12	0.712	1.656	0.648	0.534	0.83	0.704	13.134
16-01-2019	0.776	0.704	2.73	0.816	2.808	1.09	0.388	2.07	0.344	1.092	0.672	0.86	14.35
17-01-2019	0.356	3.328	2.886	0.64	3.052	0.96	0.872	1.494	0.728	0.672	0.546	0.688	16.222
18-01-2019	1.312	2.782	0.606	0.492	1.776	1.2	1.512	2.664	0.708	0.712	1.068	0.952	15.784
19-01-2019	0.684	3.384	0.522	2.034	1.888	1.008	0.564	2.88	0.57	2.142	0.972	1.224	17.872
20-01-2019	0.376	0.96	2.548	0.476	3.052	0.93	0.808	2.016	0.872	0.636	0.86	0.776	14.31
21-01-2019	0.88	3.232	2.508	0.648	2.912	0.89	0.696	2.106	0.808	0.576	1.09	1.06	17.406
22-01-2019	0.712	0.824	2.288	0.332	2.208	1.17	0.896	3.564	0.376	0.57	0.85	0.8	14.59
23-01-2019	0.808	3.296	2.6	0.372	2.832	0.87	0.384	1.854	0.648	0.636	0.95	1.07	16.32

Figure 6.4.8 Datewise, Unitwise power consumption

You have just swapped Unit in rows to Columns

PivotTables gives you Excellent Data summarization & analytical capability

You can group & subgroup fields in a PivotTable.

1. You can **group headings** in a PivotTable.
2. In the above Case Study let's say you want to group *Dates to months* to have a monthly grouping of Power consumption.
   a. **Select** the column to want to group,

b. **Right-click** the selection, and then go to **Group**

c. **Grouping** Dialog Box will appear as shown in Fig 6.4.9

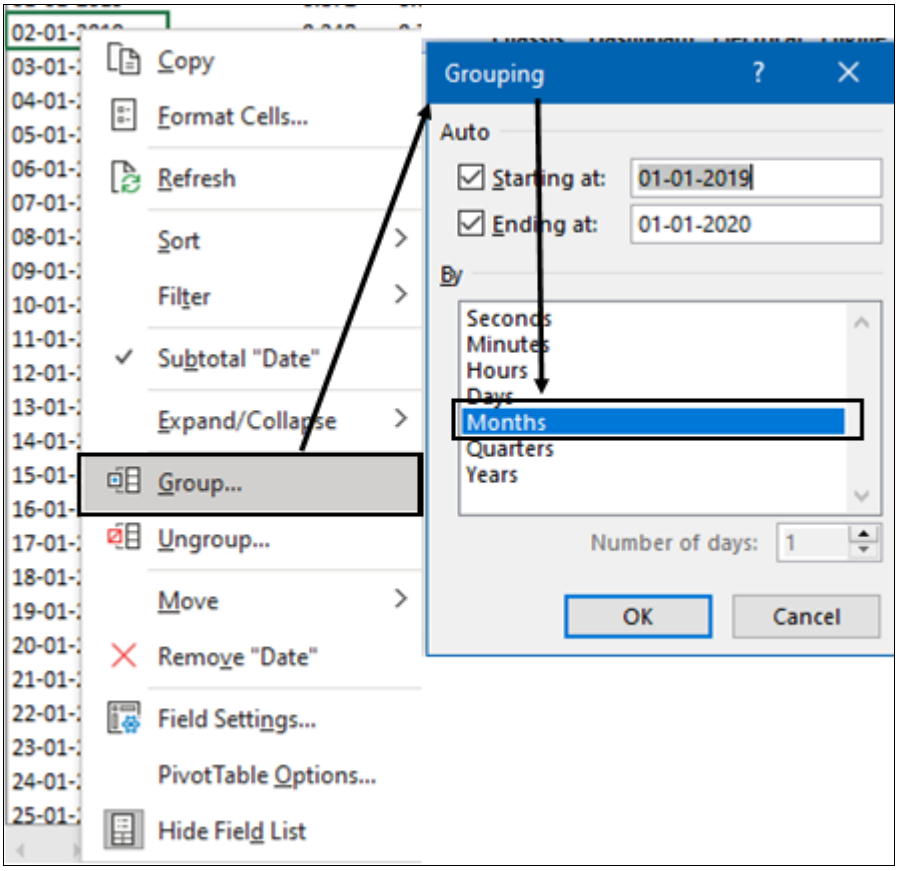


Figure 6.4.9 Grouping Options

3. Alternatively, you can go to **PivotTable Analyse>Group>Group selection** as shown in Fig.6.4.10

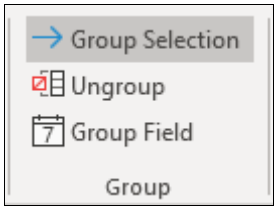


Figure 6.4.10 Grouping in Ribbon

4. In the **Grouping** dialog box, select one or more options from the '**By**' list.

5. Select **Months**, You have the option of selecting Days, Quarter, Years too.

6. To limit the dates that are grouped, you can set a **Start and End date**, by typing the dates in the 'Startinat' and 'Ending at' boxes

7. Click **OK** to close the dialog box and you have a neat monthly report.

Sum of Power	Column Labels												
Row Labels	Brake	Chassis	Dashboard	Electrical	Engine	Headlight	Paintshop	Seats	Steering	Suspension	Transmission	Windshields	Grand Total
Jan	21.604	62.608	65.032	25.028	82.728	34.228	23.384	78.666	20.394	27.726	30.706	29.454	501.558
Feb	13.726	38.625	34.651	13.21	46.842	19.59	13.683	43.957	12.931	16.505	14.562	16.535	284.817
Mar	10.405	30.967	24.813	9.555	31.868	13.573	10.066	33.063	8.459	13.048	12.068	12.237	210.122
Apr	3.478	10.139	9.054	3.612	11.651	5.152	3.977	14.366	3.186	3.969	4.548	4.461	77.593
May	8.526	24.083	24.903	10.073	32.214	13.54	9.241	36.559	8.689	12.602	12.967	11.832	205.229
Jun	18.571	58.208	56.754	21.134	75.566	31.136	23.36	74.541	19	28.056	25.883	25.862	458.071
Jul	26.843	84.242	78.963	27.46	102.537	42.838	25.856	115.771	25.001	34.267	35.852	36.259	635.889
Aug	14.744	42.049	36.468	14.082	49.573	21.675	12.86	50.048	11.57	19.508	18.154	18.335	309.066
Sep	7.674	23.152	19.808	7.38	27.532	11.905	8.263	28.574	8.187	9.596	9.692	10.479	172.242
Oct	7.285	19.092	18.396	7.157	22.981	10.683	7.857	24.549	6.445	10.637	8.906	8.858	152.846
Nov	5.169	18.237	14.637	4.131	17.935	7.93	5.285	19.47	5.093	6.635	6.837	6.23	117.589
Dec	12.952	38.126	33.333	11.496	43.532	20.406	12.816	47.506	11.358	15.909	15.492	15.886	278.812
Grand Total	**150.977**	**449.528**	**416.812**	**154.318**	**544.959**	**232.656**	**156.648**	**567.07**	**140.313**	**198.458**	**195.667**	**196.428**	**3403.834**

Figure 6.4.11 Monthwise grouping of dates

Gist: You have learnt to Group by Dates in PivotTables

Commands learnt: PivotTables- Grouping Dialog Box

Food for thought Grouping is available for time or numbers also.

DRILL –DOWN PIVOTS

Whenever you want to see the values behind a pivot field just double click on it.

If you click any value in a pivot table, Excel adds a new worksheet. In this worksheet, it copies the records that were used to calculate the pivot table cell in question, and nothing else.

CASE STUDY: Your principal CA Chandiwala, on receiving the report in the previous case study sees that consumption of power for April for Brake is less compared to other months & Units, He wants you to find details of the underlying data so that he can look for the reason if any.

Strategy:

You want to get the data behind power consumption in April relating to Unit "Brakes",

When you double click on cell B8 ie Power Consumption in Apr of "Brake", you get the data relating to Brake in a new worksheet as shown in Fig.6.4.12.

	A	B	C	D
1	Date	Unit	Power	Day
2	30-04-2019	Brake	0.11	Tue
3	29-04-2019	Brake	0.132	Mon
4	28-04-2019	Brake	0.058	Sun
5	27-04-2019	Brake	0.228	Sat
6	26-04-2019	Brake	0.221	Fri
7	25-04-2019	Brake	0.049	Thu
8	24-04-2019	Brake	0.108	Wed
9	23-04-2019	Brake	0.05	Tue
10	22-04-2019	Brake	0.071	Mon
11	21-04-2019	Brake	0.126	Sun
12	20-04-2019	Brake	0.197	Sat
13	19-04-2019	Brake	0.278	Fri
14	18-04-2019	Brake	0.104	Thu
15	17-04-2019	Brake	0.142	Wed
16	16-04-2019	Brake	0.096	Tue
17	15-04-2019	Brake	0.057	Mon

Figure 6.4.12 Detailed power consumption

Gist: You have learnt to Drill Down in PivotTables

Commands learnt: PivotTables- Drill down

LAYOUT AND FORMAT

Formatting a PivotTable is a breeze and as easy as formatting any list or table in Excel. In order to format a PivotTable, you simply need to click on any cell in the Pivot table and go to **PivotTable Design,** and Ribbon with lots of layout and formatting options is displayed as shown in Fig.6.4.13.

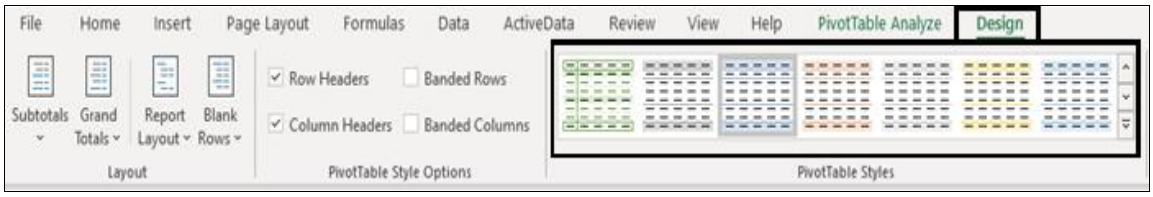


Figure 6.4.13 PivotTable Design

The Design tab on the PivotTable Tools contextual tab is divided into three groups:

- **Layout group** that enables you to add subtotals and grand totals to the pivot table and modify its basic layout
- **PivotTable Style Options** group that enables you to refine the pivot table style you select for the table using the PivotTable Styles gallery to the immediate right
- **PivotTable Styles** group that contains the gallery of styles you can apply to the active pivot table by clicking the desired style thumbnail.

6.4.3 PIVOTCHARTS

Excel makes it easy to visually summarize PivotTables by using Pivot Charts.

A pivot chart paints a picture of the data in a pivot table and can make the data easier to understand. Instead of a table full of numbers, a pivot chart can use columns, bars, or pie charts to illustrate the numbers

Making a pivot chart from a pivot table is very simple. Just click on the pivot chart icon from the **PivotTable Analyse** ribbon area and follow the wizard.

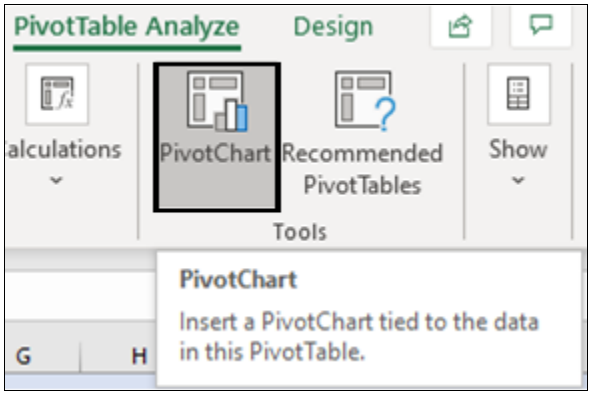


Figure 6.4.14 Pivot Charts

You get an **Insert Chart** Dialog box with various options of different Chart styles, you can select any and click OK

A **Pivot Chart** is inserted

CASE STUDY: In the previous case study CA. Chandiwala wants to gain a visual Summary of the data.

Strategy:

You can use the Pivot Chart Features of PivotTables.

1. Go to **PivotTable Analyse Tools> Pivot Chart** and a charting dialog box open as below from which you can select the type of chart you want.

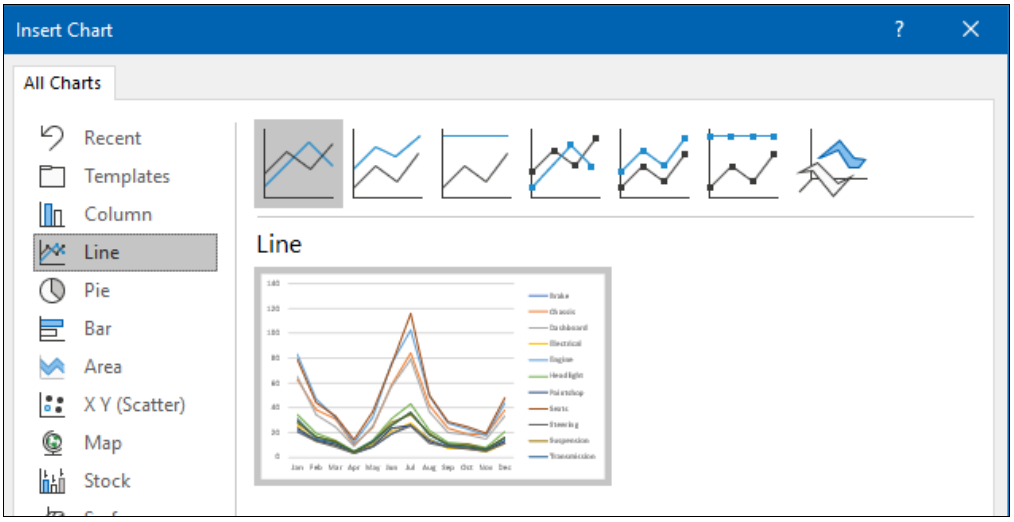


Figure 6.4.15 Insert Chart Dialog Box

2. In this case, you select a line chart and a chart appears as shown in Fig.6.4.16:

Sum of Power	Column Labels												
Row Labels	Brake	Chassis	Dashboard	Electrical	Engine	Headlight	Paintshop	Seats	Steering	Suspension	Transmission	Windshields	Grand Total
Jan	21.604	62.608	65.032	25.028	82.728	34.228	23.384	78.666	20.394	27.726	30.706	29.454	501.558
Feb	13.726	38.625	34.651	13.21	46.842	19.59	13.683	43.957	12.931	16.505	14.562	16.535	284.817
Mar	10.405	30.967	24.813	9.555	31.868	13.573	10.066	33.063	8.459	13.048	12.068	12.237	210.122
Apr	3.478	10.139	9.054	3.612	11.651	5.152	3.977	14.366	3.186	3.969	4.548	4.461	77.593
May	8.526	24.083	24.903	10.073	32.214	13.54	9.241	36.559	8.689	12.602	12.967	11.832	205.229
Jun	18.571	58.208	56.754	21.134	75.566	31.136	23.36	74.541	19	28.056	25.883	25.862	458.071
Jul	26.843	84.242	78.963	27.46	102.537	42.838	25.856	115.771	25.001	34.267	35.852	36.259	635.889
Aug	14.744	42.049	36.468	14.082	49.573	21.675	12.86	50.048	11.57	19.508	18.154	18.335	309.066
Sep	7.674	23.152	19.808	7.38	27.532	11.905	8.263	28.574	8.187	9.596	9.692	10.479	172.242
Oct	7.285	19.092	18.396	7.157	22.981	10.683	7.857	24.549	6.445	10.637	8.906	8.858	152.846
Nov	5.169	18.237	14.637	4.131	17.935	7.93	5.285	19.47	5.093	6.635	6.837	6.23	117.589
Dec	12.952	38.126	33.333	11.496	43.532	20.406	12.816	47.506	11.358	15.909	15.492	15.886	278.812
Grand Total	150.977	449.528	416.812	154.318	544.959	232.656	156.648	567.07	140.313	198.458	195.667	196.428	3403.834

Figure 6.4.16 Chart Inserted

The command buttons on the Power Chart Analyse, Design, and Format tabs attached to the Pivot Chart contextual tab make it easy to further format and customize your pivot chart.

- You can use the **Pivot Chart Analyse tab** analysis of the PivotTable.
- You can use the **Design tab** buttons to select a new **chart style** for your pivot chart or even a brand new chart type.

- You can use the **Format tab** buttons to **refine the look of any graphics,** you have added to the chart as well as select a new background color for your chart.

  Through these, you can manage the various aspects of Pivot Chart

  Gist: You have Created PivotChart based on Sales in Pivot Tables

  Commands learnt: PivotTables- Pivot Chart

  Food for Thought: If you close the PivotTable Field List, you can display it again. Right-click the PivotTable, and then click Show Field List.

  Food for Thought: If you don't see the fields that you want to use in the PivotTable Field List, refresh the PivotTable or PivotChart report to display any new fields.

6.4.4 MODIFYING PIVOT DATA

On the **Analyze** tab, in the Data group, click **Change Data Source**, and then click Change Data Source.

The **Change PivotTable Data Source** dialog box is displayed.

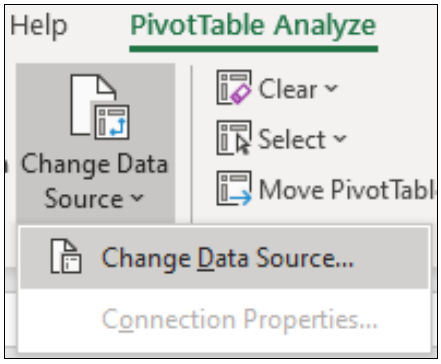


Figure 6.4.17 Change data source

To change the data source of a PivotTable to a different Excel table or a cell range, click Select a table or range, and then enter the first cell in the Table/Range text box, and click OK

6.4.5 FILTERING, SORTING, AND MULTIPLE REPORTS

When you create a new pivot table, you notice that Excel automatically adds drop-down buttons to the Report Filter field as well as the labels for the column and row fields. These drop-down buttons, called filter buttons, enable you to filter all but certain entries in any of these fields, and in the case of the column and row fields, to sort their entries in the table.

By sorting and filtering the data in a pivot table, you can focus on areas that need the most attention. With these pivot table tools, you can go from the big picture, comparing all aspects of your business, to the small details, homing in on problems for closer analysis.

To show details, add a field to the Row Labels or Column Labels area. To limit the data, add a field to the Report Filter area, and filter the data.

When you add a field to the Column Labels or Row Labels area of the pivot table, the labels are usually sorted alphabetically in ascending order. For some fields, you may prefer the labels in descending order. You can sort the labels in your pivot table to see labels in descending or ascending order depending on your selection.

Sorting could be from:

- Smallest to Largest
- Largest to Smallest

You can use conditional formatting to colour cells in a pivot table and add data bars to illustrate the amounts. You can also add icons, such as red, yellow, and green traffic lights to indicate progress or decline or to indicate good or poor results. With conditional formatting, you can color only the values that fall within a specific date range or those that are above or below a certain level.

Filtering data

When you create a Pivot table filter buttons are added to the Report Filter field(s). By selecting a particular option on the dropdown lists attached to one of these filter buttons, only the summary data for that subset displays in the PivotTable.

For Example in the pivot table created in the case study, you can filter the data based on months by clicking the months filter as shown in Fig.6.4.20 If you select “Feb” than only the data relating to Feb would be displayed.

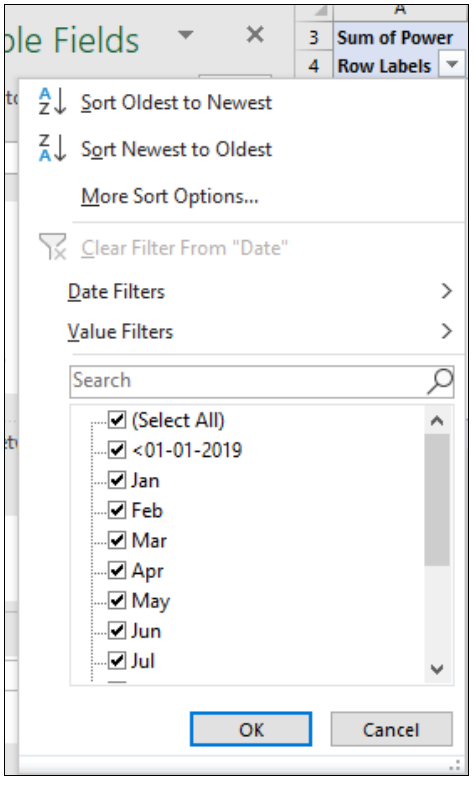


Figure 6.4.18 Filter Options

Report Filter

After you have summarized your data by creating an Excel Pivot Table, you can focus on specific portions of the data by using Report Filters.

You can add a Report Filter

In the PivotTable Field list, click on the field that you want to use as a Report Filter.

Drag the field into the Filters box, as shown in the screenshot below.

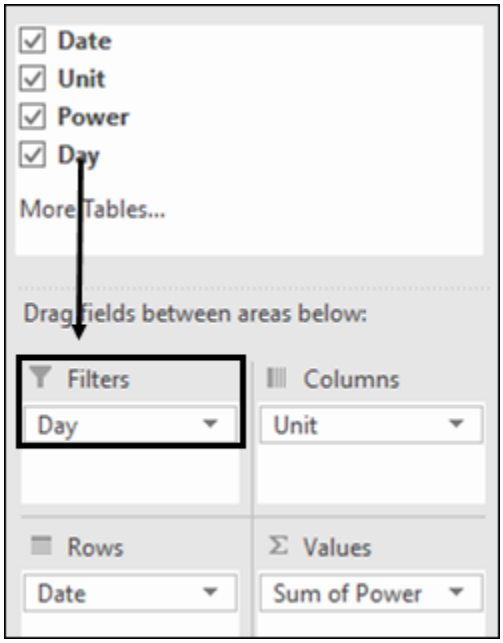


Figure 6.4.19 Report Filters

GENERATING MULTIPLE REPORTS

You can quickly create multiple pivot table reports with the **Show Report Filter Pages** feature.

Click inside your Pivot Table and in the menu ribbon under Pivot Table Analyse choose the **Options** tab and then select the **Options** drop-down and choose Show Report Filter Pages.

Each of your Filter items will have their unique Pivot Table in a separate Sheet with their Values and transactional metrics.

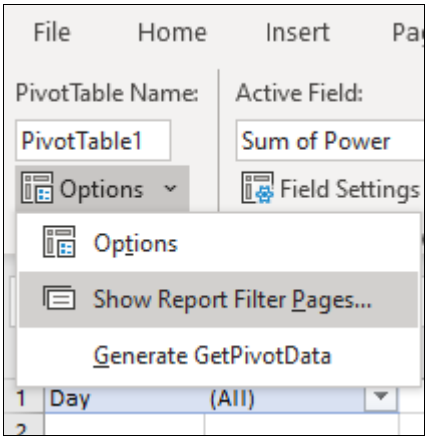


Figure 6.4.20Generate Multiple Report

CASE STUDY: In the previous case study CA Chandiwala wants you to do a WeekDaywise, Monthwise, and also unit wise Power Consumption and He also wants the report of each unit on a separate Excel sheet. You are a smart auditor how will you proceed smartly?

Strategy*:*

You can use **Report Filter** options for the same

1. Create a PivotTable by taking "Date" to **Rows**, "Day" to **Columns** & "Power " to **Values**
2. **Group** Date By Months
3. Drag Unit to Filter as shown in Figure

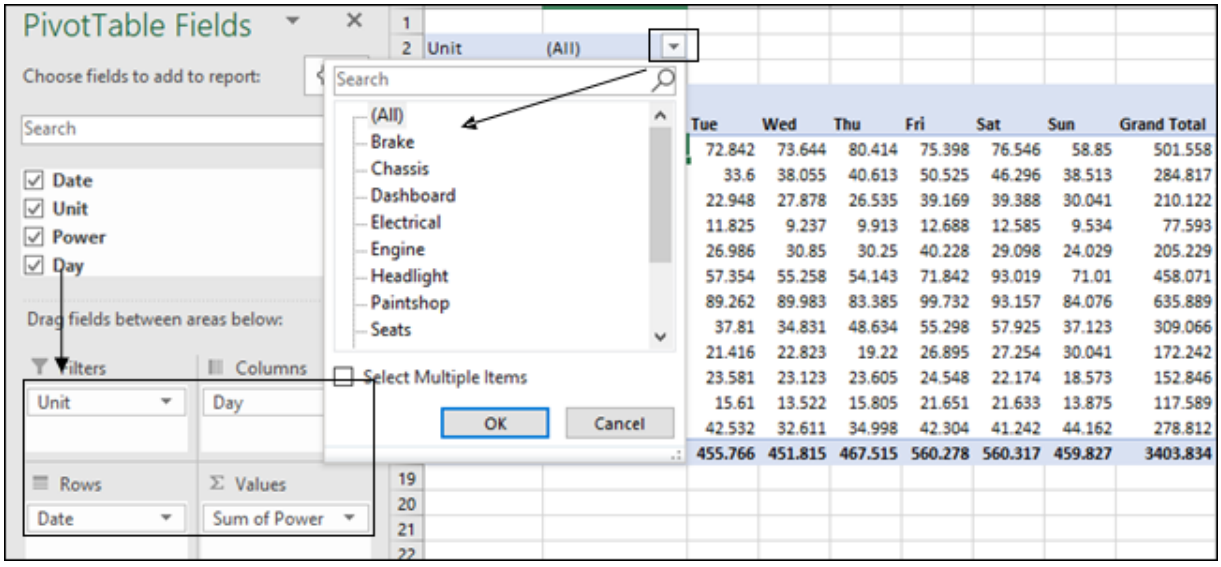


Figure 6.4.21 Filter by Unit

4. You click on Brake and data corresponding to Brake Unit is displayed

Unit	Brake							
Sum of Power	Column Labels							
Row Labels	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Grand Total
Jan	2.888	3.184	2.772	2.748	3.844	4.62	1.548	21.604
Feb	1.303	1.272	1.428	1.574	3.462	3.161	1.526	13.726
Mar	0.623	0.815	0.696	0.755	2.711	3.405	1.4	10.405
Apr	0.525	0.409	0.358	0.335	0.693	0.73	0.428	3.478
May	1.223	1.093	1.121	0.96	1.472	1.728	0.929	8.526
Jun	1.988	2.571	2.159	1.343	2.664	5.048	2.798	18.571
Jul	2.458	4.398	3.552	3.009	5.873	4.232	3.321	26.843
Aug	1.649	1.704	2.025	1.724	2.908	3.164	1.57	14.744
Sep	1.019	0.856	0.781	0.902	2.29	0.801	1.025	7.674
Oct	0.591	0.883	0.863	1.037	1.707	1.29	0.914	7.285
Nov	0.714	0.588	0.451	0.623	0.959	1.179	0.655	5.169
Dec	1.393	1.911	1.428	1.13	3.174	2.534	1.382	12.952
Grand Total	**16.374**	**19.684**	**17.634**	**16.14**	**31.757**	**31.892**	**17.496**	**150.977**

Figure 6.4.22 Filtered Data of Brake Unit

5. To generate **multiple pivot table reports** for each of the Units.
6. Under PivotTable Analyse choose the **Options** tab.
7. select the Options drop-down and choose **Show Report Filter Pages**.

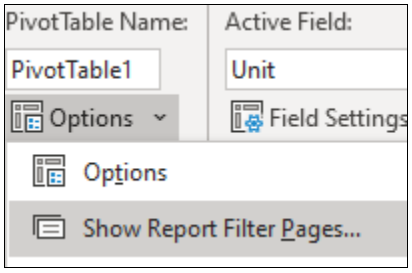


Figure 6.4.23Show Report Filter pages Option

8. Show Report Filter Pages Dialog box will appear select Unit as shown in Fig 6.4.24

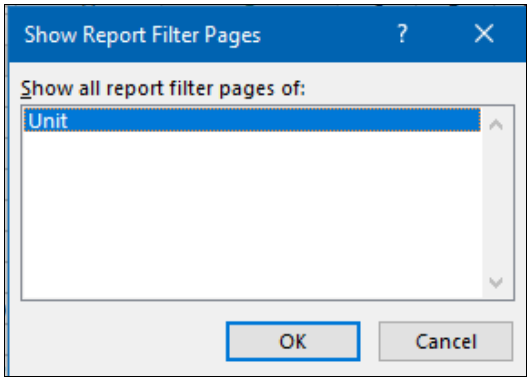
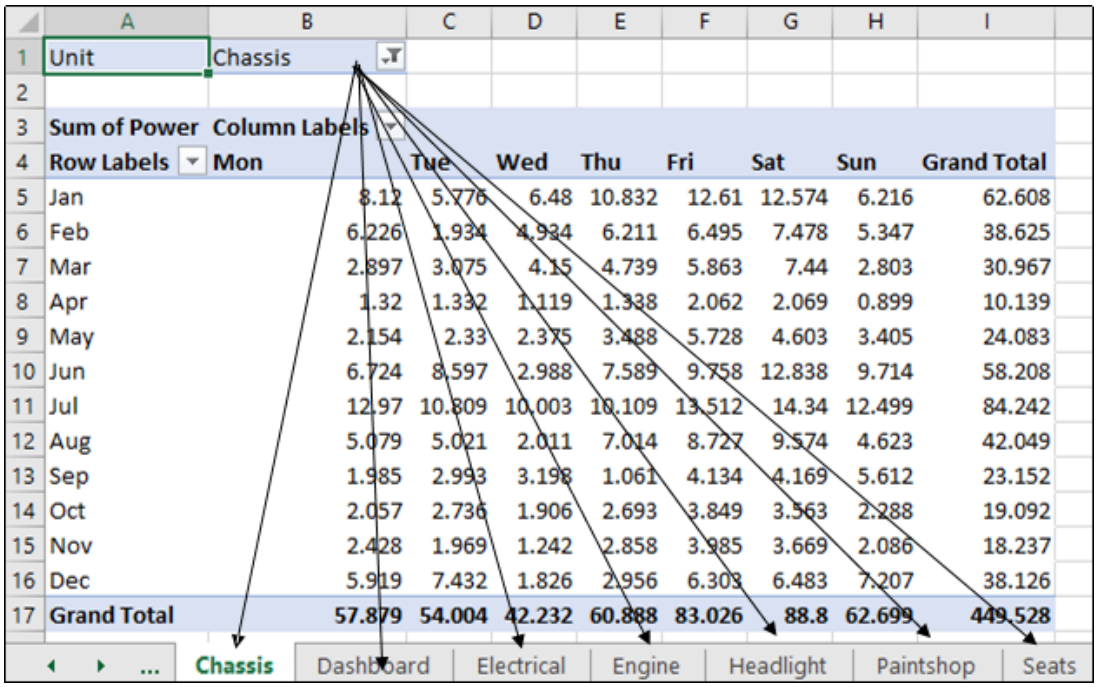


Figure 6.4.24 Show Report Filter Dialog Box

9. Each of your Filter itemsie Unit will have their unique Pivot Table in multiple separate Sheet with their Values and transactional metrics as shown in Figure 6.4.25



	A	B	C	D	E	F	G	H	I
1	Unit	Chassis							
2									
3	Sum of Power	Column Labels							
4	Row Labels	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Grand Total
5	Jan	8.12	5.776	6.48	10.832	12.61	12.574	6.216	62.608
6	Feb	6.226	1.934	4.934	6.211	6.495	7.478	5.347	38.625
7	Mar	2.897	3.075	4.15	4.739	5.863	7.44	2.803	30.967
8	Apr	1.32	1.332	1.119	1.338	2.062	2.069	0.899	10.139
9	May	2.154	2.33	2.375	3.488	5.728	4.603	3.405	24.083
10	Jun	6.724	8.597	2.988	7.589	9.758	12.838	9.714	58.208
11	Jul	12.97	10.809	10.003	10.109	13.512	14.34	12.499	84.242
12	Aug	5.079	5.021	2.011	7.014	8.727	9.574	4.623	42.049
13	Sep	1.985	2.993	3.198	1.061	4.134	4.169	5.612	23.152
14	Oct	2.057	2.736	1.906	2.693	3.849	3.563	2.288	19.092
15	Nov	2.428	1.969	1.242	2.858	3.985	3.669	2.086	18.237
16	Dec	5.919	7.432	1.826	2.956	6.303	6.483	7.207	38.126
17	Grand Total	57.879	54.004	42.232	60.888	83.026	88.8	62.699	449.528

Figure 6.4.25 Reports on Multiple Separate sheets

Gist: You have Created Report Filters and Multiple reports based on data in PivotTables

Commands learnt: PivotTables- Report Filter, Options

6.4.6 CALCULATIONS AND CALCULATED FIELD

Apart from sum PivotTable gives you the following options for summarisation

- Count
- Average
- Max
- Min
- Product
- Count Numbers
- StdDev
- StdDevp
- Var
- Varp

Apart from summarization also PivotTable gives you various options for showing the values where you can show surface Values as a

- % of Grand Total
- % of Column Total
- % of Row Total
- % Of
- % of Parent Row Total
- % of Parent Column Total
- % of Parent Total
- Difference From
- % Difference From
- Running Total in
- % Running Total in
- Rank Smallest to Largest
- Rank Largest to Smallest
- Index

Pivot has many calculation options available in PivotTables based on **Calculated Fields**

1. Click on any cell in pivot table
2. Go to PivotTable Analyse>Calculations> “Fields, Items & Sets” > “Calculated Field”

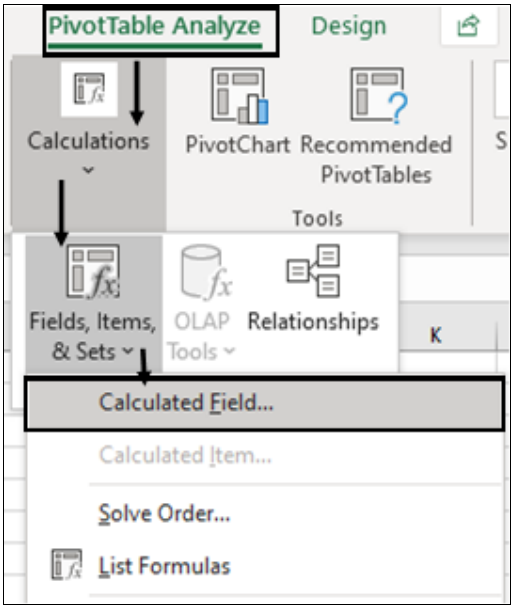


Figure 6.4.26 calculated Field in PivotTable

***CASE STUDY:** In the above case study, CA. Chandiwala wants to know the average Unit wisePower Consumption rather than Total Power Consumption, he also wants the power Consumption of each Unit as % of total power consumption. He has also obtained corresponding power consumption in the previous year and wants you to find the increase in power consumption in % terms.*

	A	B	C	D
1	Date	Unit	Power CY	Power PY
2	01-Jan	Brake	0.872	0.802
3	01-Jan	Chassis	0.656	0.604
4	01-Jan	Dashboard	2.354	2.095
5	01-Jan	Electrical	0.396	0.348
6	01-Jan	Engine	2.304	2.028
7	01-Jan	Headlight	0.830	0.739
8	01-Jan	Paintshop	0.388	0.345
9	01-Jan	Seats	4.248	3.866
10	01-Jan	Steering	0.340	0.306
11	01-Jan	Suspension	1.404	1.264
12	01-Jan	Transmission	0.800	0.736
13	01-Jan	Windshields	1.200	1.056
14	02-Jan	Brake	0.348	0.306
15	02-Jan	Chassis	0.712	0.634

Figure 6.4.27Data with previous year power consumption

***Strategy**:*

Apart from Sum, Many other summarizing options are available in Pivots

10. Create a PivotTable by taking "Unit" to **Rows**& "Power CY" to **Values**
11. By default, excel summarizes pivot data by "**sum**" or "**count**" depending on the data type**.**
12. But you can change it.
13. **Right-click** on pivot table **values**
14. Select "**Summarize Values By**" and select the "**Average"** option as shown in Fig.6.4.28.

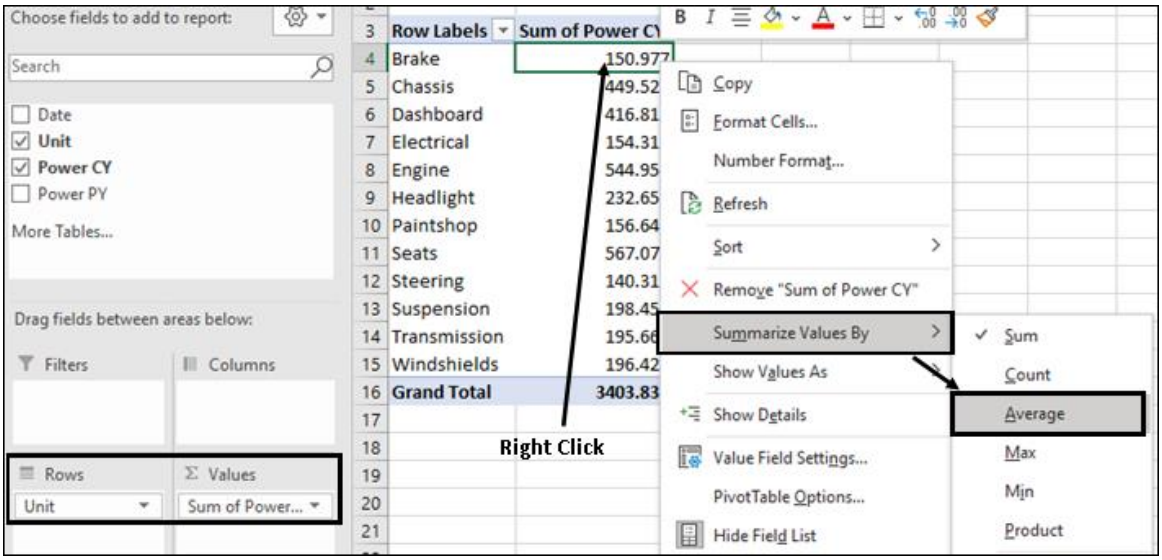


Figure 6.4.28 Summarise by "Average"

Some other options are Max, Min, Product etc.

3	Row Labels	Average of Power CY
4	Brake	0.414
5	Chassis	1.232
6	Dashboard	1.142
7	Electrical	0.423
8	Engine	1.493
9	Headlight	0.637
10	Paintshop	0.429
11	Seats	1.554
12	Steering	0.384
13	Suspension	0.544
14	Transmission	0.536
15	Windshields	0.538
16	**Grand Total**	**0.777**

Figure 6.4.29 Unitwise average power consumption

15. You drag “Power CY” to **Values** once more and in the second Column “**Summarise values by**” Sum
16. Right-click on any pivot table value in Value By sum Column select “**Show values as**”
17. Under “**Show value as”** Select **"% of Column Total**”

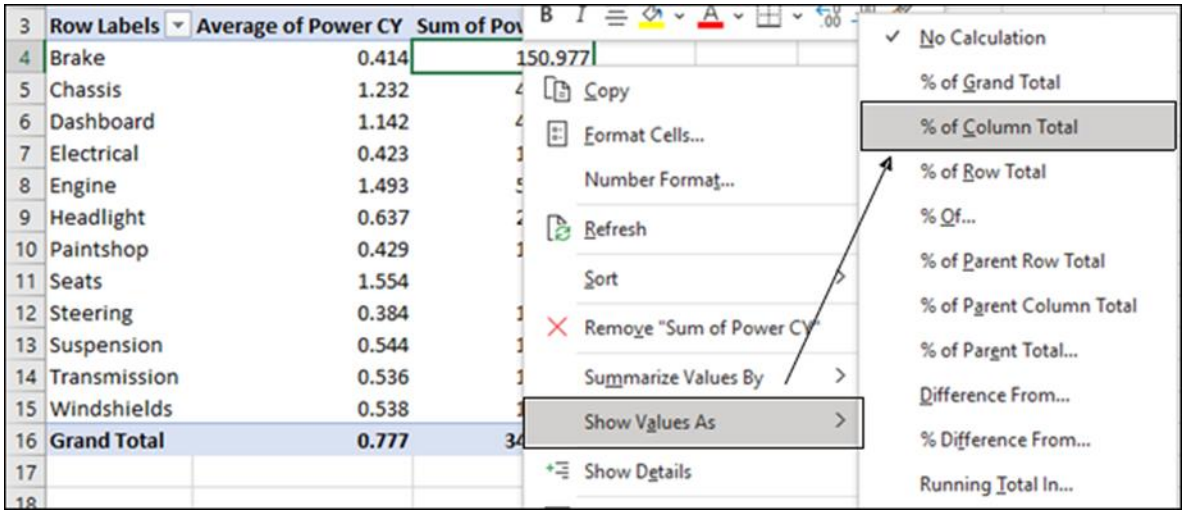


Figure 6.4.30Show values as % of Column Total

There are numerous other options to choose from as discussed above.

3	Row Labels	Average of Power CY	Sum of Power CY
4	Brake	0.414	4.44%
5	Chassis	1.232	13.21%
6	Dashboard	1.142	12.25%
7	Electrical	0.423	4.53%
8	Engine	1.493	16.01%
9	Headlight	0.637	6.84%
10	Paintshop	0.429	4.60%
11	Seats	1.554	16.66%
12	Steering	0.384	4.12%
13	Suspension	0.544	5.83%
14	Transmission	0.536	5.75%
15	Windshields	0.538	5.77%
16	**Grand Total**	**0.777**	**100.00%**

Figure 6.4.31% power Consumption

18. Go to **PivotTable Analyse> Calculations > "Fields, Items & Sets" > "Calculated Field"**

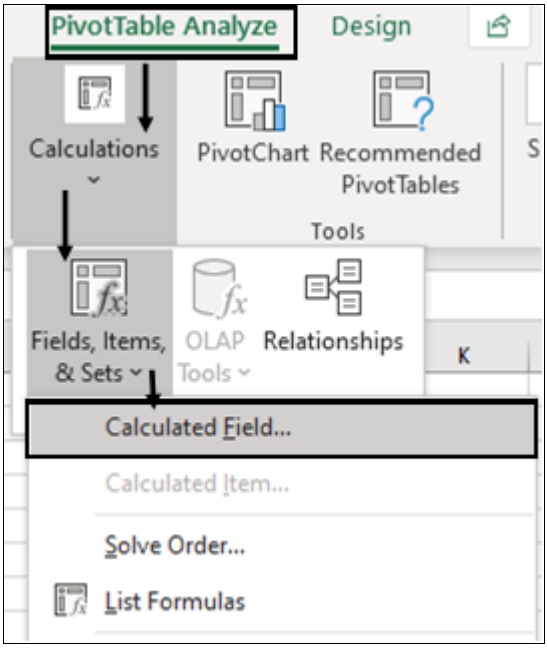


Figure 6.4.32 calculated Field

19. Define a new calculated Field by giving a **Name** "Percent Increase" and Specifying the **formula** which is = 'Power CY'/'Power PY'%

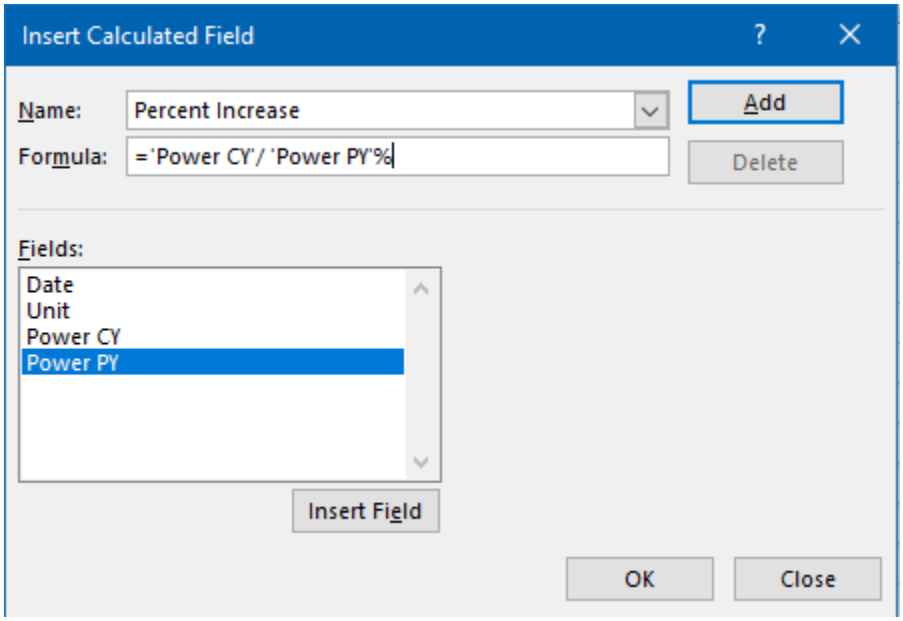


Figure 6.4.33 Insert calculated Field Dialog Box

20. You now get one more column in which you see Power consumption as a percentage of corresponding power consumption in the previous year.

3	Row Labels	Average of Power CY	Sum of Power CY	Sum of Percent Increase
4	Brake	0.414	4.44%	111.158
5	Chassis	1.232	13.21%	110.795
6	Dashboard	1.142	12.25%	111.198
7	Electrical	0.423	4.53%	111.160
8	Engine	1.493	16.01%	110.976
9	Headlight	0.637	6.84%	111.263
10	Paintshop	0.429	4.60%	111.170
11	Seats	1.554	16.66%	111.264
12	Steering	0.384	4.12%	111.094
13	Suspension	0.544	5.83%	111.185
14	Transmission	0.536	5.75%	111.148
15	Windshields	0.538	5.77%	111.000
16	**Grand Total**	**0.777**	**100.00%**	**111.100**

Figure 6.4.34 Percent Increase in power Consumption

Gist: You have learnt to show the values by different parameters as percentages etc in PivotTables

Commands learnt: PivotTables- Summarise by, Show Values As calculated fields

6.4.7 CONSOLIDATE MULTIPLE SHEETS IN PIVOTTABLE

You can consolidate data from different sources to combine data into one report. As an example, if you have a PivotTable of Sales figures for each Quarter, you can use a data consolidation to roll up these figures into an Annual report.

You can use the **Multiple consolidation ranges** option in **PivotTable** as follows

1. Click a blank cell in the workbook.
2. Use Short cut **Alt+D+P to** create PivotTable
3. Pivot table and Pivot Chart wizard will Appear
4. On the Step 1 page of the wizard, click Multiple consolidation ranges and then click Next.
5. On Step 2a page of the wizard, click **Create a Single Page Field for me**, and then click Next.
6. On Step 2b page of the wizard, do the following:
7. Go back to the workbook and select the cell range, and
8. then back on the PivotTable and PivotChart Wizard, click Add.
9. Repeat for other ranges
10. Click Next.
11. Click Finish and all the ranges are consolidated and Pivot created.

CASE STUDY: Your Principal CA Chandiwala gives you a workbook containing four sheets for each quarter sales, he wants you to consolidate in a single sheet and create a Pivot based on the same.

	A	B	C	D
1	DATE	AMOUNT		
2	24-04-2019	45,460.00		
3	29-05-2019	81,068.00		
4	05-06-2019	7,493.00		
5	06-06-2019	43,877.00		
6	09-06-2019	24,659.00		
7	29-06-2019	22,906.00		
8	30-06-2019	8,954.00		

Q1 Q2 Q3 Q4

Figure 6.4.35 Quarterwise sales

Strategy:

You can use the **Multiple consolidation ranges** option in **PivotTable** to consolidate quarterwise sales

1. Click a blank cell in the workbook.

2. Use Short cut **Alt+D+P to** create PivotTable

3. Pivot table and Pivot Chart wizard will Appear as shown in Figure

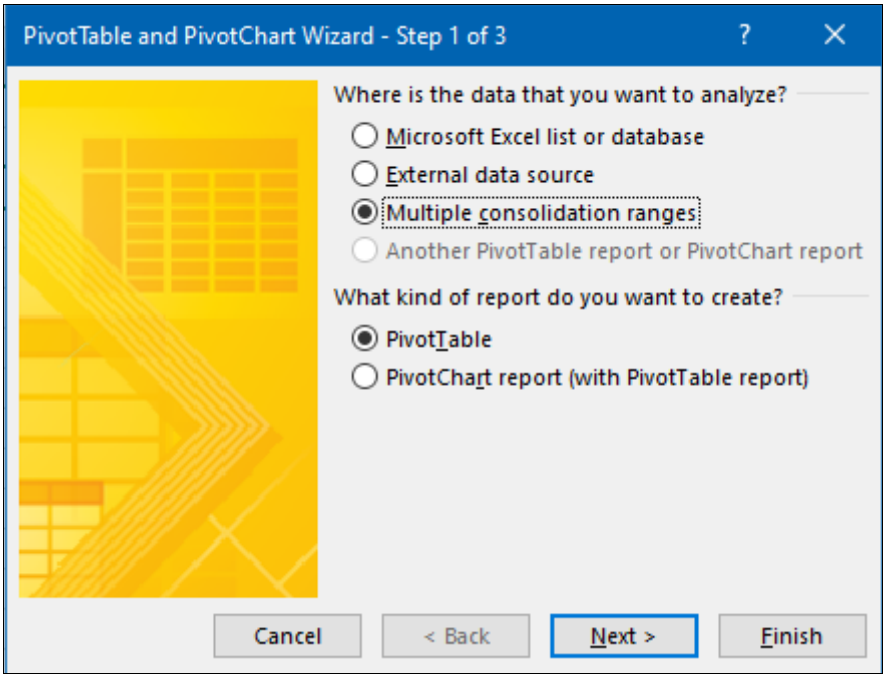


Figure 6.4.36 Pivot table and Pivot Chart wizard

4. On Step 1 page of the wizard, click Multiple consolidation ranges, And in What Kind of Report... select Pivot table and then click Next.

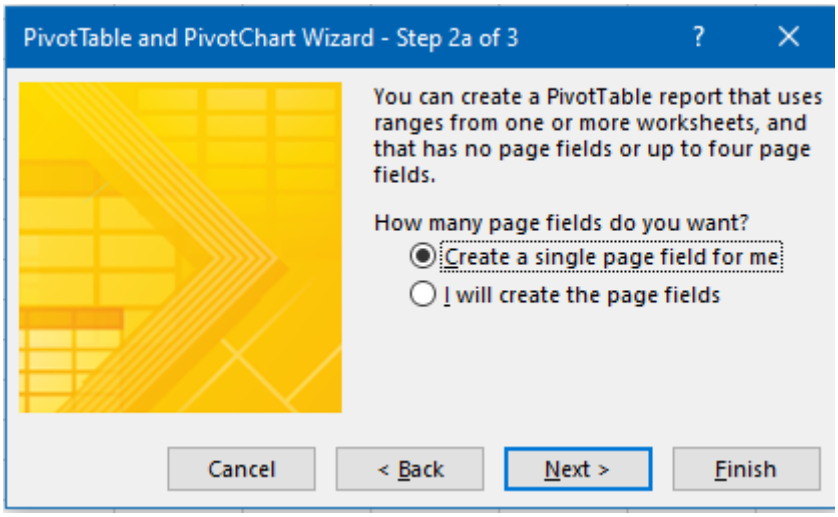


Figure 6.4.37 Step 2a of Wizard

5. On Step 2a page of the wizard, click **Create a Single Page Field for me**, and then click Next.

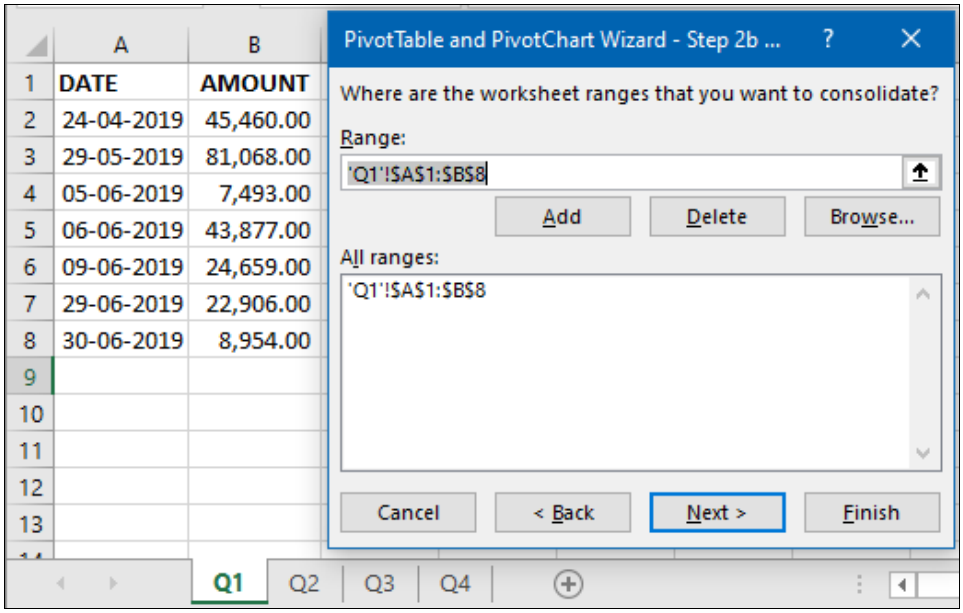


Figure 6.4.38 Step 2b Of Wizard- Select Ranges

6. On Step 2b page of the wizard, do the following:
7. Go back to the workbook and select the cell range, and
8. then back on the PivotTable and PivotChart Wizard, click Add.
9. Repeat for other ranges

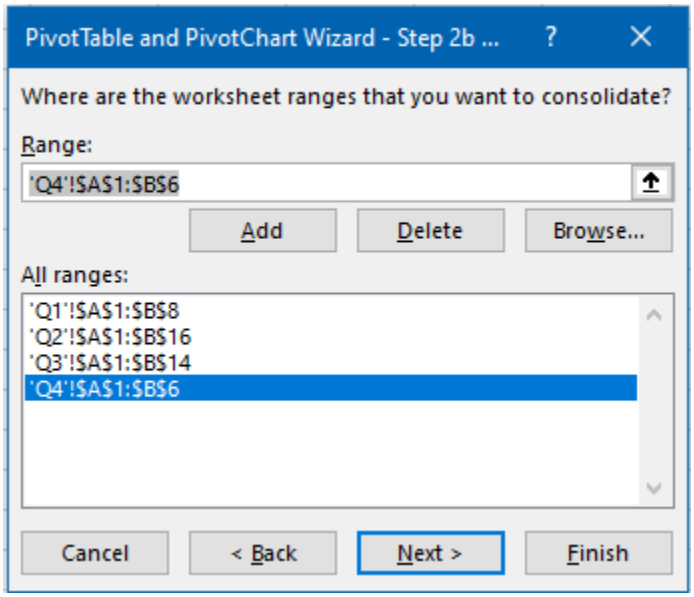


Figure 6.4.39 Select all ranges to be consolidated

10. Click Next.

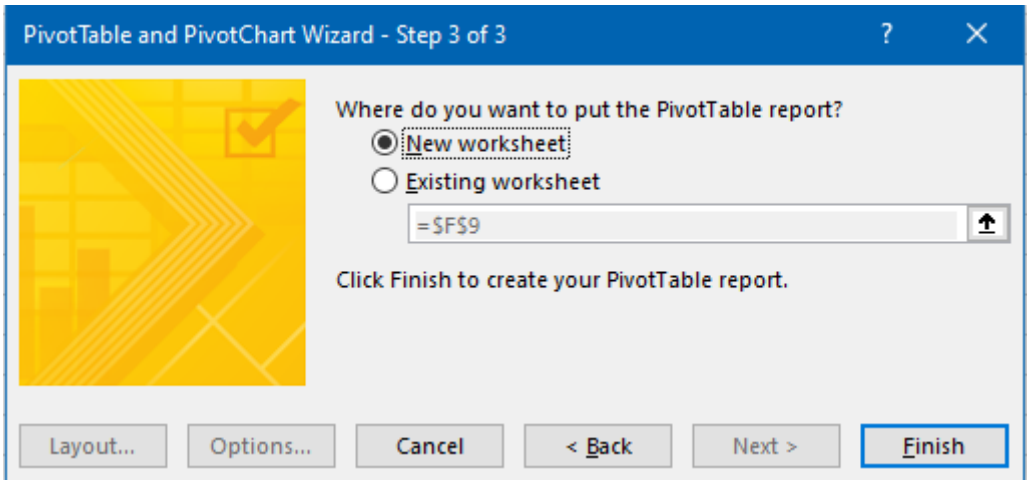


Figure 6.4.40 Specify where you want results

11. Click Finish.

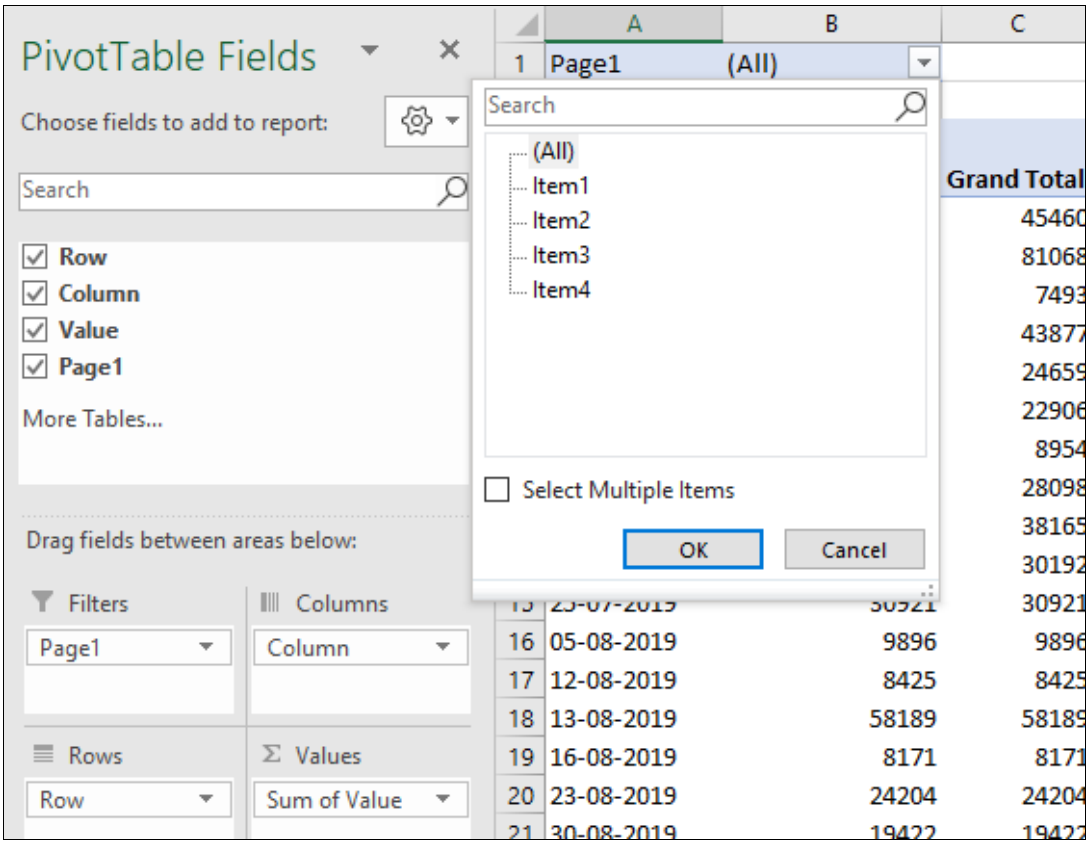


Figure 6.4.41 Consolidated Pivot table

12. A pivot table is created consolidating all sheets which appear as Report filter options.

Gist: You have learnt to consolidate Multiple Sheets into a PivotTable.

Commands learnt: PivotTables- Consolidate Multiple Ranges.

6.4.8 SLICERS AND TIMELINES

Slicers are easy to create and to use, and they make it a snap to **filter** the contents of the pivot table on more than one field, similarly **Timelines** are a kind of Slicer on based on time

The slicer/ Timeline can be accessed from **PivotTable Analyse> Filter>Insert Slicer** as shown in Figure

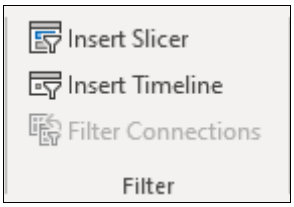


Figure 6.4.42 Insert Slicers& Timelines

CASE STUDY: As a part of the investigation your principal CA Chandiwala has given you data relating to power consumption in different units of a company for various dates as shown in Fig 6.4.43. He wants you to create slicers and Timelines for Unitwise power consumption across Months

	A	B	C	D
1	**Date**	**Unit**	**Power**	**Day**
2	01-01-2019	Brake	0.872	Tue
3	01-01-2019	Chassis	0.656	Tue
4	01-01-2019	Dashboard	2.354	Tue
5	01-01-2019	Electrical	0.396	Tue
6	01-01-2019	Engine	2.304	Tue
7	01-01-2019	Headlight	0.830	Tue
8	01-01-2019	Paintshop	0.388	Tue
9	01-01-2019	Seats	4.248	Tue
10	01-01-2019	Steering	0.340	Tue
11	01-01-2019	Suspension	1.404	Tue
12	01-01-2019	Transmission	0.800	Tue

Figure 6.4.43 Datewise data

Strategy:

You can achieve your objective with a pivot table; you can use Slicers Function in Pivot tables.

1. You will first **create** a Pivot table with Dates in Rows and Power in Values.
2. You will then **Group** dates into **months.**

3. Next you will Go to Pivot table analyse >Filter> Insert Slicers

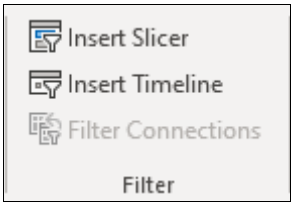


Figure 6.4.44 Pivot table analyse >Filter> Insert Slicers

4. An **Insert Slicer Dialog Box** will appear

5. Select **Unit**

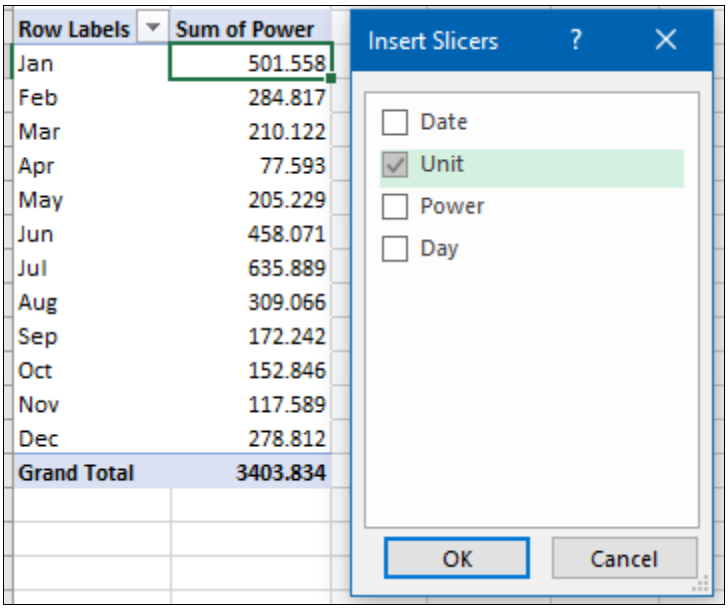


Figure 6.4.45 Insert Slicers Dialog Box

6. A **Slicer** based on Unit will appear, when you click on any of those Unit relevant Data is filtered.

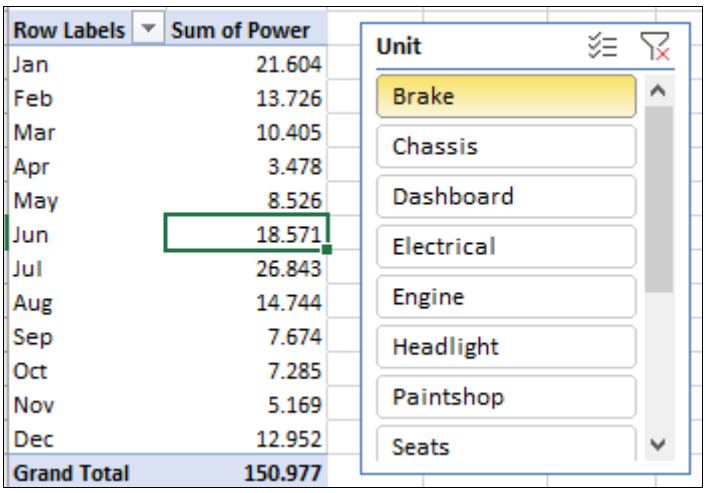


Figure 6.4.46 slicer based on Unit

7. To Insert Timeline, you will go to **PivotTable Analyse>Filter > Insert Timeline**

8. **Insert Timelines Dialog Box** will appear select **Date**

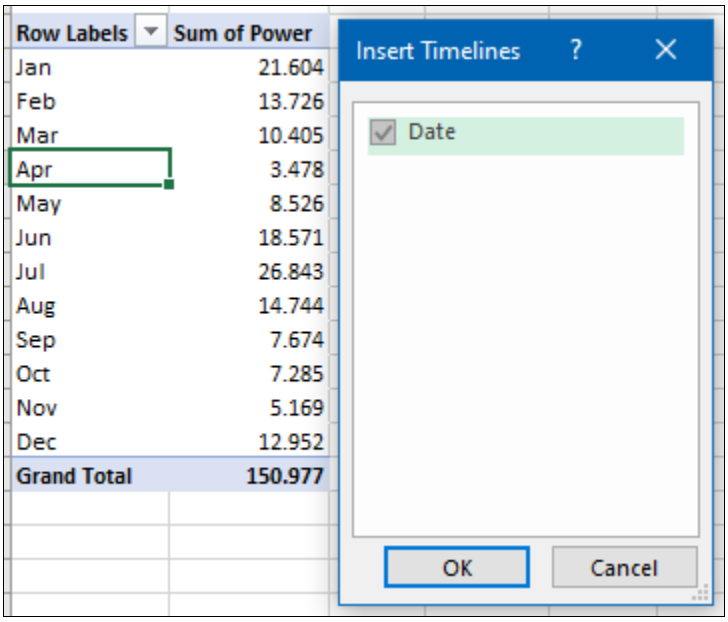


Figure 6.4.47 Insert Timelines

9. A **Timeline** based on Dates will appear, when you click on any of the period relevant Data is filtered.

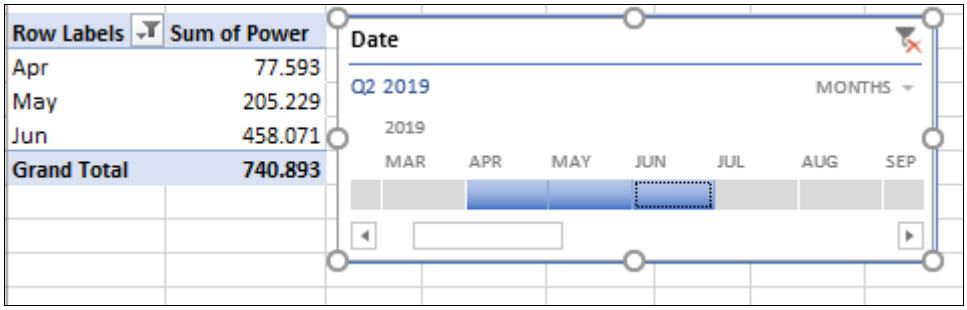


Figure 6.4.48 Timeline based Filtering

Gist: You have learnt insert Slicers and Timelines into a PivotTable.

Commands learnt: PivotTables-Slicers/Timelines.

6.4.9 CREATING A DYNAMIC DASHBOARD

Microsoft Excel is an excellent choice for creating and distributing dashboards because of its ease of use. A dashboard is mostly a single Page/Screen view with KPIs that helps managers and business leaders in getting a high-level overview and take some decisions based on it. It usually consists of charts, tables, views that are backed by data. Dashboards are often sometimes referred to as management dashboards or information dashboards or dashboard reports.

These dashboards could be static or interactive (where the user can make selections and change views and the data would dynamically update).

Creating an Excel Dashboard is a multi-step process and there are some key things you need to keep in mind when creating it.

The first step in dashboard creation is, you need to be clear about the objectives of the dashboard.

As a best practice, divide your Excel workbook into three parts

- **Data** – You can have one or more than one worksheet that contains the raw data.
- **Calculations** –You will make all calculations.
- **Dashboard** – This is the sheet that has the dashboard. Ideally, it is a single page view that shows the required analysis backed by data.

Dashboards-Best Practices and things to avoid

- **Naming your Charts/Section**: the dashboard is like a story and Charts are Characters name them and have some order to make it intelligible.
- **Restrict Movement in the dashboard area:** Hide all unnecessary rows or columns to make sure the user doesn't go beyond the dashboard
- **Use Freeze panes wherever required**.
- **Provide a User Help with proper legends**.

- **Use Space Wisely.**
- **Don't clutter your Dashboards**: Focus on the objective of the dashboard and only place relevant piece of information
- **Don't keep extra data in your workbook**: If you need that data, create a copy of the dashboard and keep it as the backup.

CASE STUDY: Your principal CA Chandiwala has given you data relating to power consumption in different units of a company for various dates. He wants you to create a dashboard where he could interactively monitor for Unitwise month-wise and Weekdaywise power consumption and he also wants that whenever new data is added Dashboard is Refreshed.

	A	B	C	D
1	**Date**	**Unit**	**Power**	**Day**
2	01-01-2019	Brake	0.872	Tue
3	01-01-2019	Chassis	0.656	Tue
4	01-01-2019	Dashboard	2.354	Tue
5	01-01-2019	Electrical	0.396	Tue
6	01-01-2019	Engine	2.304	Tue
7	01-01-2019	Headlight	0.830	Tue
8	01-01-2019	Paintshop	0.388	Tue
9	01-01-2019	Seats	4.248	Tue
10	01-01-2019	Steering	0.340	Tue
11	01-01-2019	Suspension	1.404	Tue
12	01-01-2019	Transmission	0.800	Tue

Figure 6.4.49 Power Consumption data.

Strategy:

You can build a Dynamic Dash Board using features like Table, Pivots, Slicers & Charts within Excel.

Since the objective is Unitwise month-wise and Weekdaywise power consumption, you will build dashboard components with this in mind.

1. You will first **set the data as a Table (Insert >Table)** so that if more data is added Table with expand dynamically and name it as **Tbl_Power**

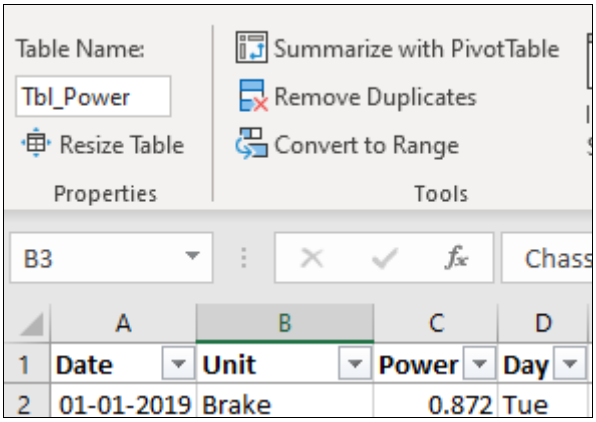


Figure 6.4.50 Set up as Table

2. You will then create a PivotTable and Drag dates to **Row** and Power to **Values**
3. You will then **Group** dates as Months
4. Go to **PivotTable analyze> Pivot table name** and Name The Pivot as **Tbl_Month**

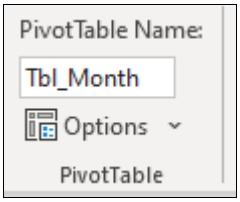


Figure 6.4.51 Name the PivotTable

5. Click on Pivot Chart and Select a Line Chart and a Pivot Chart is created as shown in figure.

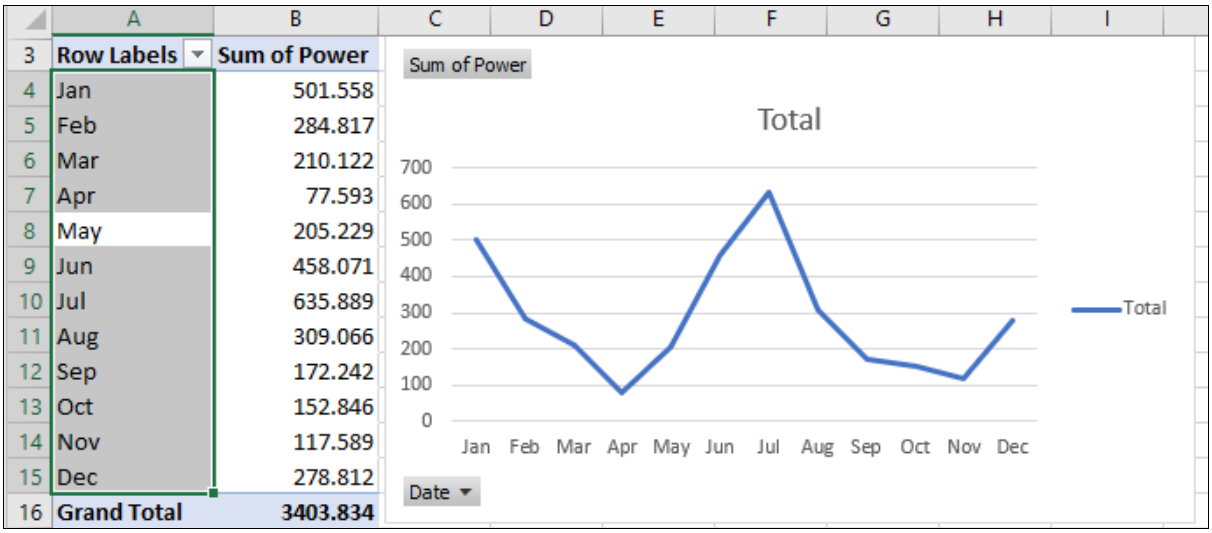


Figure 6.4.52 Monthwise Power Consumption with Chart

6. You will then create one more PivotTable in a new worksheet and Drag days to **Row** and Power to **Values**

7. Go to **PivotTable analyze> Pivot table name** and Name The Pivot as **Tbl_Day**

8. Click on Pivot Chart and Select a Pie Chart and a Pivot Chart is created as shown in figure.

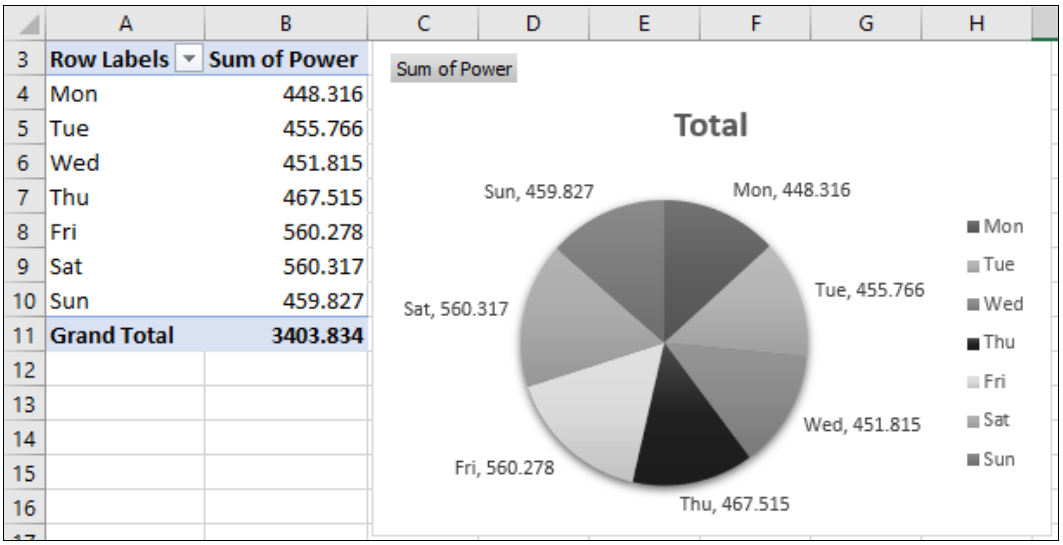


Figure 6.4.53 Day-wise Power Consumption

9. You will create Another PivotTable and Drag "Unit" to **Row** and Power to **Values**

10. Go to **PivotTable analyze> Pivot table name** and Name The Pivot as **Tbl_Unit**

11. Click on Pivot Chart and Select a Bar Chart and a Pivot Chart is created as shown in the figure.

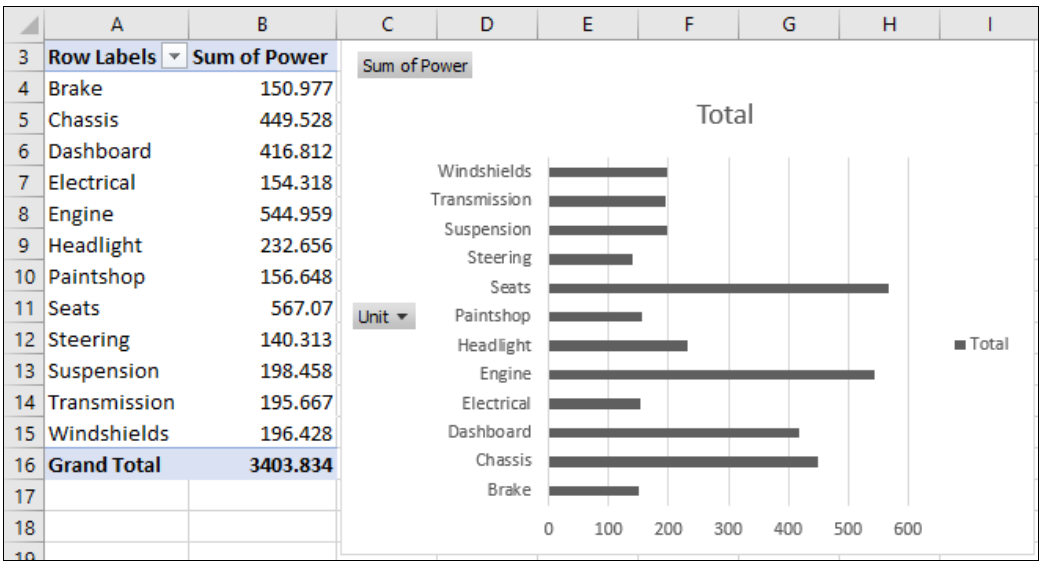


Figure 6.4.54 Unitwise Consumption

12. In three sheets you have built the components of a Dashboard.

13. Insert a Blank sheet and In Row 1 give it a name Power Consumption Dashboard

14. In the Next Step cut the charts from the three sheets and Paste it in the Blank Sheet and place them properly as shown in figure.

15. Give Title to the Charts

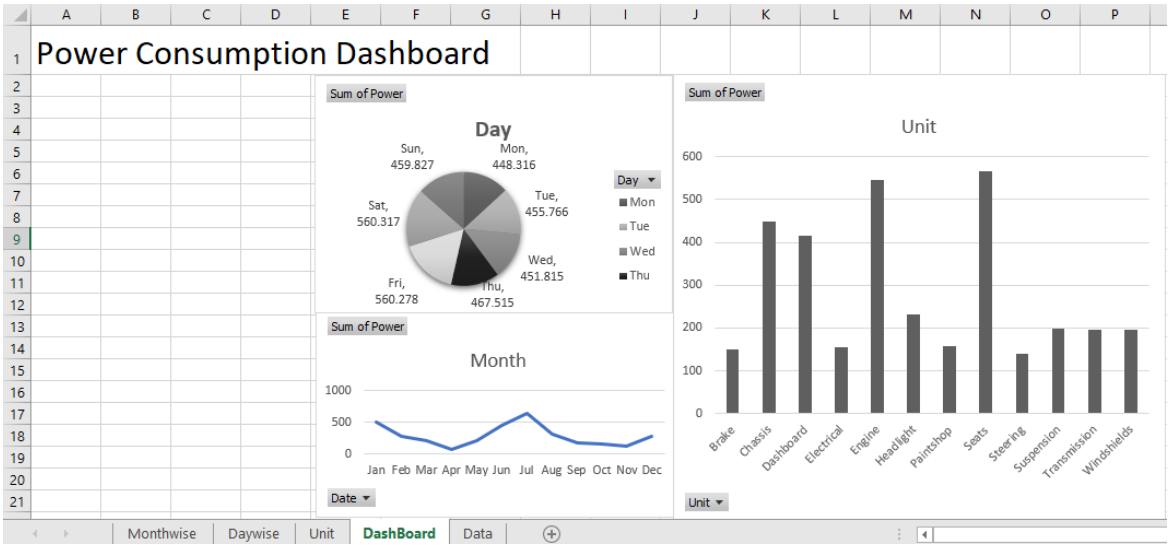


Figure 6.4.55 Arrange Charts on the dashboard

16. To make them interactive click on Monthwise Chart and **Pivot Chart analyse insert slicers** for **unit** and **day** as shown in figure

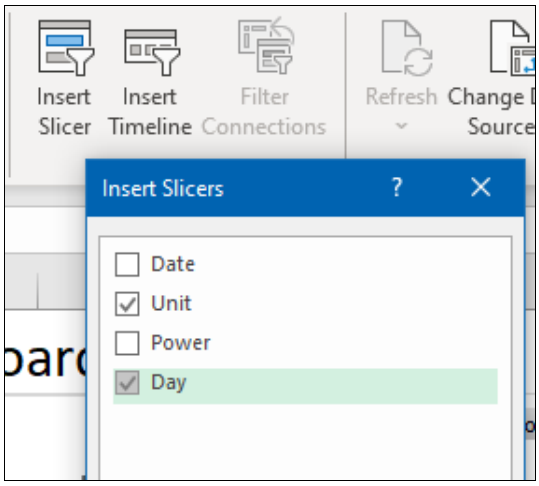


Figure 6.4.56 Insert Slicers

17. Slicers will be created arrange as shown in Figure

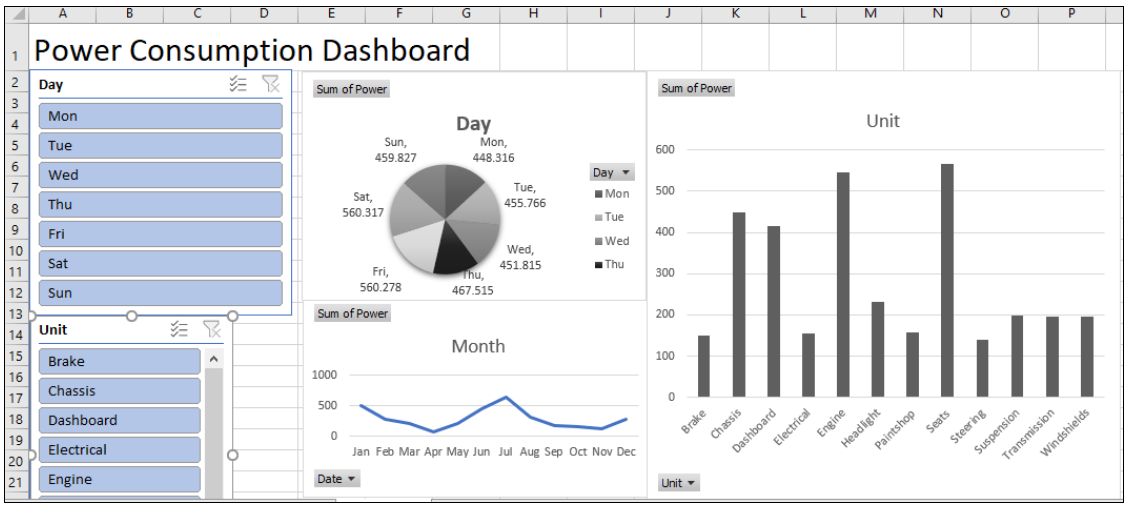


Figure 6.4.57 Arrange slicers

18. Click on Unit slicer, Slicer tab will appear, in Buttons group increase Columns to 2

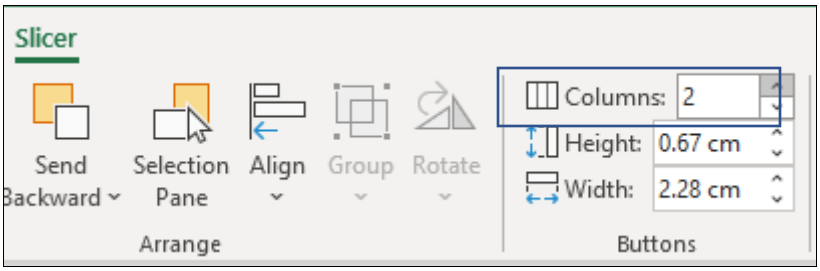


Figure 6.4.57 Under Buttons increase columns to 2

19. Unit Slicer will be in two columns as shown in Figure

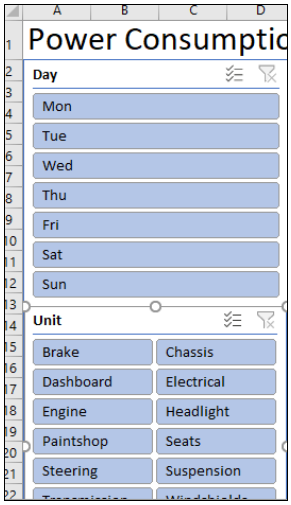


Figure 6.4.58 Slicer in Two columns

20. Go to Slicer Ribbon and click on **Report Connections**
21. **Report Connections Dialog Box** will appear select all pivot tables.
22. **Slicer** will **get linked** with all Pivots**,**
23. Repeat the process for **Day Slicer.**

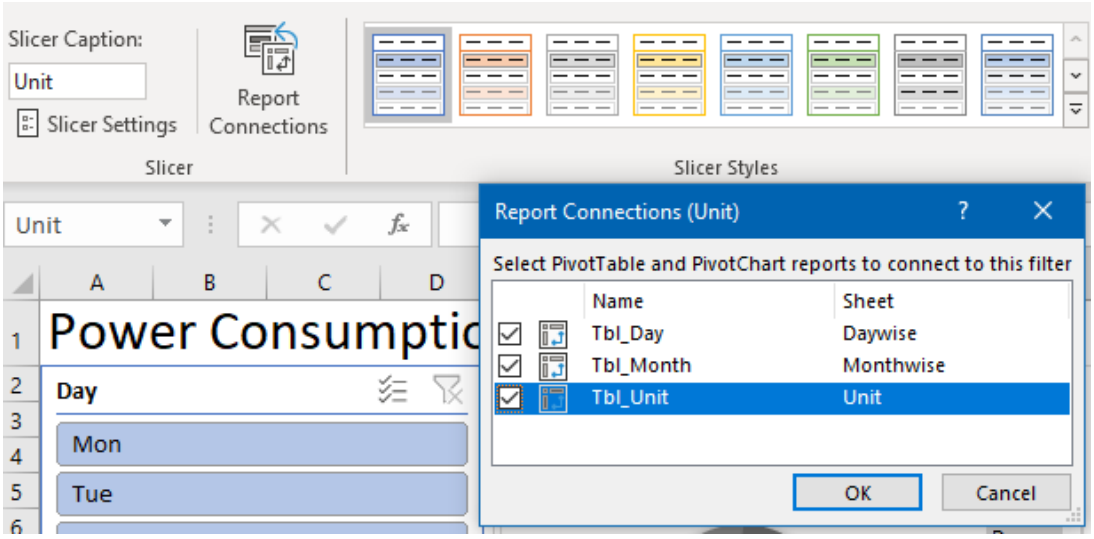


Figure 6.4.59 Report Connections

24. Go to View tab, Unselect gridlines to have a neat view

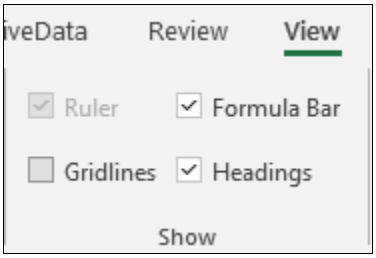


Figure 6.4.60 Remove Gridlines

25. Your Dynamic Dashboard is ready,
26. Select Mon and Tue in **Day Slicer** and Brakes and Chasis in **Unit slicer** and see all charts respond accordingly.

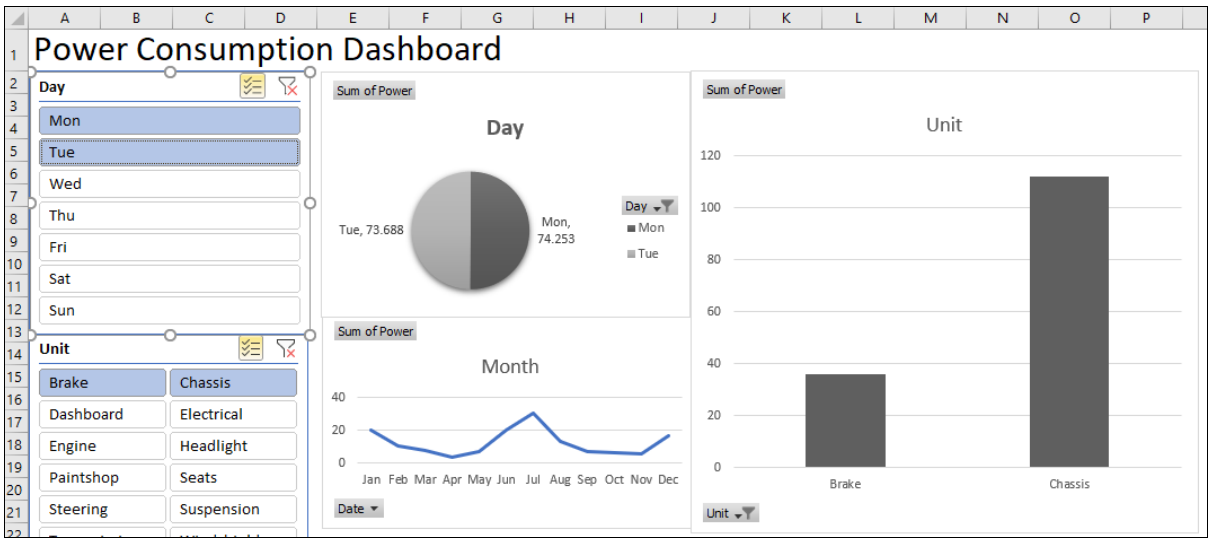


Figure 6.4.61Dynamic Dashboard with figures for power consumption on Mon/ Tue for Brake / Chassis

Gist: You have created a power Consumption dashboard.

Commands learnt: Tables, PivotTables, Charts, Slicers.

6.4.10 DATA CONNECTIONS, DATA MODELLING

Data Modelling is a process of formulating data in a structured format. It assists in analysing data easily which helps in meeting business requirements.

A data model describes information in a systematic way that allows it to be stored and retrieved efficiently, it emphasizes on what data is needed and how it should be organized rather than what analysis can be done. Data Model is more like an architect's building plan that set a relationship between various data items.

In the same way, the Data Model is used for building a model where data from various sources can be combined by creating relationships among the data sources. A Data Model assists in integrating the tables, which ultimately enables wide-ranging analysis using PivotTables, Power Pivot, and Power View.

Have you ever wondered how to create a report that connects multiple related tables in a single PivotTable?

A Data Model is a feature, just for that, that is for integrating data from multiple tables, effectively building a relational data source inside an Excel workbook.

When you import two or more tables from a database a Data Model is created. The relationships between those tables already existing are used to create the Data Model in Excel.

Apart from a **Data Model** you can **create connections** directly connect to the data source

You can Build a PivotTable from the data model rather than a single Excel table that offers many advantages.

- You can create a PivotTable that uses fields from multiple tables.
- The formulas you can write are much powerful than available in a traditional PivotTable.

- You can pick and choose rows and columns using named sets.
- You can create connections directly connect to the data source rather than copy/paste data into a worksheet,
- You can use a Get & Transform query (to clean the data before it arrives,
- You can connect to multiple data sources (eg, a csv file, a database table, and an Excel workbook) in a single model.
- Once built, you can just Refresh the report in later periods.

CASE STUDY: As a part of the investigation your principal CA Chandiwala has given you expense data relating to a period but the data has Accounting Codes rather than Expense Names, He has to repeatedly refer to codes with names which are in a different area in the workbook, and still he is not able to do much meaningful analysis. He wants you to find a way to build Pivot Tables based on Expense Account Names rather than Accounting Codes.

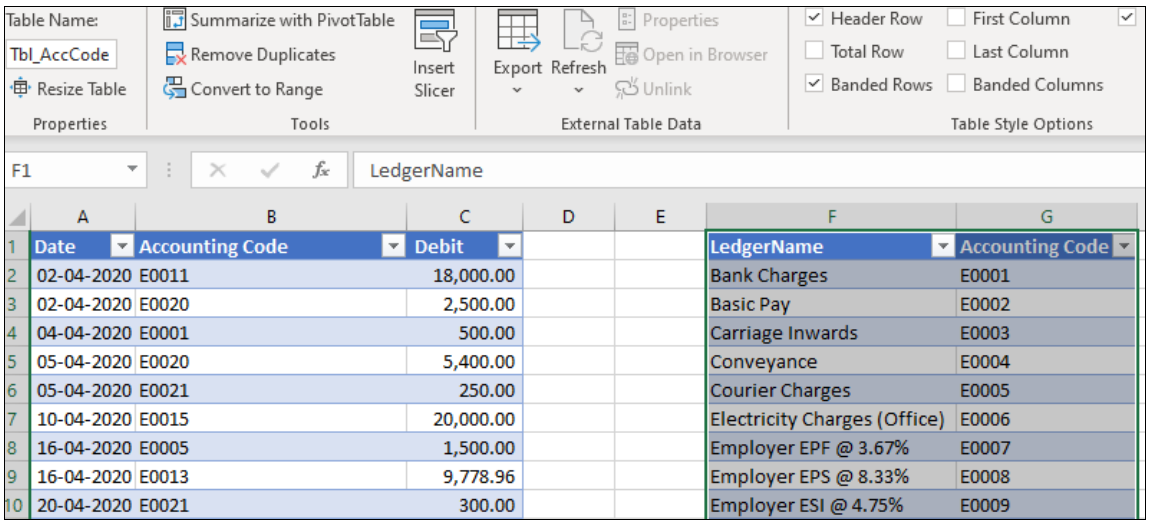
	A	B	C	D	E	F	G
1	**Date**	**Accounting Code**	**Debit**			**LedgerName**	**Accounting Code**
2	02-04-2020	E0011	18,000.00			Bank Charges	E0001
3	02-04-2020	E0020	2,500.00			Basic Pay	E0002
4	04-04-2020	E0001	500.00			Carriage Inwards	E0003
5	05-04-2020	E0020	5,400.00			Conveyance	E0004
6	05-04-2020	E0021	250.00			Courier Charges	E0005
7	10-04-2020	E0015	20,000.00			Electricity Charges (Office)	E0006
8	16-04-2020	E0005	1,500.00			Employer EPF @ 3.67%	E0007
9	16-04-2020	E0013	9,778.96			Employer EPS @ 8.33%	E0008
10	20-04-2020	E0021	300.00			Employer ESI @ 4.75%	E0009
11	25-04-2020	E0013	3,691.27			Fees & taxes	E0010
12	28-04-2020	E0021	325.00			Godown Rent	E0011

Figure 6.4.62 Expense data

Strategy:

You can achieve your objective data model feature in Pivot tables.

1. First set up your data as tables.
2. To create a table, select any cell in the range and press CTRL+T. Specify a name for your table from the design tab.



	A	B	C	D	E	F	G
1	Date	Accounting Code	Debit			LedgerName	Accounting Code
2	02-04-2020	E0011	18,000.00			Bank Charges	E0001
3	02-04-2020	E0020	2,500.00			Basic Pay	E0002
4	04-04-2020	E0001	500.00			Carriage Inwards	E0003
5	05-04-2020	E0020	5,400.00			Conveyance	E0004
6	05-04-2020	E0021	250.00			Courier Charges	E0005
7	10-04-2020	E0015	20,000.00			Electricity Charges (Office)	E0006
8	16-04-2020	E0005	1,500.00			Employer EPF @ 3.67%	E0007
9	16-04-2020	E0013	9,778.96			Employer EPS @ 8.33%	E0008
10	20-04-2020	E0021	300.00			Employer ESI @ 4.75%	E0009

Figure 6.4.63 Set up data As Tables

3. Now, go to the **data ribbon** & click on the **relationships** button.

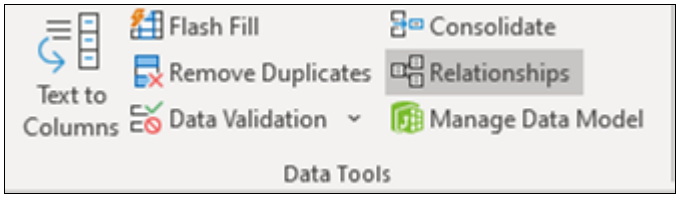


Figure 6.4.64Relationship on Data Tab

4. Click New to create a new relationship.
5. Select the Source table & column name and Map it to the target table & column name, in this case accounting Code.
6. It does not matter which order you use here. Excel is smart enough to adjust the relationship.

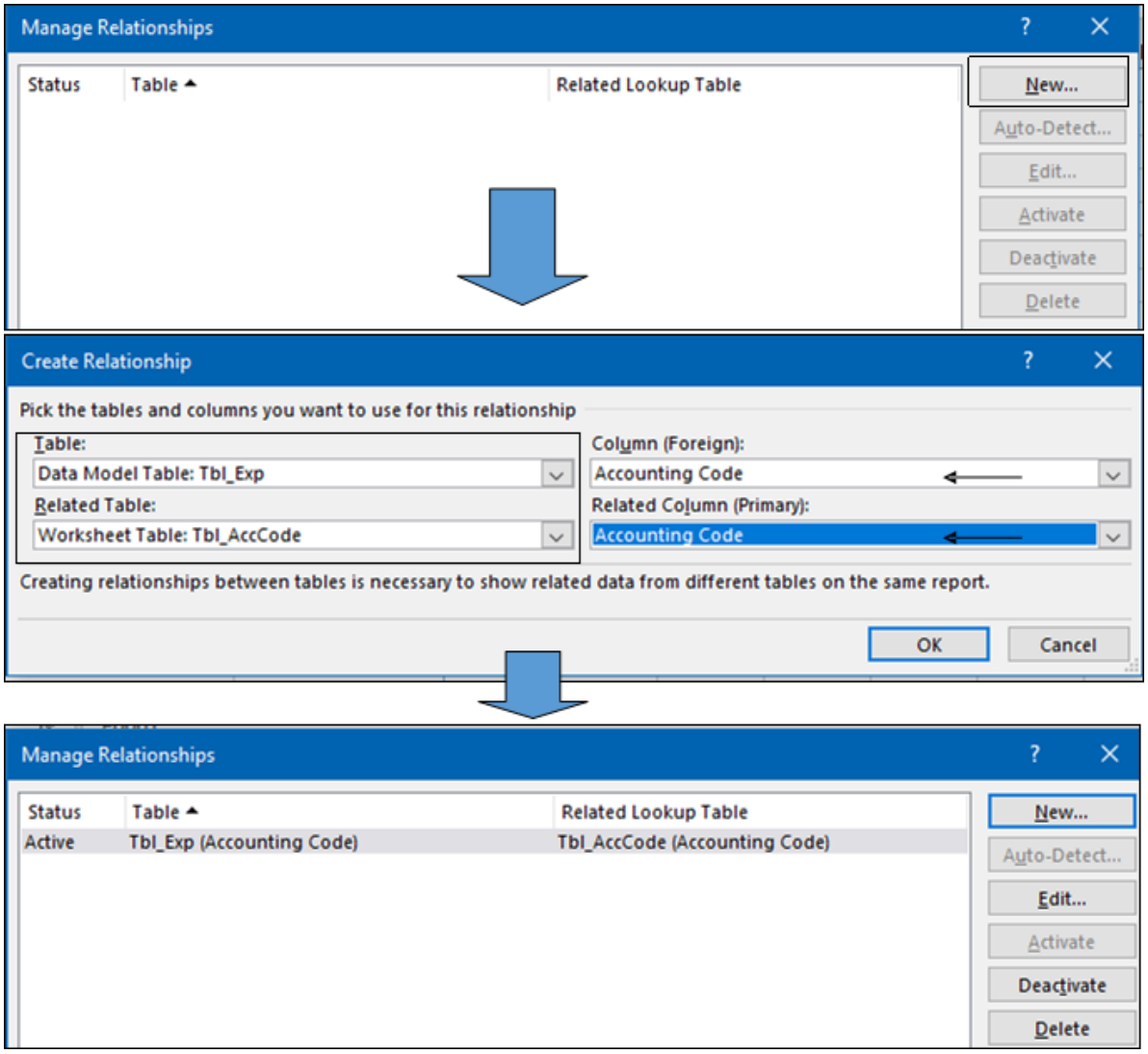


Figure 6.4.65Create Relationship

7. Select any table and insert a PivotTable (Insert > Pivot table)
8. Check the "**Add this data to data model**" checkbox.
9. If you observe it also has a Facility to create Data Connections.

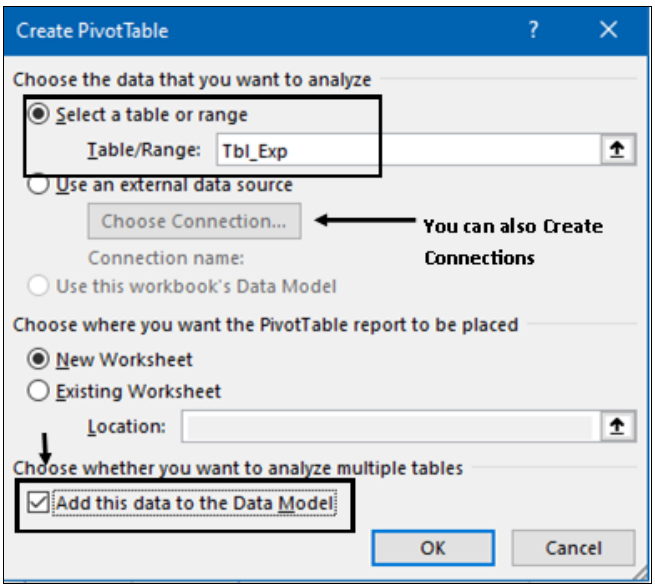


Figure 6.4.66 Crete Pivot with Add to Data Model

10. A **PivotTable** is created,

11. Drag Date to **Rows** and Debit to **Values**

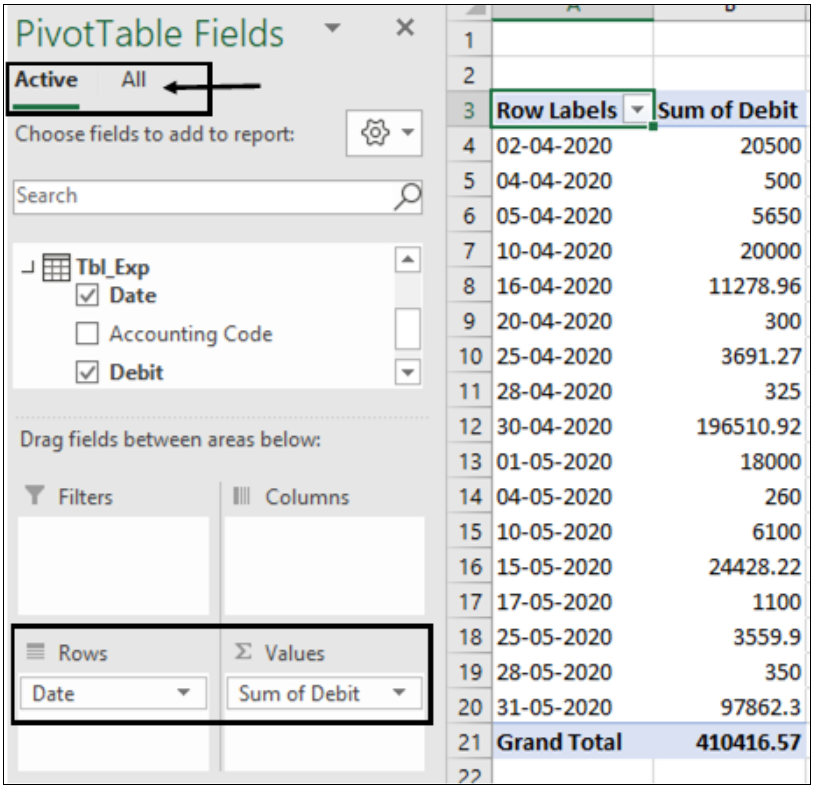


Figure 6.4.67Pivot table

12. In your pivot table field list, check “ALL” instead of “ACTIVE” to see all table names.

13. Select the field “Ledger name” from Tbl_Acc Code to **Column** create a combined pivot report.

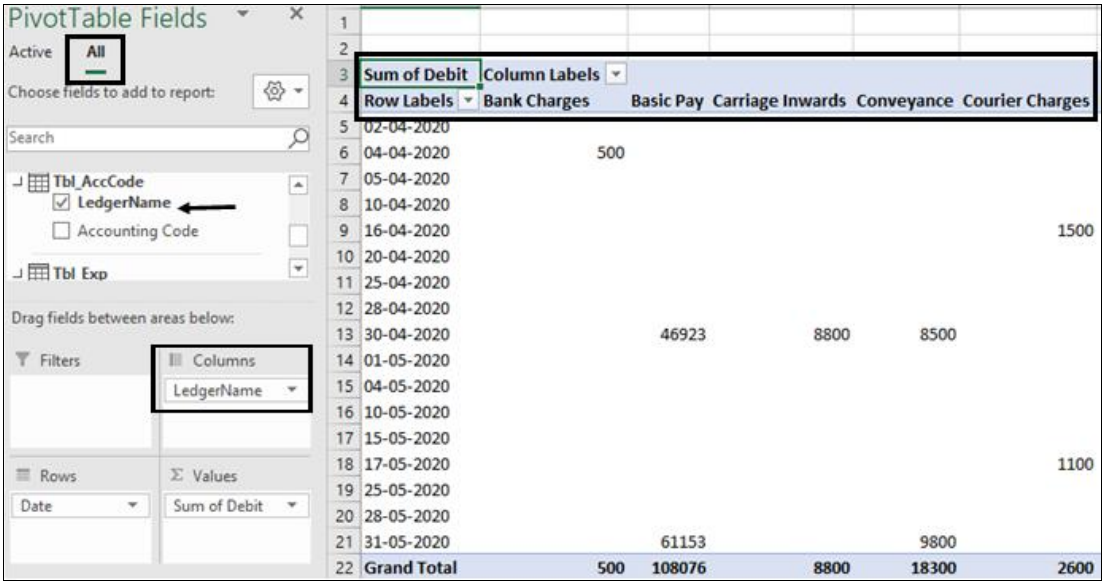


Figure 6.4.68Pivot with Ledger names

Gist: You have created a PivotTable using data Model out of two related Tables.

Commands learnt: PivotTables- Data Models.

6.4.11 POWER PIVOT AND POWER QUERY OVERVIEW

Power Query and Power Pivot complement each other.

Power Pivot is an efficient, powerful tool that comes with Excel as an Add-in. Power Pivot is an in-memory data modelling component that provides highly compressed data storage and extremely fast aggregation and calculation With Power Pivot, you can load hundreds of millions of rows of data from external sources and manage the data effectively. This makes it possible to perform the calculations, analyze the data, and arrive at a report to draw conclusions and decisions.

The difference between regular **PivotTables** and **Power Pivot PivotTables** is simply “power”,

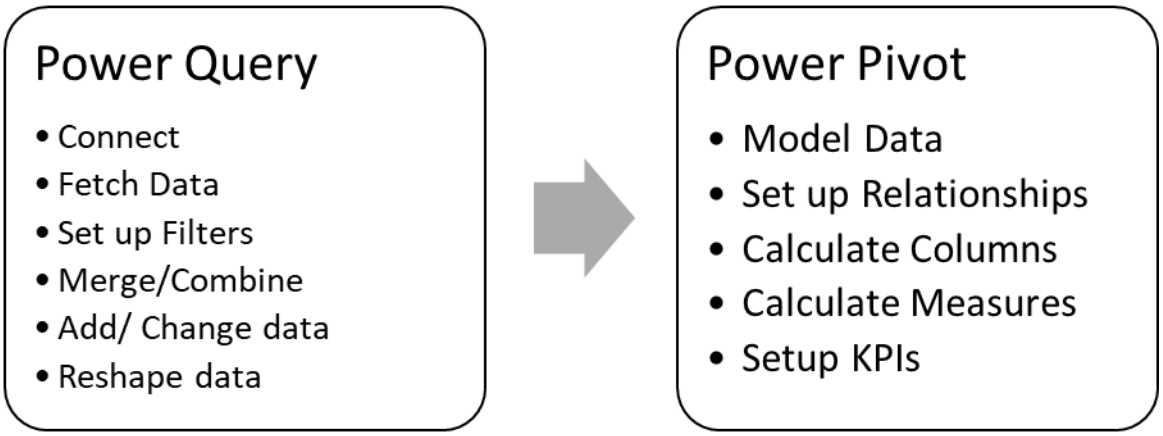
Power Query is a data processing software. It is an ETL (Extract, Transform, Load) self-service tool that works like an Excel add-in. It allows users to extract data from different sources, manipulate the specified data into a form that matches their needs, and load it into Excel.

You can use Power Query to

- Connect to several types of data sources
- Bring and combine data that is it can append, merge, join from various places

- Derive new columns of data
- Format, remove alternately reduce data
- Transform data like transpose, group, pivot, un-pivot, etc
- Write formulas to do advanced manipulation of data

The output of the Power Query can go to Excel or Power BI (another Microsoft application). That is why Power Query is available in both applications.



Power View

Power View feature in Excel facilitates interactive data visualization, exploration, and presentation that encourages intuitive ad-hoc reporting. Through Power View, Huge data sets can be analysed on the fly using multiple visualizations. In Power View data visualizations work very dynamically, thereby facilitating ease of presentation of the data with a single Power View report.

Power View is based on the Data Model you have studied earlier.

The core of the Power View is in its various types of data visualizations that will enable you to portray the data, visualize, and explore, all in a dynamic way. You can handle large data sets spanning several thousands of data on the fly switching from one visualization to another, drilling up and drilling down the data displaying the essence of the data.

You can have different Power View visualisations:

- Table
- Matrix
- Card
    - Charts
    - Line Chart

- Bar Chart
- Column Chart
- Scatter Chart
- Bubble Chart

- Map

Power Pivot is be a very powerful feature in Excel and is particularly useful.

- If Excel files are big and slow. Excel cannot cope with the amount of data.
- Where data is so huge, it exceeds the 1,048,576 rows available in a worksheet.
- If data is in external databases and you must write queries, to get at the data and start analysing it.
- The Data is spread over multiple tables and for relationship you ae using lot of formulas like VLOOKUP and INDEX & MATCH to bring the data into one table so that it can be analysed using PivotTables.
- PivotTable Calculated Fields and Items cannot do the complicated calculations you need.
- Unlike regular Excel PivotTables, Power Pivot can store multiple tables of data in the one model.
- You simply create relationships between those tables like you would with Access or other relational databases

Power Pivot can collate data from various data sources to perform the required calculations.

Power Pivot uses the Data Model

The power of the Power Pivot lies in its use of the Data Model. The data is stored in the form of data tables in the Data Model. You can create relationships between the data tables to combine the data from different data tables for analysis and reporting. You have in the previous section understood about Data Models

Memory Optimization

Power Pivot Data Model uses special storage, which is highly compressed when data is loaded into memory that makes it possible to store hundreds of millions of rows in memory.

You can load data directly into the Data Model thereby having an efficient highly compressed form.

Compact File Size

In a Data Model, it occupies very little space on the hard disk. Through the data Model, you can decrease file size sometimes to 10% thereby increasing the speed.

Power PivotTables

You can create the Power PivotTables from the Power Pivot window. The PivotTables so created are based on the data tables in the Data Model, making it possible to combine data from the related tables for analysis and reporting.

Power Pivot Charts

You can create the Power Pivot Charts from the Power Pivot window. The Pivot Charts so created are based on the data tables in the Data Model, making it possible to combine data from the related tables for analysis and reporting.

DAX Language

The strength of Power Pivot also comes from the DAX Language, it is somewhat like excel functions whereas excel functions work on cells DAX works on Columns. This can therefore be used effectively on the Data Model to perform calculations on the data in the data tables. You can have Calculated Columns and Calculated Fields defined by DAX that can be used in the Power PivotTables and Power Pivot Charts.

CASE STUDY: As a part of the investigation your principal CA Chandiwala has given you huge expense data file in CSV format with around 257499 rows of data. In the CSV file the field Account_Idis not formatted properly and has some unnecessary spaces. He gives you another file with Table containing Account_ID masters, which has Expenditure. He wants you to find a smart way to build analyse this huge data on various parameters.

	A	B	C
1	Date	Account_ID	Amount
2	28-12-2019	1	1310.6
3	18-01-2019	7	6754.3
4	12-10-2019	12	7082.4
5	12-12-2019	1	748.9
6	09-11-2019	6	609.1
7	24-01-2019	7	5119
257492	18-08-2019	2	496.5
257493	10-12-2019	2	16682.9
257494	16-11-2019	1	11233.6
257495	19-11-2019	7	7394.2
257496	08-11-2019	6	2639.5
257497	30-05-2019	5	6622
257498	19-04-2019	1	13480.3
257499	28-04-2019	2	9533.1

	A	B
1	Account_ID	Expenditure
2	1	Advertising Expenses
3	2	Bank Charges
4	3	Basic Pay
5	4	Business Promotion Expenses
6	5	Carriage Inwards
7	6	Conveyance
8	7	Courier Charges
9	8	Entertainment Expenses
10	9	Gifts
11	10	Maintenance Charges
12	11	Postage & Telegraph
13	12	Printing & Stationery
14	13	Repairs & Maintenance
15	14	Telephone Expenses
16	15	Transportation Expns
17	16	Travel Expenses

Figure 6.4.69 Expense.CSV and Accounting Heads.XLSX

Strategy:

You can achieve your objective using Power Query and Power Pivots

1. Open a new Excel sheet
2. Go to **Data tab > Get Data> From File > From Text/CSV** as shown in Fig 6.4.70

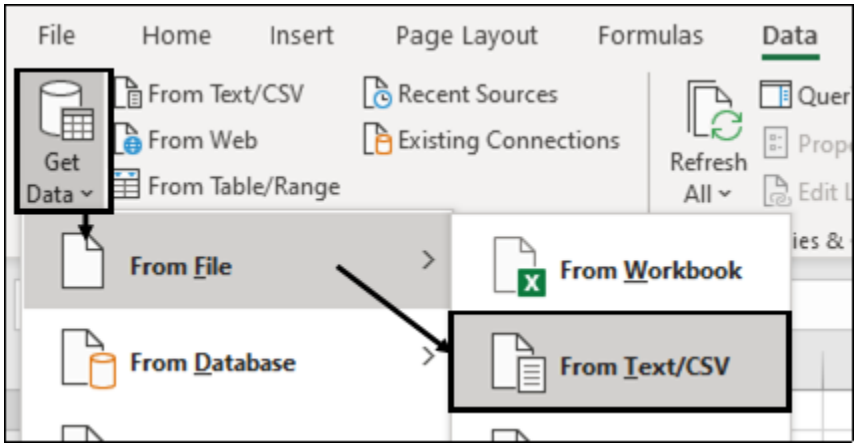


Figure 6.4.70 Get data from CSV

3. A Data Query Preview window will Open

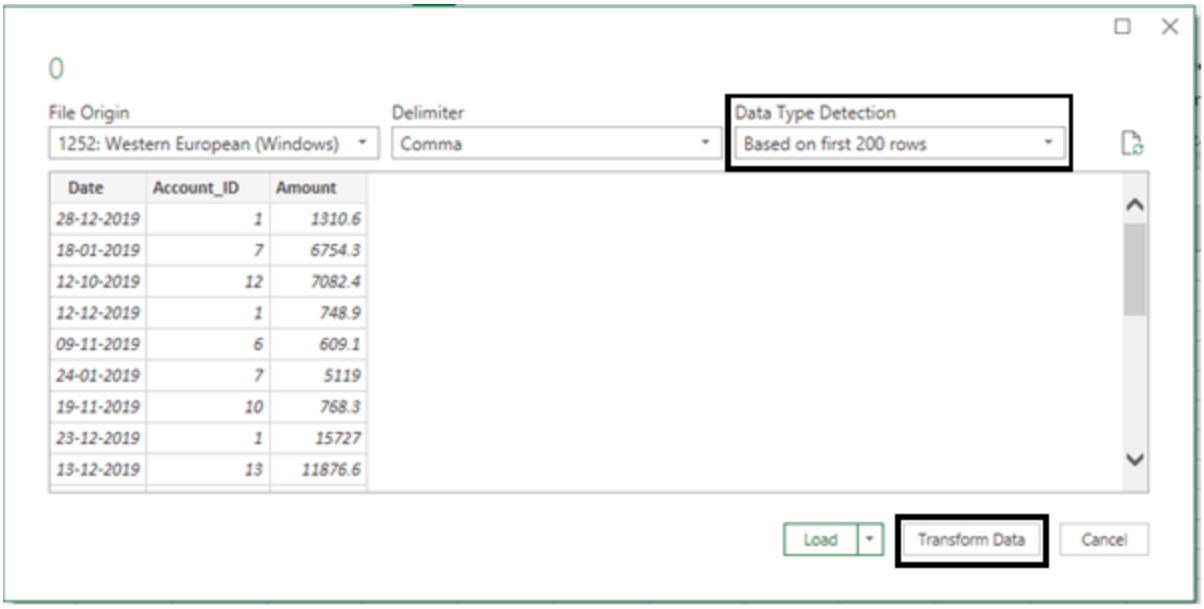


Figure 6.4.71 Preview Window

4. Select **Transform data**
5. **Power Query Editor** will open
6. Some steps it has automatically done, like:
   - Promoted Headers
   - Cleaned Spaced in Account_ID
   - Converted AccounT_ID to Numeric

7. You can also see it has a Different Language as Compared to Excel

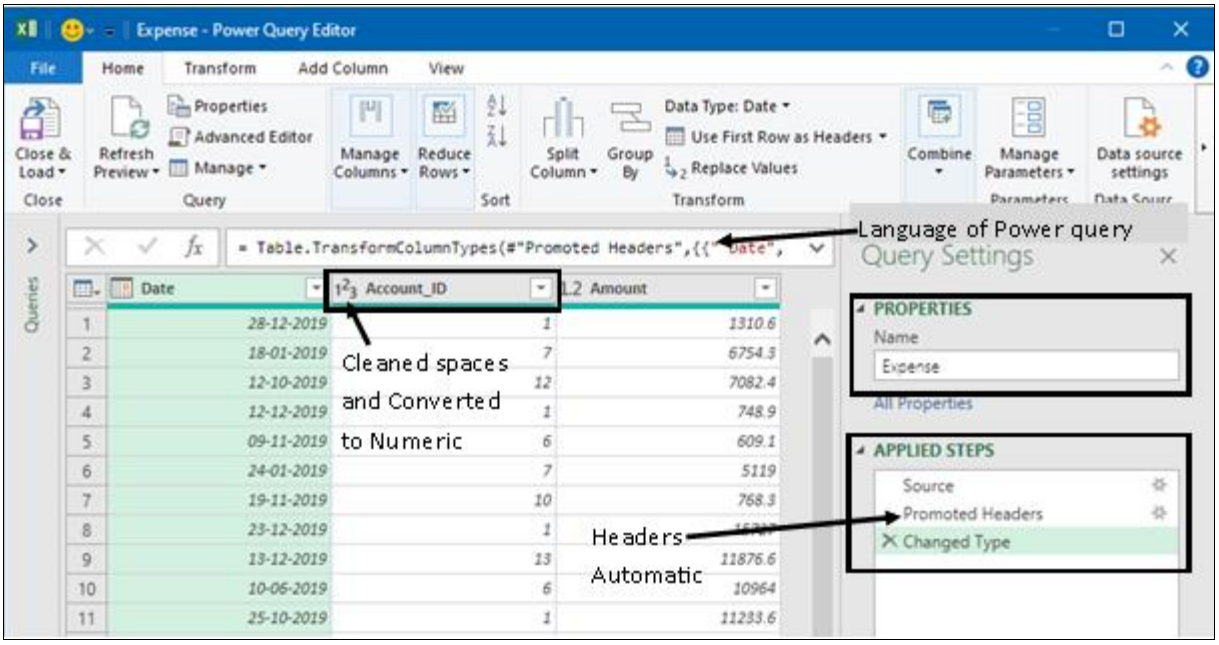


Figure 6.4.72 Power Query Editor

8. You want to do a datewise analysis, You can add a Column based on Day of the week.

9. In Power Query Formulas work on whole Column rather than Cells.

10. Select the **Add Column tab>Date>Day>Name of Day** (There are many more options)

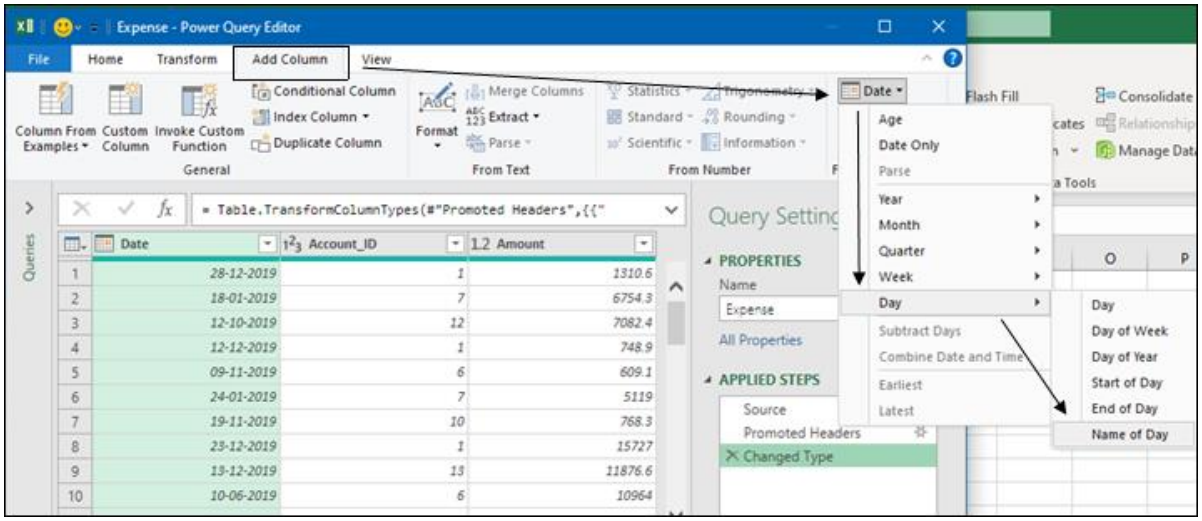


Figure 6.4.73 Add Column

11. A column with Weekday will be inserted

12. A step is also automatically inserted, You can undo steps whenever you want

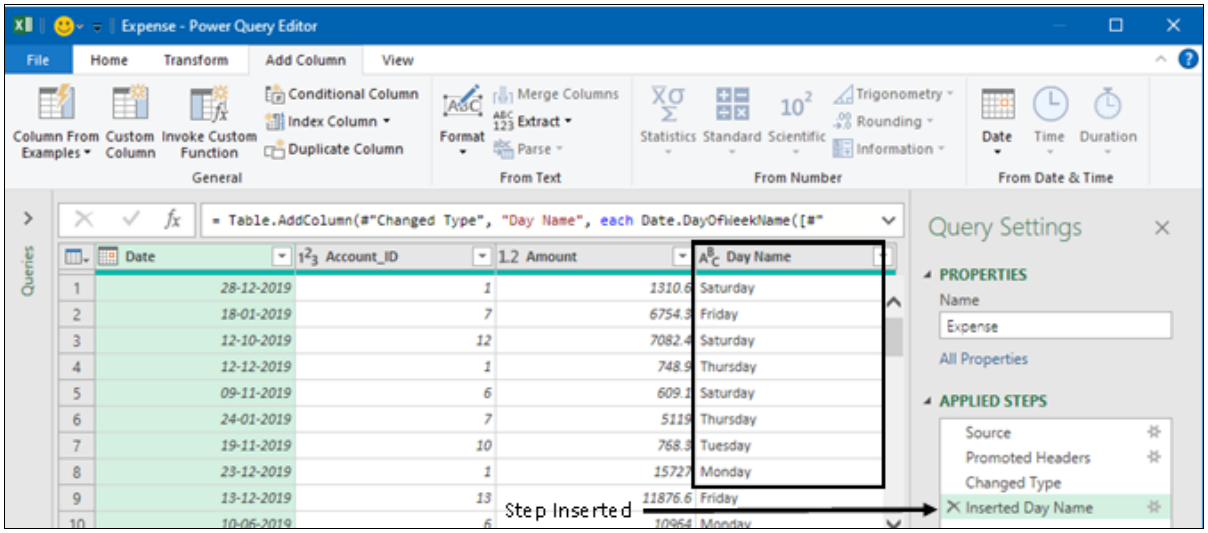


Figure 6.4.74 weekday Column Inserted

13. Similarly, you will add a Column for **Name of Month** also

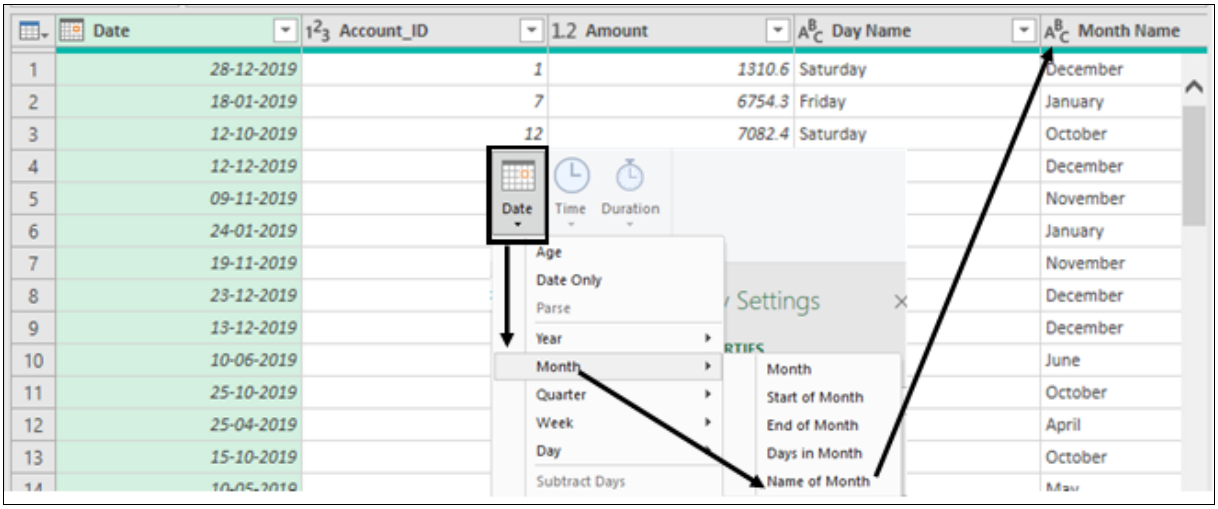


Figure 6.4.75 Name of Month Column Inserted

14. You have seen that Transforming, Cleaning is very easy with Power Query with just a couple of clicks.

15. In the Home Tab, you have the option to load to an Excel or create a Connection.

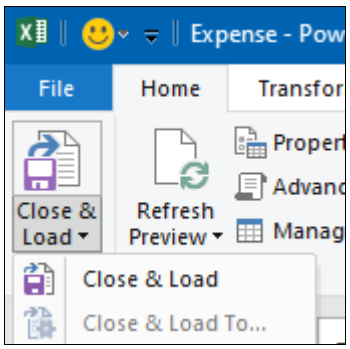


Figure 6.4.76 Close and Load

16. Once you Load File as a Connection it will appear as a connection in the data Tab> queries and Connections.

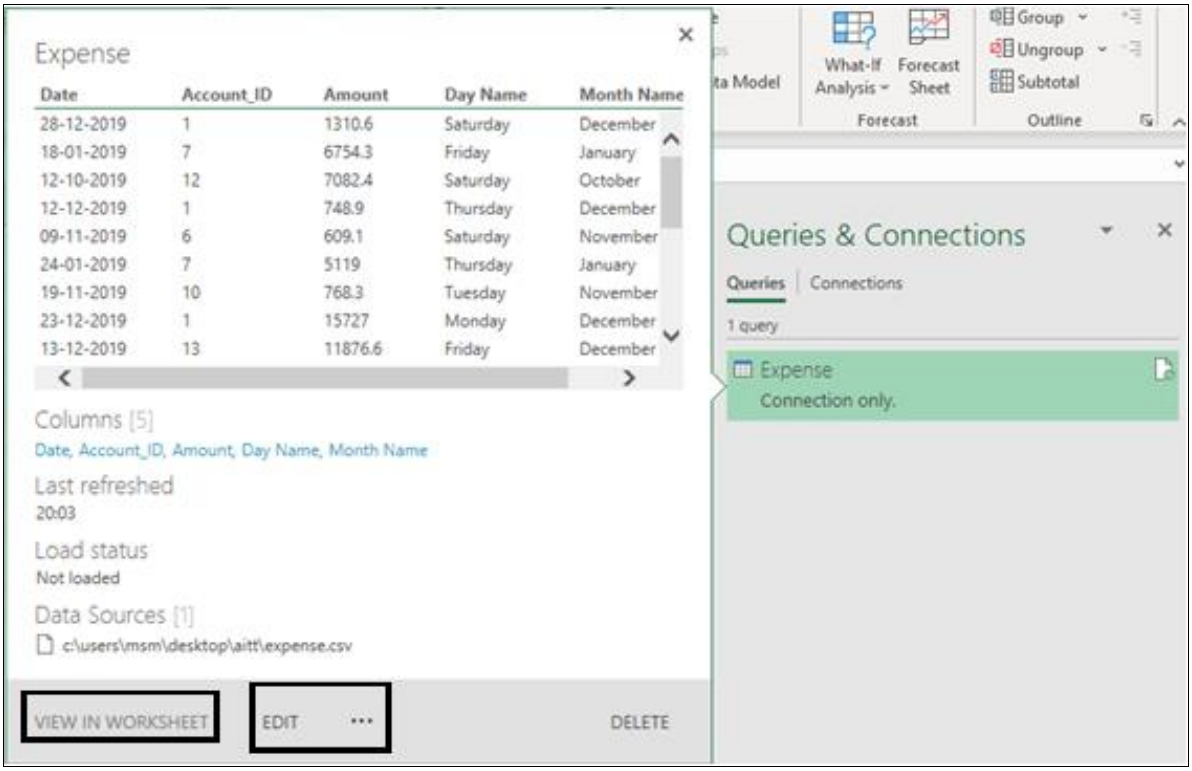


Figure 6.4.77 Queries in Data Tab as a Connection

17. You Right Click on **Connection** and select **Load To**
18. **Import Data** Dialog Box will appear as shown in Figure
19. Select **Only Create Connection** and **Add this to data Model.**
20. Your "Expense" sheet is now Loaded onto the Model.

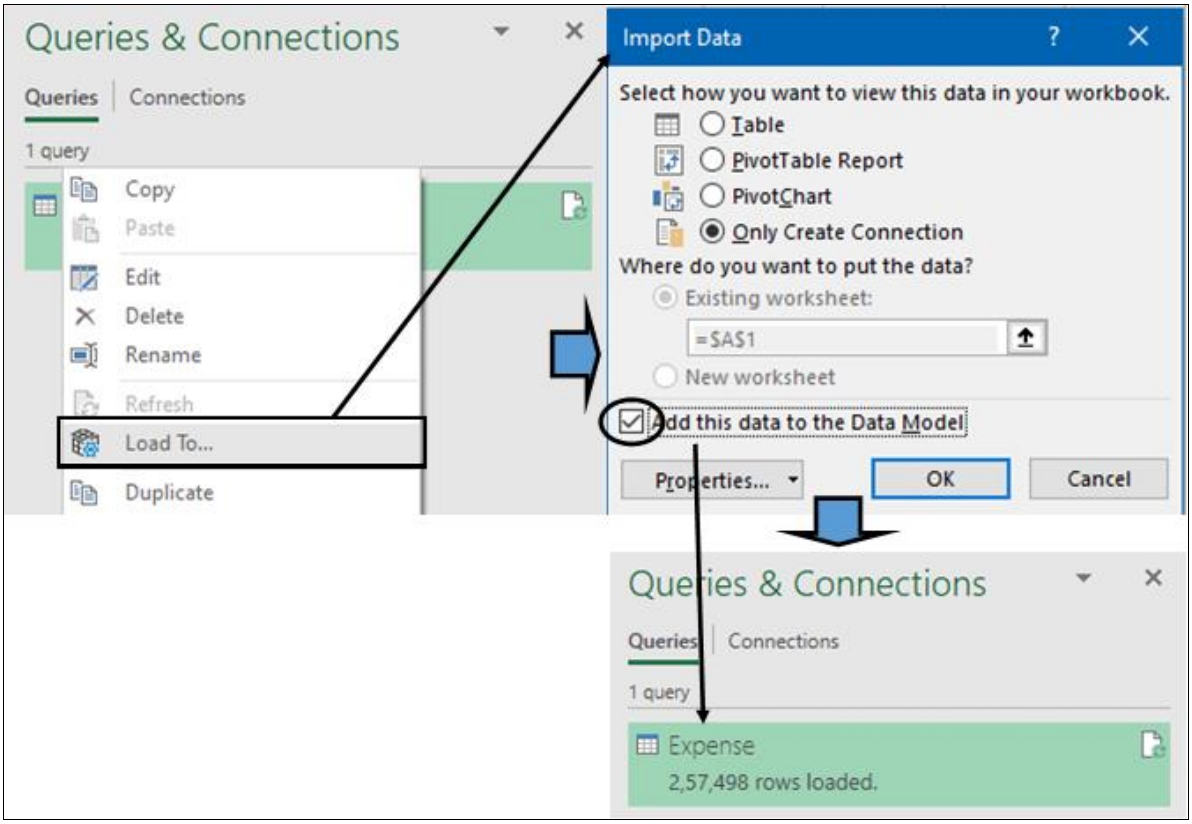


Figure 6.4.78 Loading Connection to Model

21. Click on the **POWER PIVOT** tab on the Ribbon.
22. Click on **Manage** in the **Data Model** group.
23. Power Pivot window appears, in Data View.
24. Power Pivot has two views −
    - Data View − It displays all the tables in the Data Model with fields displayed in columns and data as records in the rows, with a calculation area below each table. The table tabs look similar to the Excel worksheet tabs with names.
    - Diagram View − It displays all the tables as boxes with table names as caption and the fields listed in the box. You can drag the tables to align them, resize them to make all the fields visible, and create relationships by just clicking on the fields and connecting them with lines.
25. You can see all the 257498 Records have been loaded.

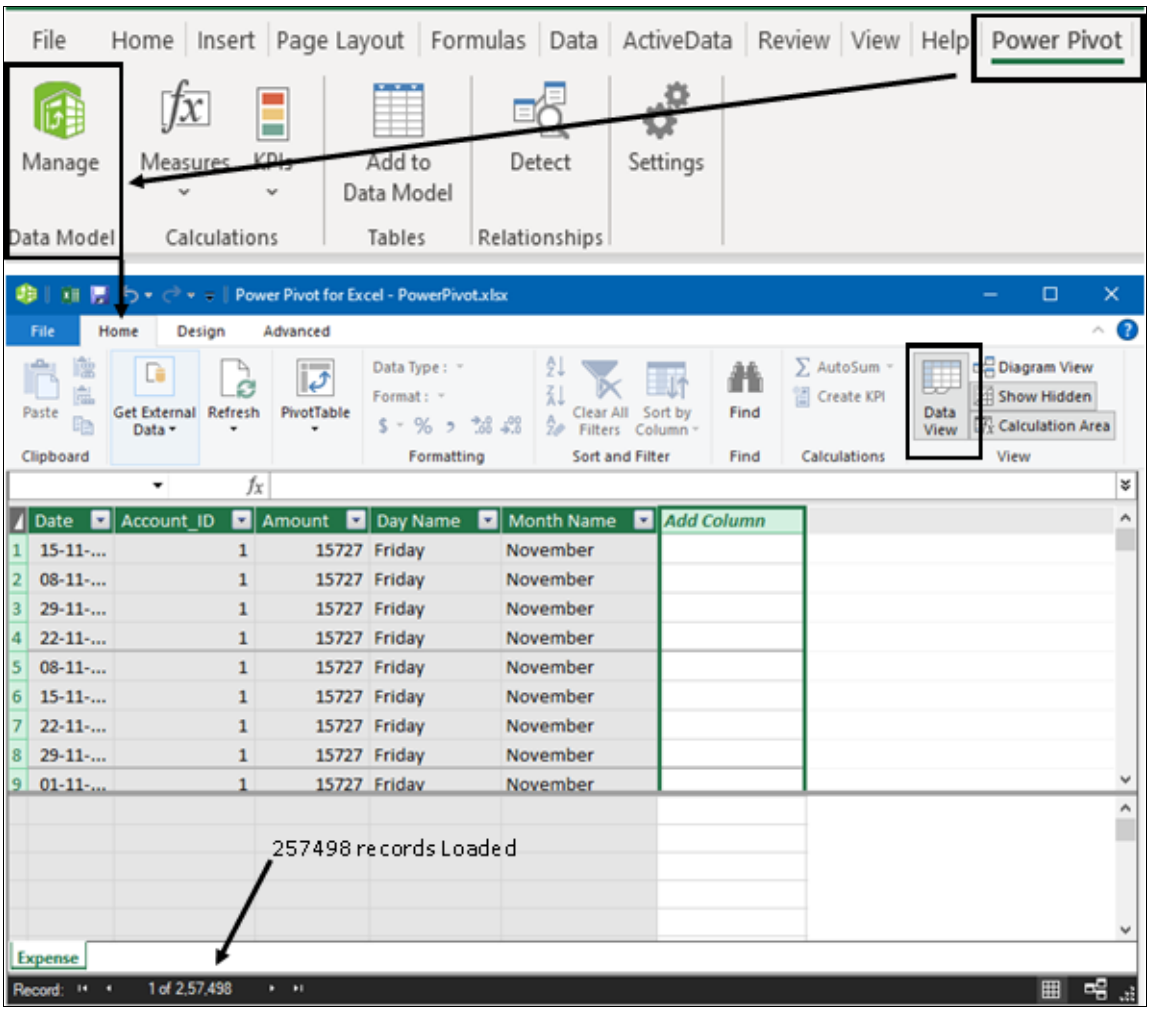


Figure 6.4.79 Power Pivot Data Model

26. Next, You want to import **Account_Id** master data with corresponding Expense Heads.
27. Click on Get External data > From other Sources
28. **Table Import Wizard** Dialog Box Appears
29. Select **Excel file**
30. In the next step of **Table Import Wizard** give the path to the Excel file.
31. Select **Use first Row as Column headers.**
32. Click Next
33. Select the Table to be imported **Acc_ID**
34. Click Next
35. The Table is also Imported to Data Model.

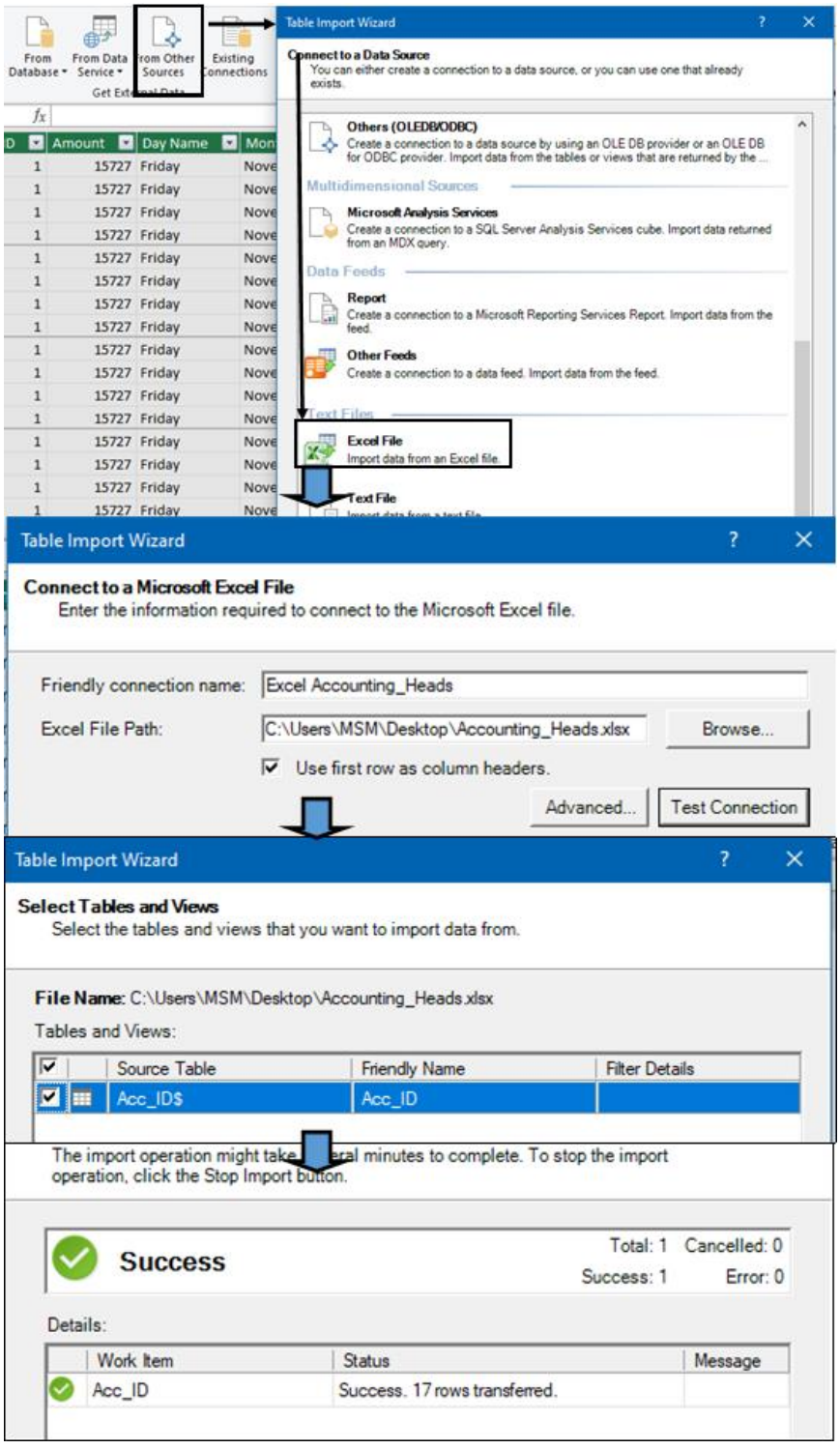


Figure 6.4.80 Import Table to data Model

	Account_ID	Expenditure
1	1	Advertising Ex...
2	2	Bank Charges
3	3	Basic Pay
4	4	Business Prom...
5	5	Carriage Inwards
6	6	Conveyance
7	7	Courier Charges
8	8	Entertainment...
9	9	Gifts
10	10	Maintenance ...
11	11	Postage & Tele...
12	12	Printing & Stati...
13	13	Repairs & Mai...
14	14	Telephone Exp...
15	15	Transportation...
16	16	Travel Expenses

Figure 6.4.81 Master Added to data Model

36. You can use **the data from different tables for analysis** and reporting only when relationships exist among them.
37. You can **view the relationships** between tables from the **diagram view** in the Power Pivot window.
38. Click **Diagram View** in the View group.
39. Resize the diagram using the scroll bar so that you can see all the tables in the Data Model in the diagram.
40. All the tables in the Data Model appear with their field lists. The relationships among the tables are denoted by the lines connecting them.
41. You can build a relationship by dragging between related fields.

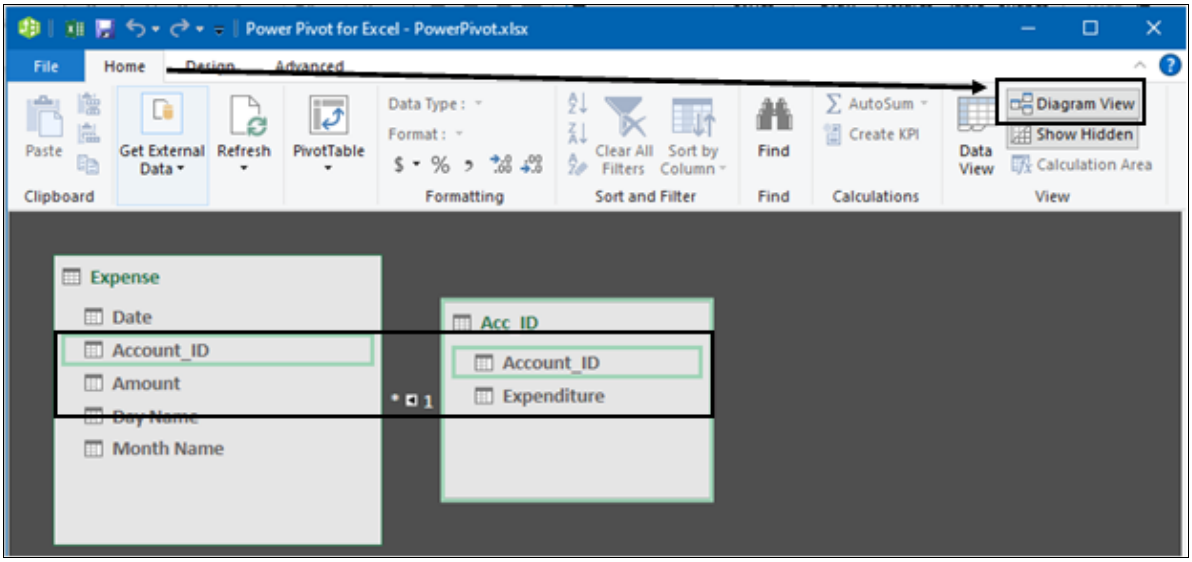


Figure 6.4.82 Diagram View

42. Next Click on the PivotTable option on Home Tab
43. Select PivotTable
44. Create PivotTable dialog Box will Open
45. Select Location where you want to create a Pivot

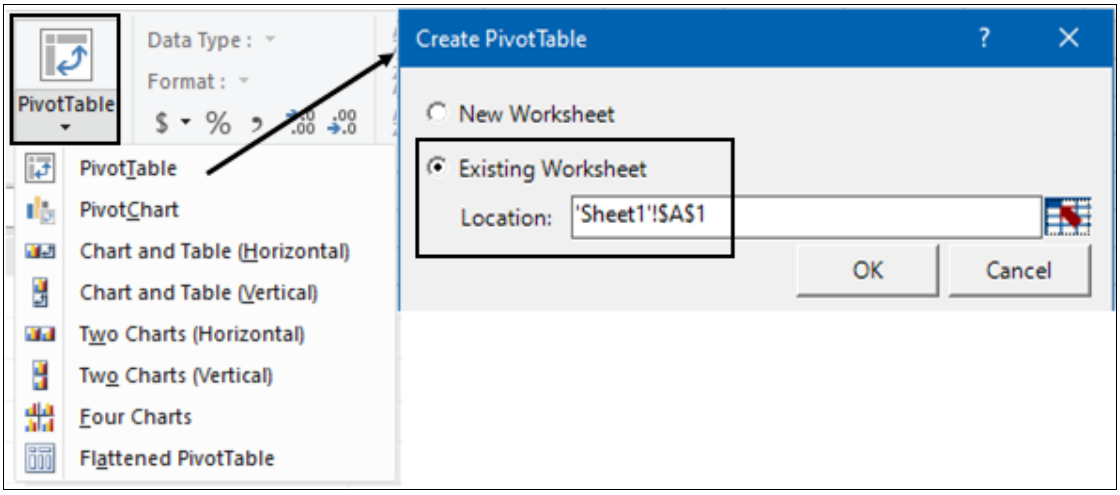


Figure 6.4.83 Create PivotTable

46. A PivotTable is created and now you can see Two databases in Field List.
47. You have the option to Drag fields from any of the databases since Power Pivot has created a relationship.
48. Drag Expenditure From Acc_ID database to Rows

49. Drag Amount from Expense database to Values

50. A Pivot Table is created as shown in Figure

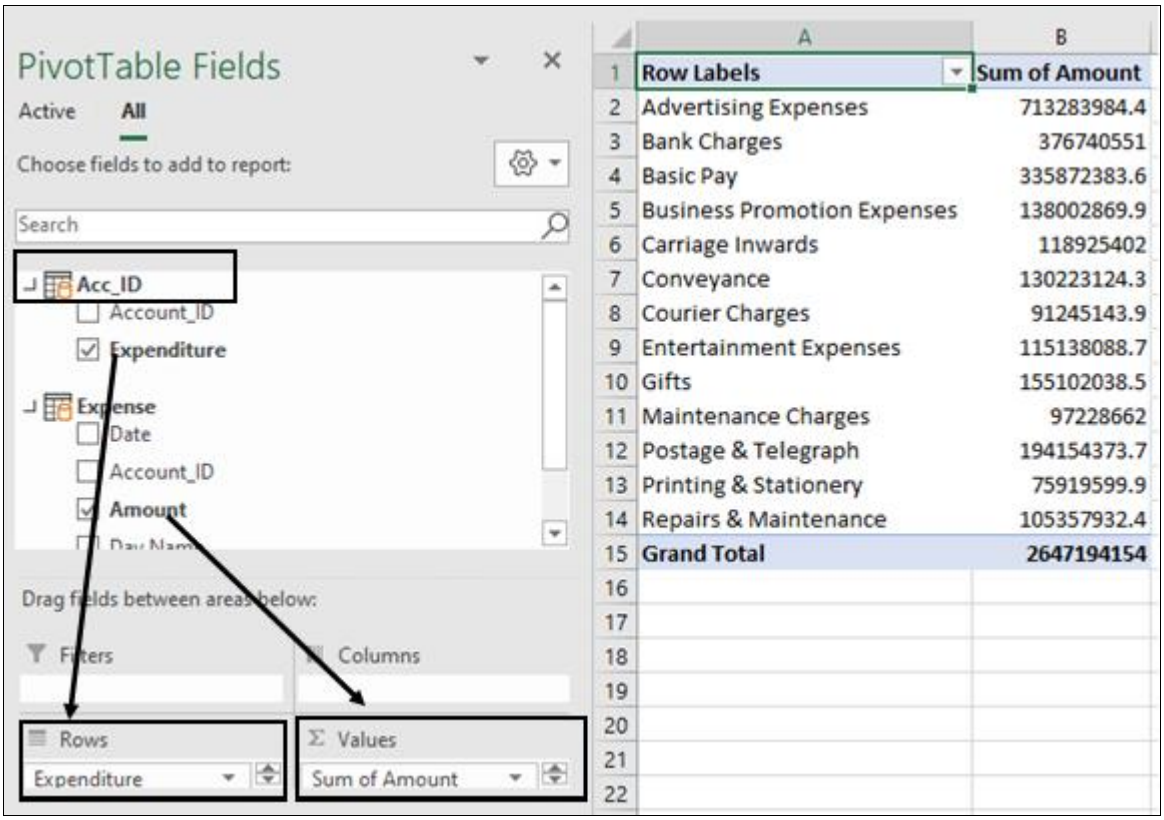
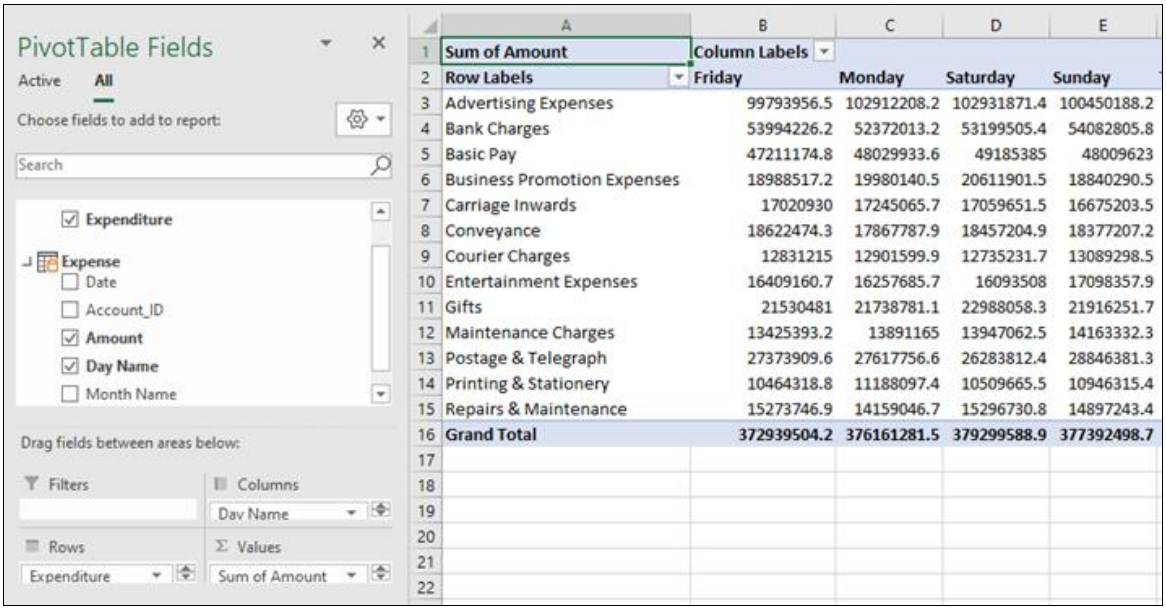


Figure 6.4.83 Pivot table created

51. You want To analyse by Weekday

52. Drag “Day Name” to Columns,

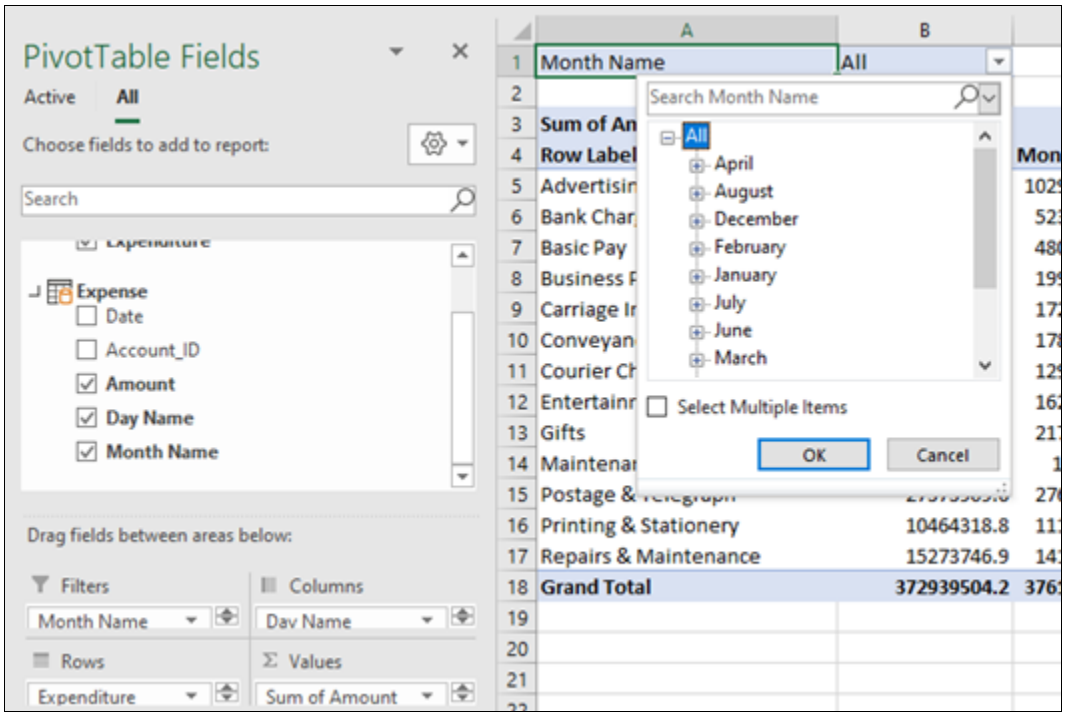
53. You get an Expensewise, Weekdaywise analysis.



	A	B	C	D	E
1	Sum of Amount	Column Labels			
2	Row Labels	Friday	Monday	Saturday	Sunday
3	Advertising Expenses	99793956.5	102912208.2	102931871.4	100450188.2
4	Bank Charges	53994226.2	52372013.2	53199505.4	54082805.8
5	Basic Pay	47211174.8	48029933.6	49185385	48009623
6	Business Promotion Expenses	18988517.2	19980140.5	20611901.5	18840290.5
7	Carriage Inwards	17020930	17245065.7	17059651.5	16675203.5
8	Conveyance	18622474.3	17867787.9	18457204.9	18377207.2
9	Courier Charges	12831215	12901599.9	12735231.7	13089298.5
10	Entertainment Expenses	16409160.7	16257685.7	16093508	17098357.9
11	Gifts	21530481	21738781.1	22988058.3	21916251.7
12	Maintenance Charges	13425393.2	13891165	13947062.5	14163332.3
13	Postage & Telegraph	27373909.6	27617756.6	26283812.4	28846381.3
14	Printing & Stationery	10464318.8	11188097.4	10509665.5	10946315.4
15	Repairs & Maintenance	15273746.9	14159046.7	15296730.8	14897243.4
16	Grand Total	372939504.2	376161281.5	379299588.9	377392498.7
17					
18					
19					
20					
21					
22					

Figure 6.4.84 Expensewise, Weekdaywise pivot

54. You can insert Report Filters by Month Name as shown in Figure



	A	B
1	Month Name	All
2		
3	Sum of An	
4	Row Label	Mon
5	Advertisin	102
6	Bank Char	52
7	Basic Pay	48
8	Business P	19
9	Carriage In	17
10	Conveyan	17
11	Courier Ch	12
12	Entertainr	16
13	Gifts	21
14	Maintenar	1
15	Postage &	27
16	Printing & Stationery	10464318.8
17	Repairs & Maintenance	15273746.9
18	Grand Total	372939504.2
19		
20		
21		

Figure 6.4.85 Filter by Month

55. You have analysed by various parameters in Power Pivots

You must have realised the power of Power Pivots.

Gist: You have used the Power Query and Power Pivot to analyse Huge data based on different parameters.

Commands learnt: PivotTables- Data Models.

6.5 SUMMARY

Visualisation is a powerful and efficient way to analyse and understand data. You have understood the use of Conditional Formatting to format cells based on condition thereby isolating Exceptions.

Through the use of PivotTables, you can examine the data for similarities, differences, highs, and lows. By making a few quick changes to the pivot table, you can see your data from a completely different angle.

You have in this chapter learnt to Create and Modify PivotTables. Sorting and filtering Pivot Tables helps you in seeing the small picture out of the broader view of the data. The data can be grouped based on values, Dates, headings, etc. Using Slicers and Time Lines Pivots can be filtered easily and interactively. Charts give a visual Feel to the data and you can see trends easily. PivotTables,

To get a visual Dynamic picture of data you have learnt build Dashboards. Power Query and Power Pivots are very powerful tools in arsenal of Excel.

REFERENCES

[1] Michael Alexander, Richard Kusleika, John Walkenbach, 'Excel 2019 Bible', Wiley,2018

[2] Greg Harvey, 'Excel 2019 All-in-One for Dummies, John Wiley & Sons,2018

[3] www.chandoo.org

[4] www.contextures.com

CHAPTER

7 PROTECTION AND DATA TOOLS IN EXCEL

LEARNING OBJECTIVES

- Understanding Protection of Cell, Worksheet
- Understanding Protection of a shared workbook
- Gain an understanding of Get and Transform Feature
- Understand the conversion of Text to Columns through Data Tool
- Understanding Data Consolidation
- Understand Forecasting through What-if -Analysis
- Understanding the use of Solver for optimisation problems

7.1 INTRODUCTION

The data can be protected by simply locking it down, preventing anyone from changing it.

Excel has Workbook protection features to help reduce errors that prevent accidental or intentional modification of data. Using them, you can:

- Prevent people from changing the structure of a worksheet (inserting or deleting cells, columns, or rows).
- Prevent people from changing a worksheet's formatting (including the number format or other formatting details like column width and cell color).
- Prevent people from editing certain cells.
- Prevent people from entering data in a cell unless it meets certain criteria.
- Provide additional information about a cell in a pop-up tip box.
- Prevent people from editing or even seeing the spreadsheet's formulas.
- Prevent people from moving to cells they do not need to edit or inspect.

The reasons for putting in protection could be many ranging from, you don't want other people to tamper with data or to prevent errors especially when the workbooks are shared with data validation, you can lock out certain types of errors.

You can lock down individual cells, a range of cells, or an entire worksheet to prevent anyone from changing its data. You can also protect a whole workbook, to prevent changes to its structure, such as adding worksheets or changing the workbook window's size. Finally, when needed, you can prevent unauthorized users from even opening a workbook at all.

Often the data you want to analyse is in a different format and present in an external application and sources, Get and transform also called Power Query gives you all the facilities to connect to various source Data and transform it in a form, making it suitable for analysis.

There are many Data Tools grouped under the Data Tab on Ribbon. Text to Columns is a fantastic Tool to split data in one column to different columns based on a delimiter or fixed length so that some meaningful analysis can be made. Consolidation is another tool under Data Tools that facilitates consolidating data from different ranges that may be on the same sheet or in separate sheets.

What-if-analysis in Excel is a part of forecasting tools in Excel that aids you to run reverse calculations, perform sensitivity analysis, and compare various scenarios.

Using these tools you can make informed decisions, which are based on data, the outcome of the business or project or task is always more in control. Decision making is an important task for any professional or manager, therefore, What-if analysis in Excel is used by almost every management professionals, to make better, faster, and more accurate decisions based on data. Goal Seek, Data Tables, and Scenario Manager are very helpful in doing reverse Calculations doing sensitivity analysis, and comparing scenarios. The solver is an add-in that provides. optimisation solutions.

7.2 PROTECTION FEATURES

7.2.1 PROTECT AN EXCEL WORKBOOK WITH PASSWORD

You can prevent other users from viewing your Excel files. For this, you can encrypt your workbook with a password that will prevent other people from opening it.

In the workbook, you want to encrypt,

- Click File > I**nfo**,
- Click the **Protect Workbook** button,
- Select **Encrypt with Password** in the drop-down menu.

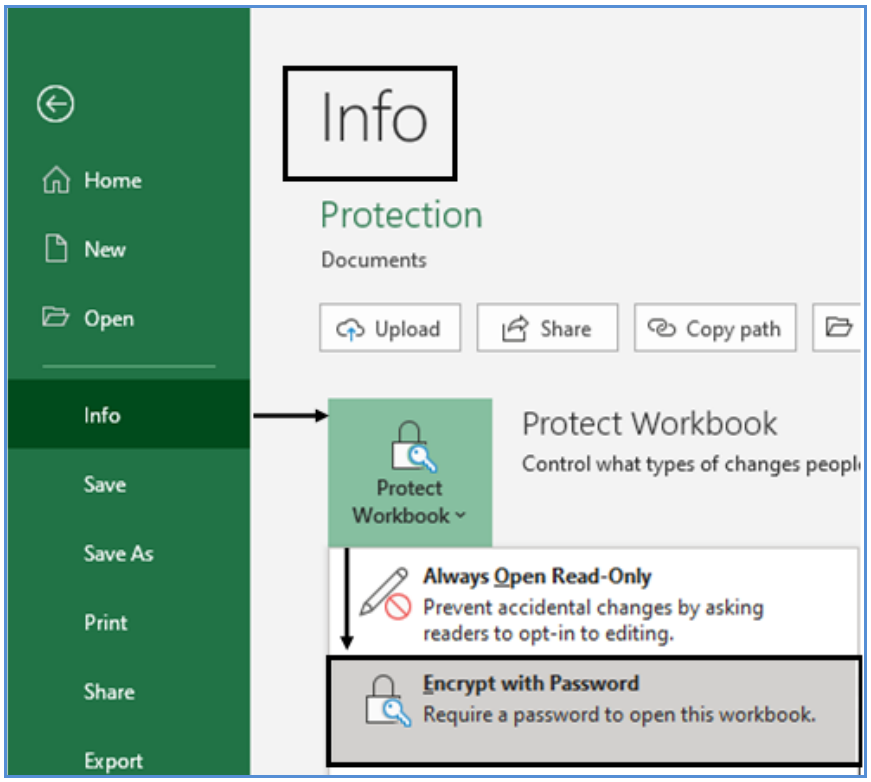


Figure 7.2.1 Select Encrypt with Password

- The **Encrypt Document** Dialog Box will appear,
- You **enter a password** in the Password box,
  - By default, Microsoft Excel uses AES 128-bit advanced encryption.
  - You can supply a password of up to 255 characters.
- Click OK.
- **Confirm Password** Dialog Box will appear
- Retype the password in the Re-enter password box,
- Click OK.

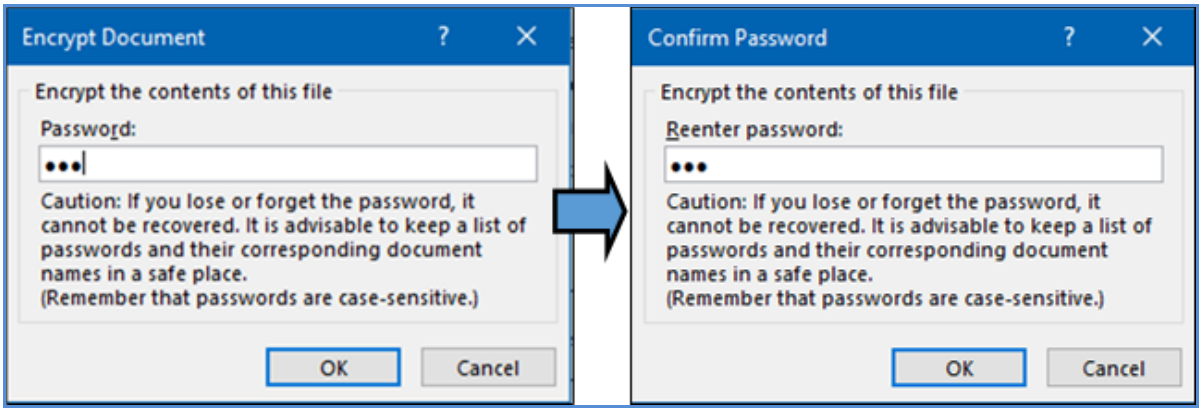


Figure 7.2.2 Encrypt Document Dialog Box

- Save your workbook.

To remove the workbook encryption, repeat the above steps, delete the password from the Password box, click OK, and then save your workbook.

PASSWORD PROTECT OPENING/ MODIFYING EXCEL WORKBOOKS

There is another method through which you can both prevent others from opening your workbook, alternately you can configure, so that users may open the Workbook but not modify the workbook.

In your Excel workbook,

- Click the **Save As** button,
- Click the **Tools** button at the bottom of the window.
- In the Tools drop-down menu, select **General Options**.
- In the General Options dialog window, type a password in the Password to **Open/ modify box**.
- You can also select the **Read-only** option, in case you want someone to open the file as read-only.
- Retype the password in the Confirm Password dialog,
- Click OK.
- If you've entered 2 passwords you will have to confirm both.

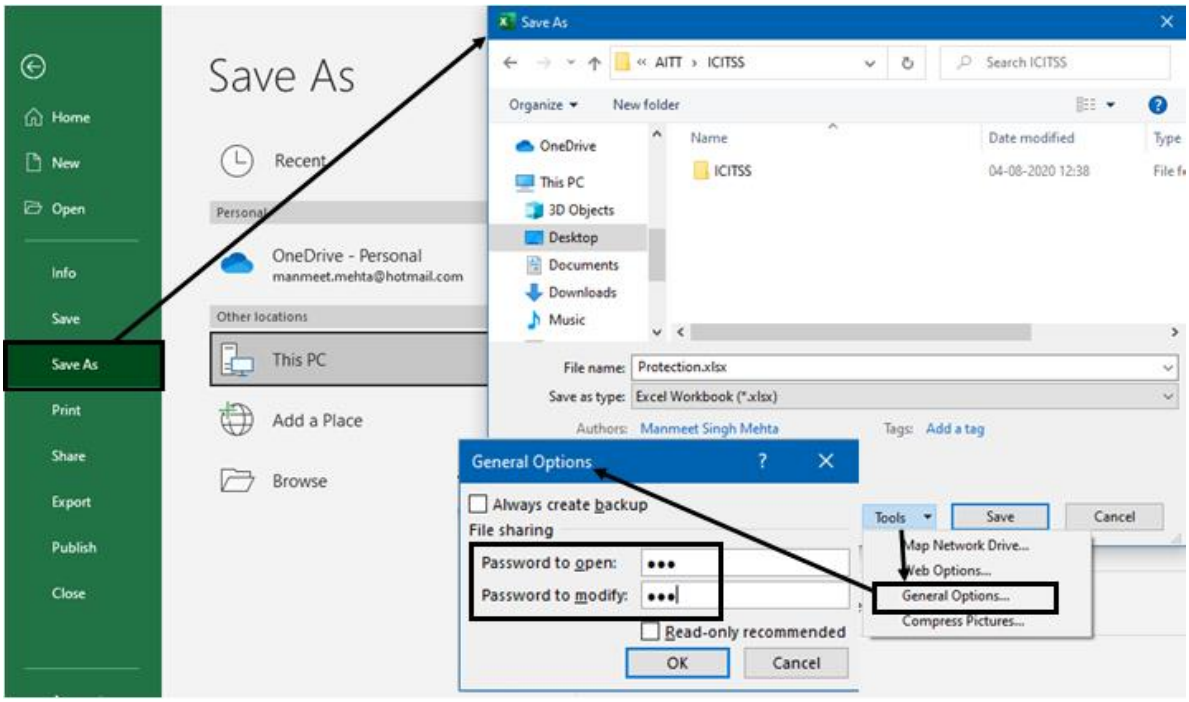


Figure 7.2.3 Password protect to Open/ modify

- Click the **Save** button in the **Save As** dialog box.

7.2.2 PROTECT A SHARED WORKBOOK

You can also protect an entire workbook. If protected, Excel prevents people from inserting, moving, or removing worksheets.

Workbook protection works hand-in-hand with worksheet protection. If you use workbook protection but not worksheet protection, people can still edit all the cells in your worksheets. However, they can't delete the worksheets or add new ones. On the other hand, if you use workbook protection and worksheet protection, people can't tamper with your data or the structure of your workbook.

To protect the workbook.

- Select Review>Changes>Protect Workbook

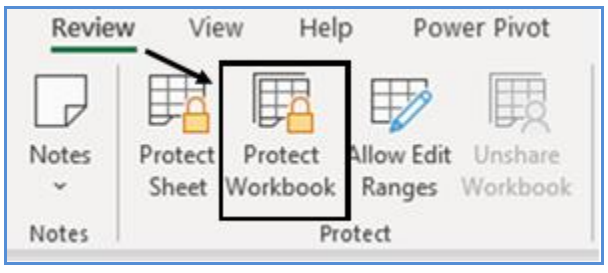


Figure 7.2.4 Protect Workbook under Review Tab

- Protect Structure and Windows Dialog Box will appear

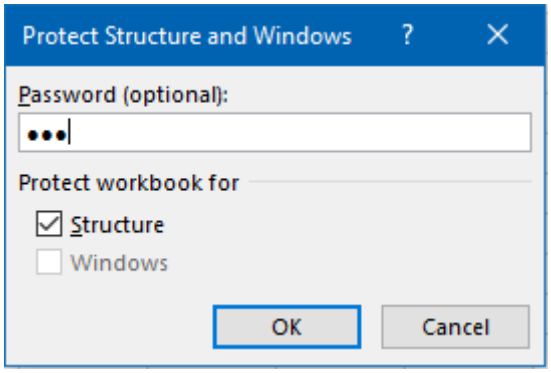


Figure 7.2.5 Protect structure and windows

- Under **Protect Structure and Windows** dialog box, you have two options:
- To protect the **structure** of a workbook, select the Structure checkbox.to protect from
  - Viewing worksheets that you have hidden.
  - Moving, deleting, hiding, or changing the names of worksheets.
  - Inserting new worksheets or chart sheets
  - Moving or copying worksheets to another workbook.
  - In PivotTable reports, displaying the source data for a cell in the data area, or displaying page field pages on separate worksheets.
  - For scenarios, creating a scenario summary report.
  - In the Analysis ToolPak, using the analysis tools that place results on a new worksheet.
- To keep **workbook windows** in the same size and position every time the workbook is opened, select the Windows checkbox. to protect from
  - Changing the size and position of the windows for the workbook when the workbook is opened.
  - Moving, resizing, or closing the windows.

  You can use these options to protect a workbook.
- In the Password box, type a password, and click the OK button.
- Retype the password in the Confirm Password window
- Click OK.

7.2.3 PROTECT CELL DATA

Excel's data validation tools help make sure any invalid data does not end up in the worksheets. But they do not protect the worksheets against things like accidentally deleted formulas, mistakenly scrambled formatting,

and unintentionally modified values. To defend against these dangers, you need to use Excel worksheet protection features.

How it works: each cell can have one of two special settings:

- **Locked**- When a cell is locked, you cannot edit it.
- **Hidden**- When a cell is hidden, its contents do not appear in the formula bar. The cell still appears in the worksheet, but if the cell uses a formula, it is not visible.

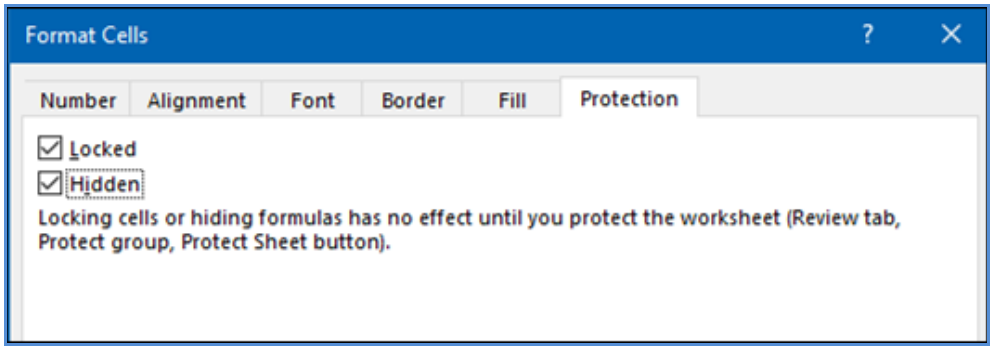


Figure 7.2.6 Locked and Hidden cell

You can use these settings *individually or together*. When a cell is **both locked and hidden**, you **can't edit** it or **view** it in the formula bar.

On the other hand, if a cell is **hidden but not locked**, people **can edit** the cell but can never tell whether the cell uses a formula because Excel keeps that information secret.

The most important thing you need to understand about locked and hidden cells is that **these settings come into effect only when the worksheet is protected**. If the worksheet isn't protected, Excel doesn't use these settings at all.

In other words, you need to take two steps to build a protected worksheet.

- First, you specify which cells are locked and hidden,
- Then you finish up by **protecting the sheet** so your settings take effect.

Interestingly, every cell in the worksheet **starts in an unhidden and locked** state. If you switch on worksheet protection without making any changes, your whole worksheet becomes read-only.

Excel uses this approach for a reason. Typically, you'll use worksheet protection to make sure the person using your workbook can edit only a few select cells. It's much easier to designate the few cells that are editable than it is to try to select every single cell that needs to be locked.

7.2.4 PROTECTING A WORKSHEET USING PASSWORD

The protection of a worksheet in Excel is not a very strong security feature. The main Focus is that someone should not accidentally open it or change worksheet structure or cells

Worksheet Protection is mainly to prevent serious keying in errors that could damage the worksheet rather than security against reading content.

By default, all cells are protected but protection does not take effect until the sheet itself is protected.

When you protect a worksheet by locking its cells, adding a password to edit the unlocked elements is optional. In this context, the password is only intended to allow access to certain users while helping to prevent changes by other users.

Protecting a worksheet involves **protecting cells** and **protecting the worksheet**.

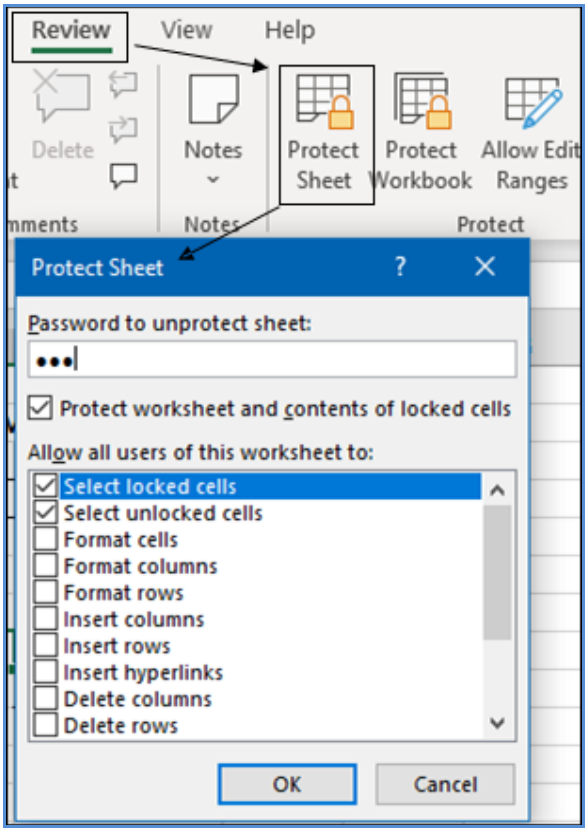


Figure 7.2.7 Protect worksheet

Actions of users which can be enabled or blocked

Select locked cells. Turn off this checkbox if you want to prevent people from moving to locked cells.

Important: if you can't select a locked cell, there's also no way to copy and paste the information in the cell to another worksheet or program. If you want the people using your workbook to be able to do this, you should keep the "Select locked cells" setting switched on.

Select unlocked cells. Turn off this checkbox if you want to prevent people from moving to unlocked cells.

Format cells, Format columns, and Format rows. Turn on these checkboxes if people need to be able to format individual cells or entire columns and rows. If you allow row and column formatting, Excel also permits

people to hide rows and columns. However, Excel never lets anyone change the locked and hidden settings of a cell while it is protected.

Insert columns and Insert rows. Turn on these checkboxes if you want to let people insert new rows or columns.

Insert hyperlinks. Click this checkbox if you want to let people insert hyperlinks in unlocked cells. This setting can be dangerous because a hyperlink can point to anything from another worksheet to a malicious Web page.

Delete columns and Delete rows. Turn on these checkboxes to bestow the ability to remove columns or rows. Use this setting at your peril; because it lets people decimate your worksheet—for example, removing entire ranges of data even if they contain locked cells.

Sort. Click this checkbox to let people sort unlocked cells

Use AutoFilter. Click this checkbox to let Users use filtering on any tables in the worksheet.

Use PivotTable reports. Click this checkbox to let Users manipulate any pivot tables in the worksheet.

Edit objects. Click this checkbox to let Users edit or delete embedded objects in the worksheet.

Edit scenarios. Click this checkbox to let Users edit or delete what-if scenarios.

CASE STUDY: Your principal CA Chandiwala wants you to protect the cells where you have put Formulas so that nobody can accidentally change formulas and also no one can see the formulas.

	A	B	C	D
1	Galaxy			
2	Trading, Profit & Loss Account for the period ended 31st March, 2020			
3				
4	PARTICULARS	AMOUNT	PARTICULARS	AMOUNT
5	OPENING STOCK	72,07,133.00	SALES	18,02,36,987.00
6	PURCHASE	17,41,19,967.00	CLOSING STOCK	77,73,432.00
7	Direct Exp.			
8	Gross Profit	66,83,319.00		
9		18,80,10,419.00		18,80,10,419.00
10				
11	Administrative Expenses	16,54,987.00	Gross Profit	66,83,319.00
12	Selling Expenses	12,36,987.00		
13	Finance Expenses	15,98,721.00		
14	Depreciation	1,96,200.00		
15	Net Profit	19,96,424.00		
16		66,83,319.00		66,83,319.00

Figure 7.2.8 Protect Formulas in the sheet

Strategy:

The strategy needed to protect your worksheet:

1. First, unlock all the cells into which you want people to type information.

- you can perform this procedure one cell at a time, or
- You can select an entire range of cells.

2. Select a**ll cells**
3. Type **CRTL** +1
4. Select Protection Under Format Cells
5. Click the **Protection** tab.
6. Turn off the **Locked** & **Hidden** checkbox, and then click **OK**.

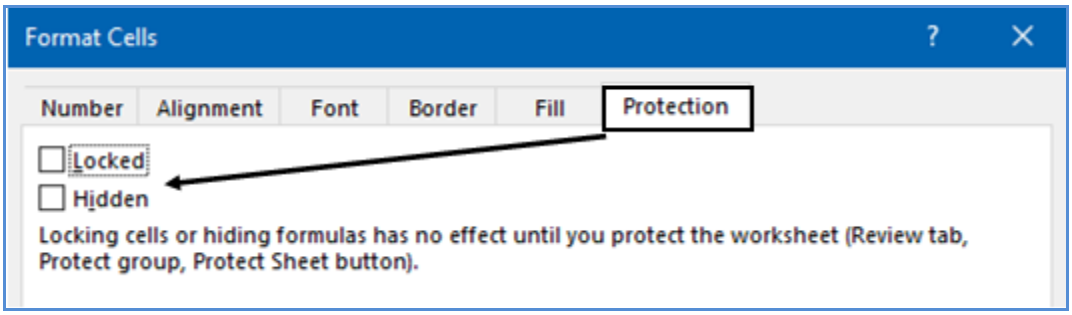


Figure 7.2.9 Unlock & unhide the entire sheet

2. You now need to hide formulas that you do not want the person using the workbook to see.
3. Select entire sheet
4. Press **F5** > **Special**
5. Select formulas as shown in figure

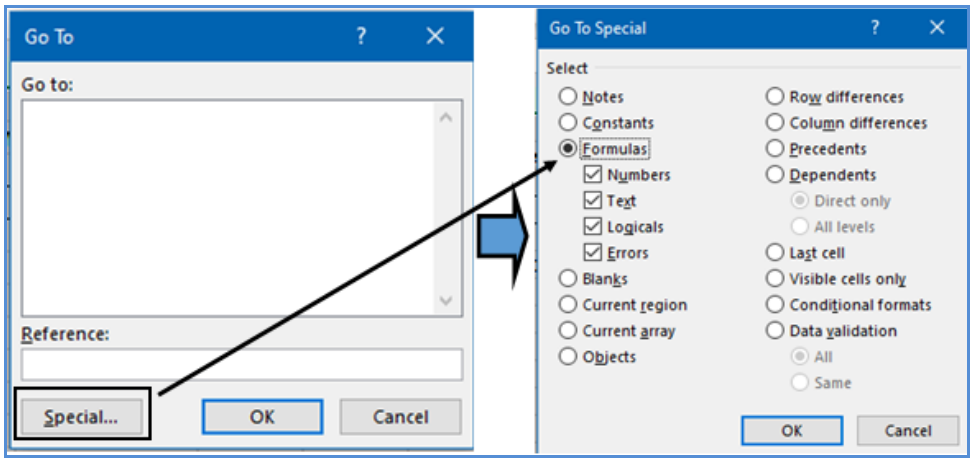


Figure 7.2.10 Go To Special Formulas

6. All the cells with Formulas will be selected as shown in Figure

	A	B	C	D
1	Galaxy			
2	Trading, Profit & Loss Account for the period ended 31st March, 2020			
3				
4	PARTICULARS	AMOUNT	PARTICULARS	AMOUNT
5	OPENING STOCK	72,07,133.00	SALES	18,02,36,987.00
6	PURCHASE	17,41,19,967.00	CLOSING STOCK	77,73,432.00
7	Direct Exp.			
8	Gross Profit	66,83,319.00		
9		18,80,10,419.00		18,80,10,419.00
10				
11	Administrative Expenses	16,54,987.00	Gross Profit	66,83,319.00
12	Selling Expenses	12,36,987.00		
13	Finance Expenses	15,98,721.00		
14	Depreciation	1,96,200.00		
15	Net Profit	19,96,424.00		
16		66,83,319.00		66,83,319.00

Figure 7.2.11 Cells with formula selected

7. Once you make your selection, **right-click** it, and then choose **Format** Cells. The **Format Cells dialog** box appears as shown in Fig 7.2.12

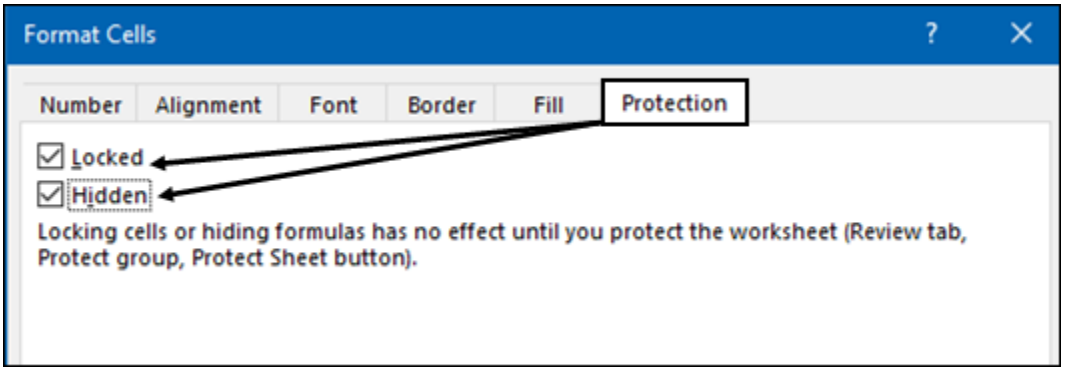


Figure 7.2.12 Lock and Hide cells

7. This time, click the **Protection** tab, **check** on the **Locked** and **Hidden** checkbox, and then click **OK**.
8. Once you have finished unlocking and hiding
9. To protect the sheet. Select **Review >Protect >Protect Sheet**.
10. In the Protect Sheet dialog box that appears, turn on the checkbox labeled **"Protect worksheet and contents of locked cells.**"

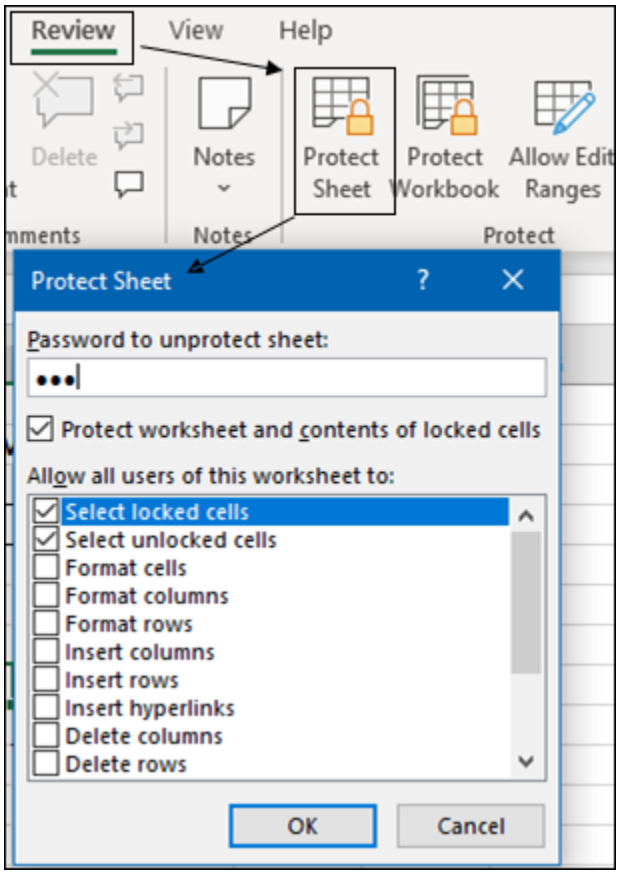


Figure 7.2.13 Protect sheet Select Locked/unlocked cells

11. In addition to protecting the contents of unlocked cells, this dialog box lets you **toggle on or off a list of Excel actions** that you want to let people using your worksheet perform.

12. From the "**Allow all users of this worksheet to**" list, turn on the things that you want people using your worksheet to be able to do.

13. If you want to stop other people from unprotecting the worksheet, specify a password in the "**Password to unprotect sheet**" text box.

14. Use strong passwords that combine uppercase and lowercase letters, numbers, and symbols. Passwords should be 8 or more characters in length.

15. If ever you want to unprotect it. Select **Review>Protect>Unprotect Sheet**. And enter the password.

Gist: You have learnt to protect & Hide cells and Worksheets so that any user of Worksheet does not accidentally change or view Formulas.

Commands Learnt: Review> Protect worksheet

7.3 GET AND TRANSFORM

When you are analysing , data often arrives as an Excel spreadsheet, and pretty often, it is a data dump into a CSV or a query into a SQL database. Most of the time, the data is arranged in a different layout or does not have the required layout or data points for analysis. You need to scrub and clean this data as discussed in chapter 6.

Excel now has an entire set of data transformation tools called **Get and Transform** also known as Power Query (Discussed in the previous Chapter) with it you can use its extract, transform, and load (ETL) functionality to seamlessly link to the data sources and get to insights quicker.

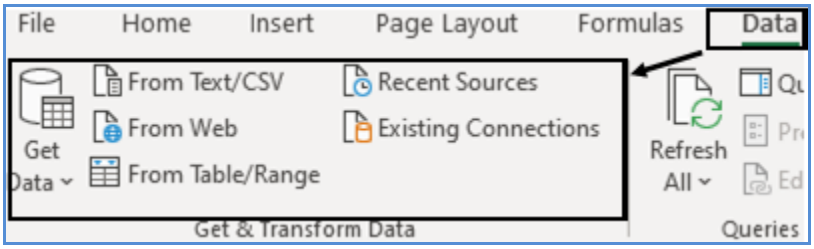


Figure 7.3.1 Get & Transform Data

7.3.1 DATA CONNECTIONS WITH ODBC

You can use a query to connect to a single data source, such as an Access database, or you can connect to multiple files, databases, ODBC, or Web sites.

CASE STUDY: Your principal CA Chandiwala wants you to create a connection with Ledger Balances in TALLY ERP so that the data in excel can be refreshed and Balance is updated regularly.

Strategy:

The strategy needed to use **GET AND TRANSFORM** to create a connection with ODBC to TALLY ERP Application

1. First, **open Tally** it should have **ODBC Active** as shown in Figure

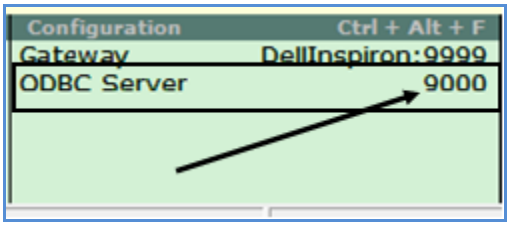


Figure 7.3.2 Tally ODBC

2. Open a new Excel sheet.
3. Click **Data > Get External Data > From Other Sources > From ODBC**, as shown in Figure

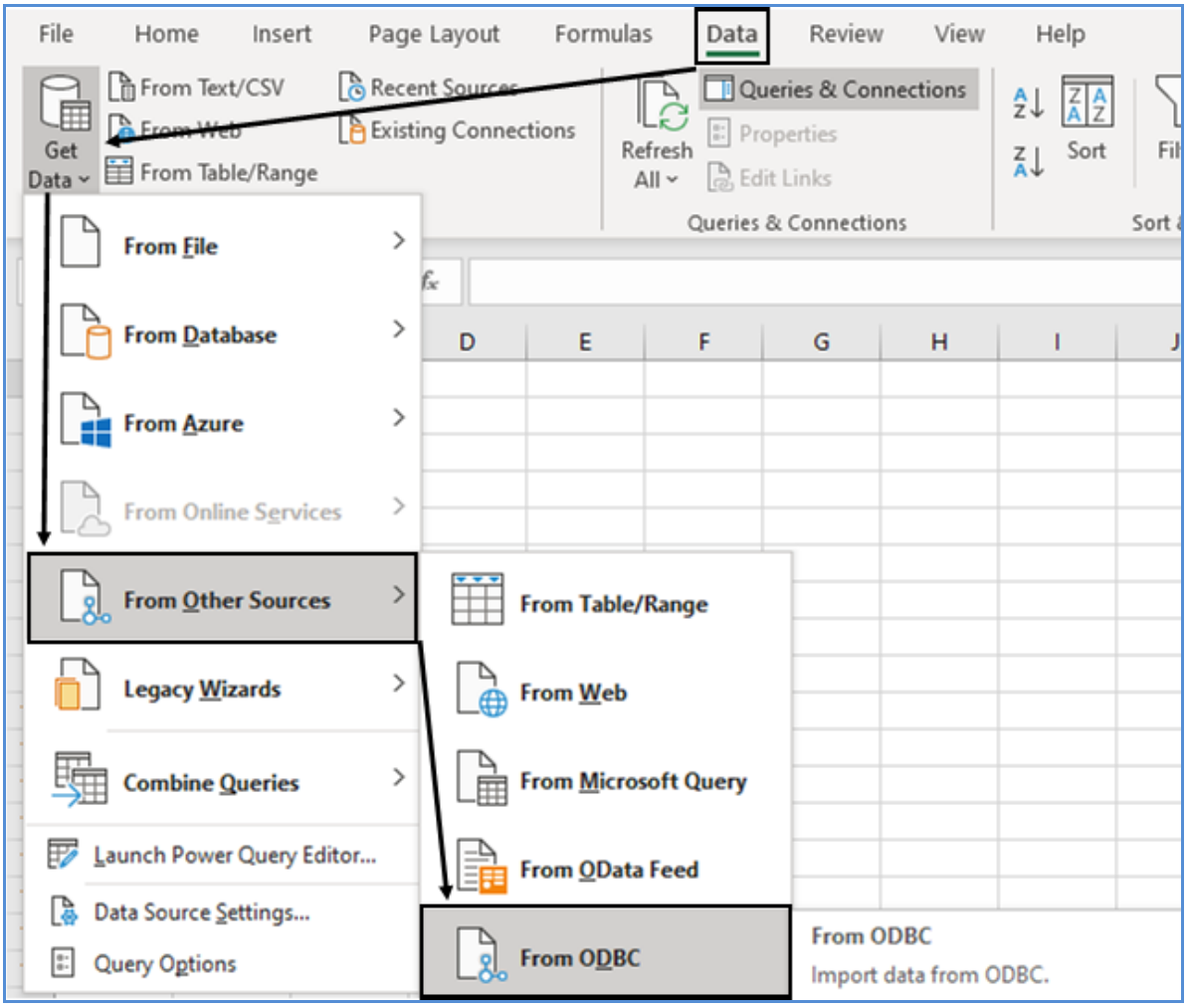


Figure 7.3.3 Get External data

4. A **From ODBC** Dialog box will open

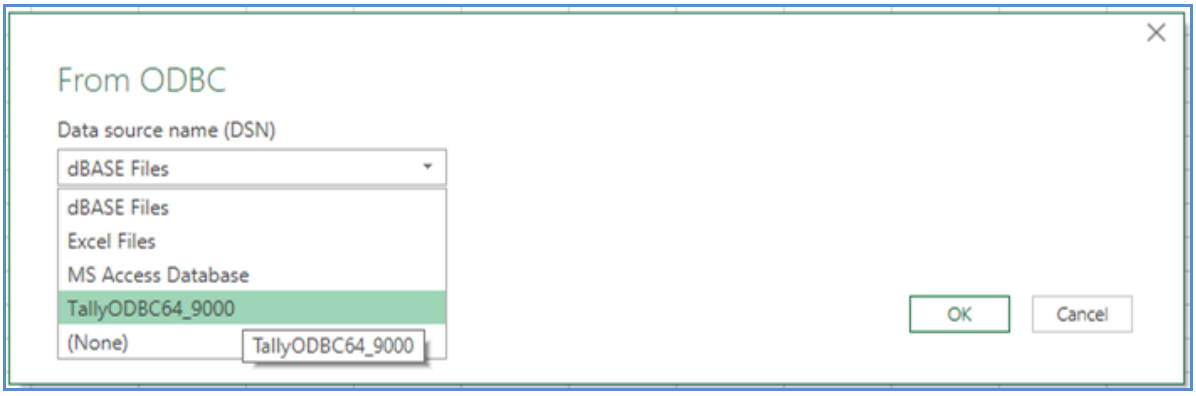


Figure 7.3.4 From ODBC

5. Select **TallyODBC64_9000** or any other ODBC data Connection you choose.

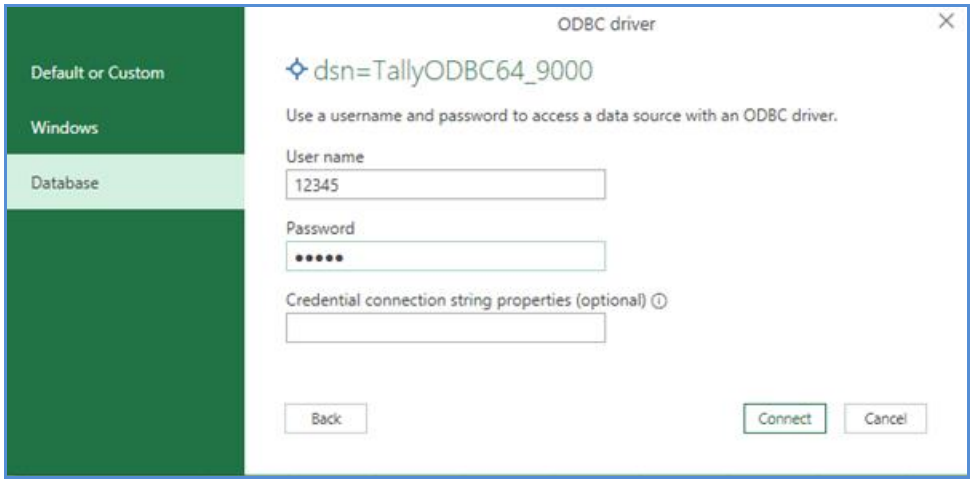


Figure 7.3.5 ODBC Credential

6. Provide the **Tally Credentials** and click Connect
7. It will create a connection with Tally Database and show the tables available
8. Select the table from which you want to create a data connection.

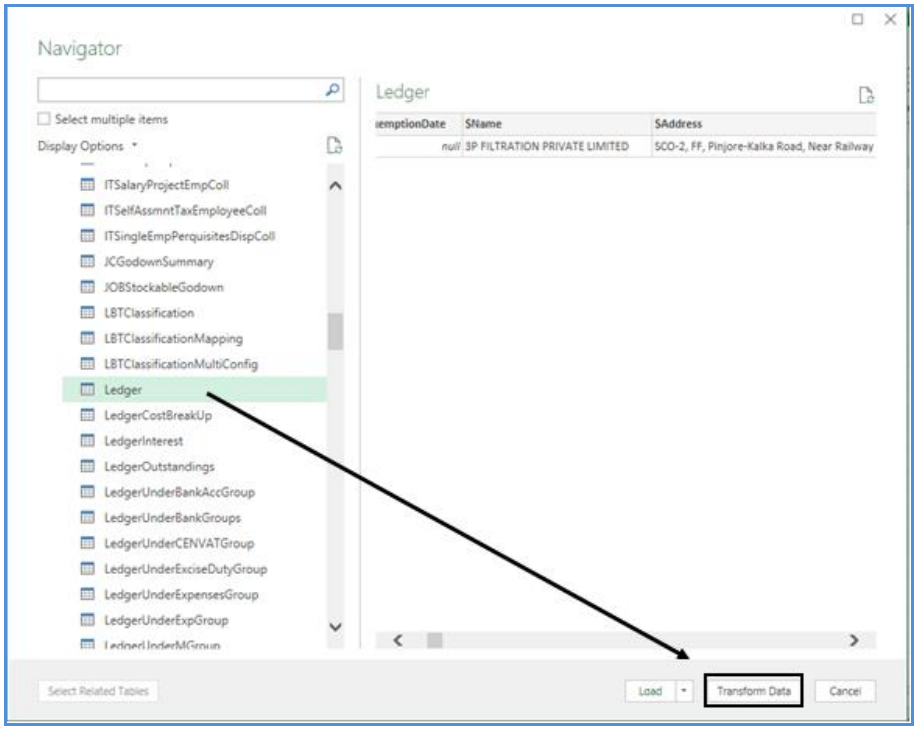


Figure 7.3.6 Select Table

9. Click on **Transform Data.**
10. The **power query editor** will open.
11. Click on **Choose columns**
12. Select **$Name** and **$_Closing Balance** as shown in Figure
13. **Remove other Columns**
14. Only Name and closing Balance columns will remain.

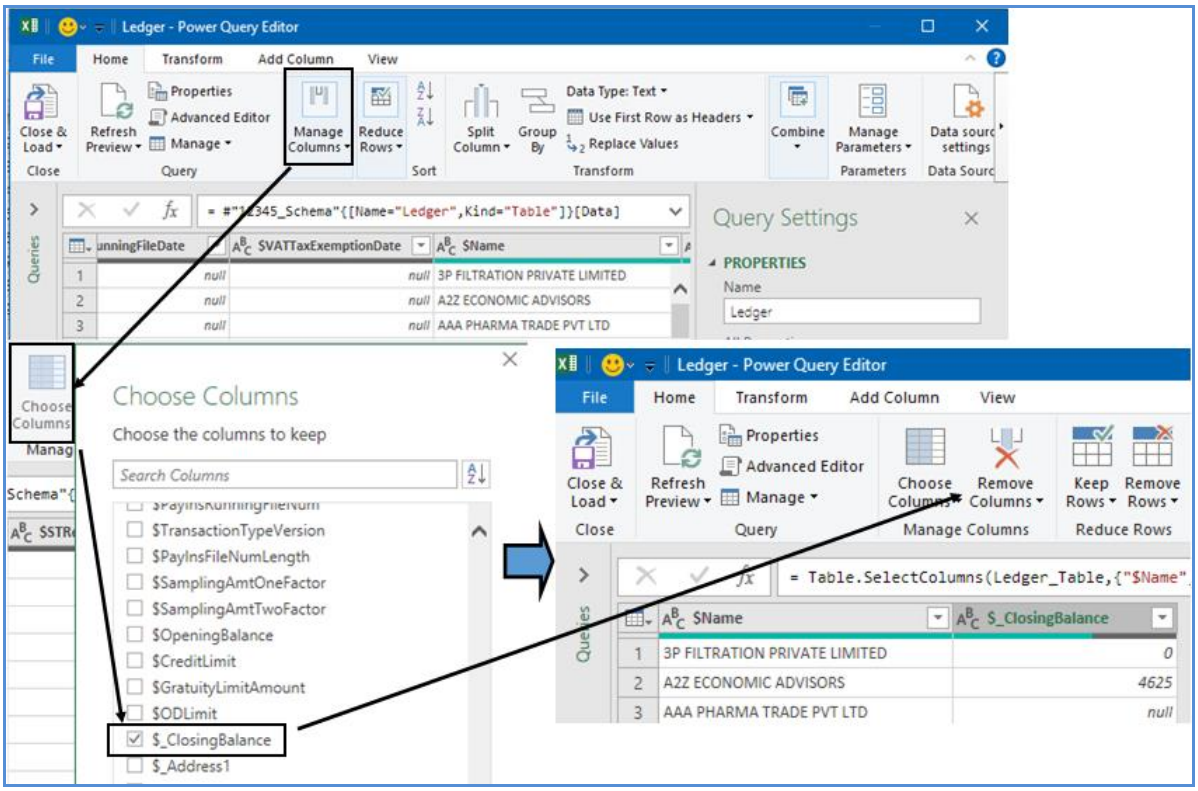


Figure 7.3.7 Power Query Editor

15. Click on **Close and load to**
16. **Import Data** Dialog Box will open
17. Select **only create connection.**

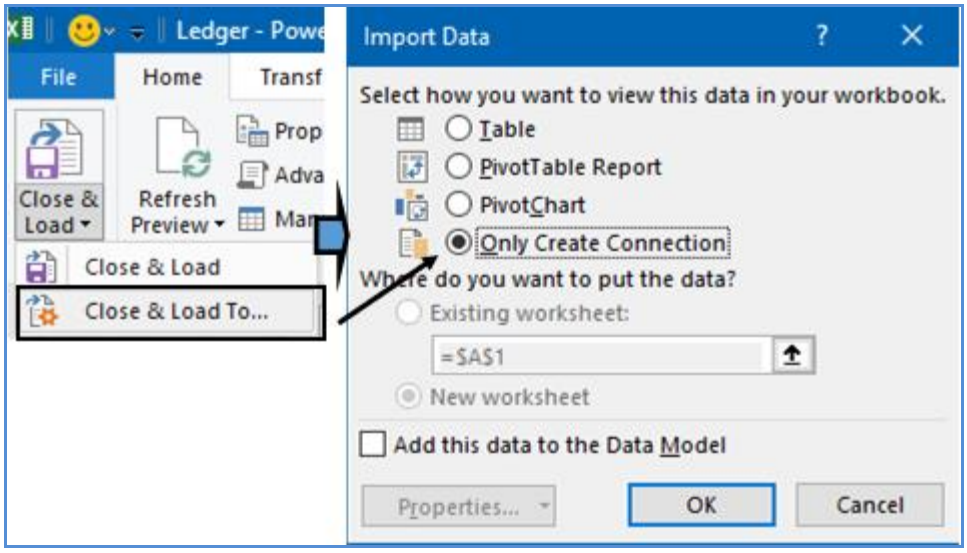


Figure 7.3.8 Load to Excel

18. It will be loaded to Excel as a connection

19. You can then **right-click** and **Load to** Excel as a Table

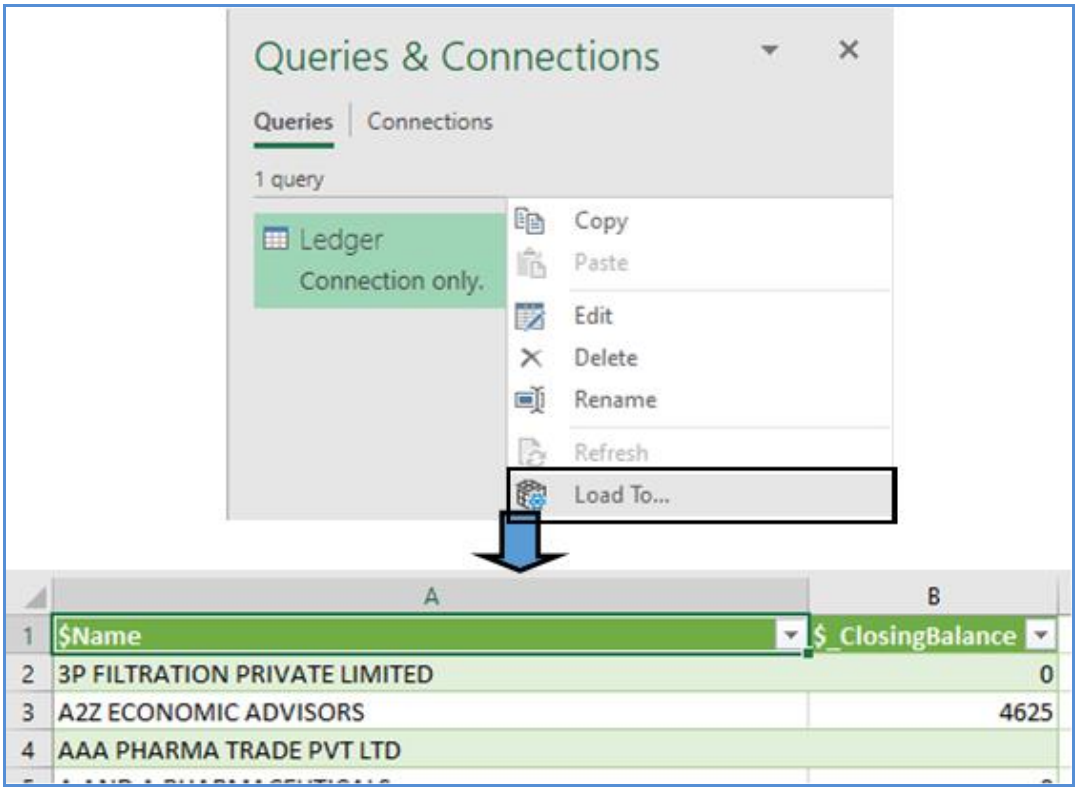


Figure 7.3.9 Ledger Accounts and Balances Loaded to Excel

20. The Tally data is loaded onto Excel as a connection which you can refresh.

Gist: You have used Get and Transform to connect Excel to ODBC of Tally and build a refreshable Query

Commands Learnt: Get & Transform data> Connect to ODBC

7.3.2 IMPORT FROM TEXT, XML AND WEB

Get & Transform lets you perform a series of steps to transform your Excel data.

If the Data is inside a text file, it's very easy to import data from text and right into Power Query.

You have to import data from a text file.

- Go to Data > New Query > From File > From Text

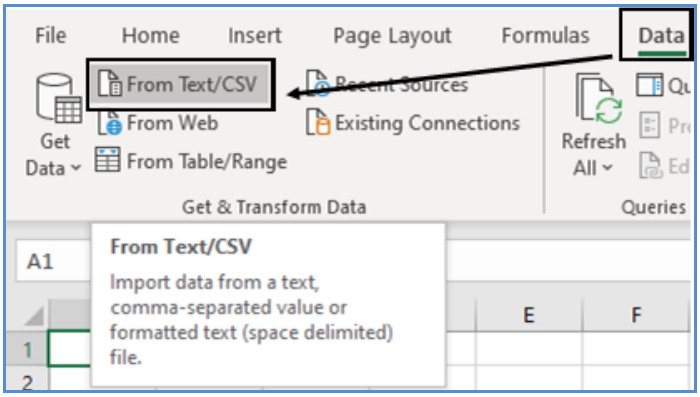
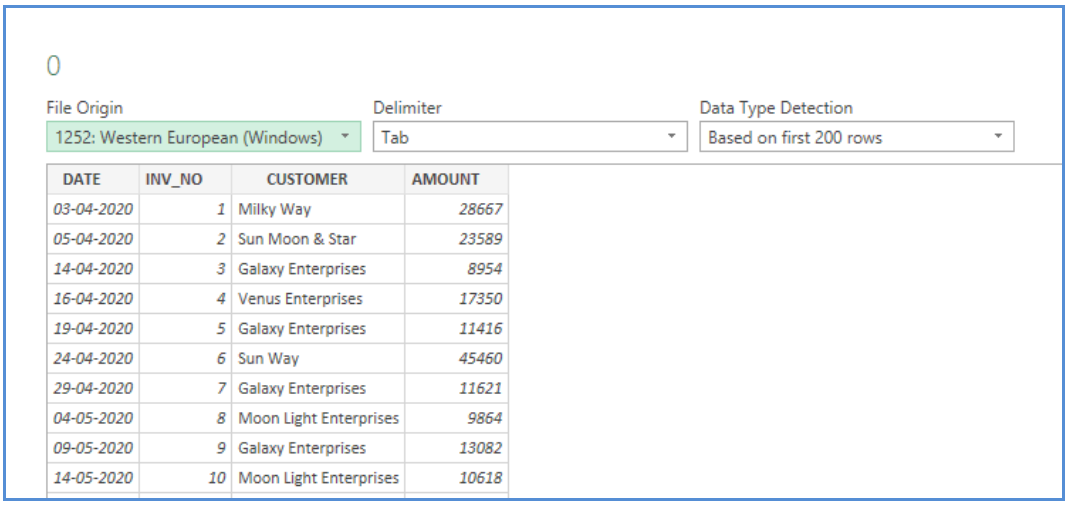


Figure 7.3.10 Get & Transform from TEXT

- Select the **text file** (with extension .txt) that contains the data.
- Click Import.



DATE	INV_NO	CUSTOMER	AMOUNT
03-04-2020	1	Milky Way	28667
05-04-2020	2	Sun Moon & Star	23589
14-04-2020	3	Galaxy Enterprises	8954
16-04-2020	4	Venus Enterprises	17350
19-04-2020	5	Galaxy Enterprises	11416
24-04-2020	6	Sun Way	45460
29-04-2020	7	Galaxy Enterprises	11621
04-05-2020	8	Moon Light Enterprises	9864
09-05-2020	9	Galaxy Enterprises	13082
14-05-2020	10	Moon Light Enterprises	10618

Figure 7.3.11 Preview

- A preview of the text data will be shown. If it is as desired, press **Transform**
- This will open up the **Power Query Editor**.
- Go to Home > Use First Row As Headers.
- This will give your table the correct Column Headers.

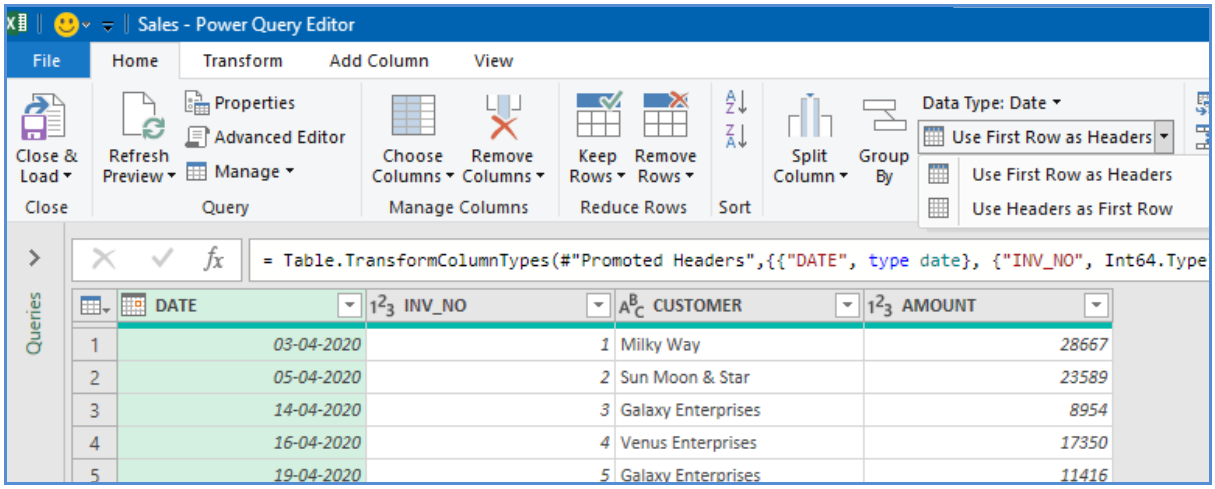


Figure 7.3.12 Power Query editor

- Click **Close & Load** from the **Home** tab.
- It will open up a new worksheet in your Excel workbook with the imported table.

	A	B	C	D
1	DATE	INV_NO	CUSTOMER	AMOUNT
2	03-04-2020	1	Milky Way	28667
3	05-04-2020	2	Sun Moon & Star	23589
4	14-04-2020	3	Galaxy Enterprises	8954
5	16-04-2020	4	Venus Enterprises	17350
6	19-04-2020	5	Galaxy Enterprises	11416
7	24-04-2020	6	Sun Way	45460
8	29-04-2020	7	Galaxy Enterprises	11621
9	04-05-2020	8	Moon Light Enterprises	9864
10	09-05-2020	9	Galaxy Enterprises	13082
11	14-05-2020	10	Moon Light Enterprises	10618
12	19-05-2020	11	Earth Limited	59705
13	20-05-2020	12	Milky Way	21472

Figure 7.3.13 Text Imported to Excel

Import from XML Steps are similar

- Data > Get Data > from File >from XML

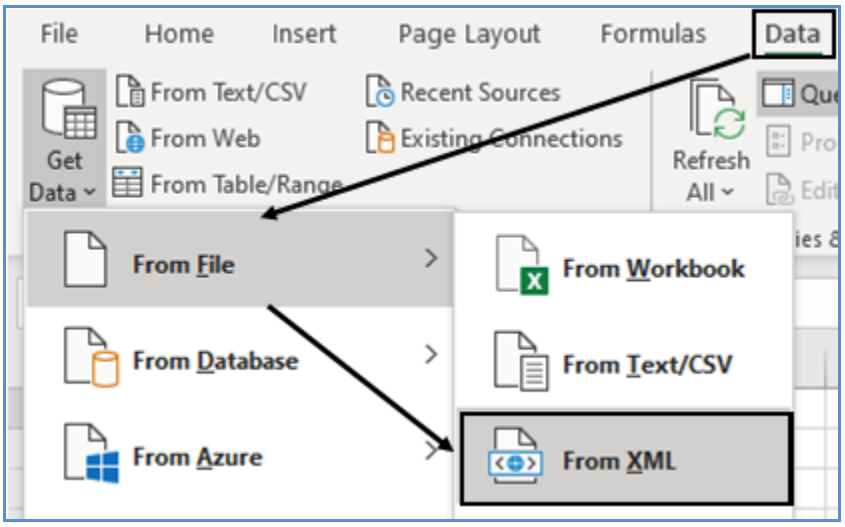


Figure 7.3.14 Import from XML

Import from Web

Through Get & Transform, you can link to a web page

As an example, you want data of forex rates from the Economic Times Website

- Copy the URL from the webpage (assuming you are previewing the page in a browser).
- Open a **new Excel Sheet**
- Select Data (tab) > Get & Transform (group) > From Web.

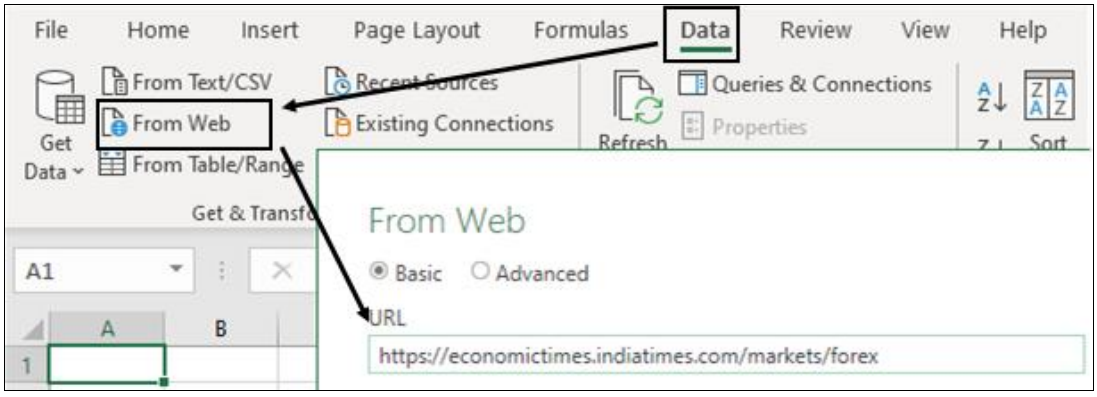


Figure 7.3.15 Get & Transform From Web

- In the From Web dialog box, paste the URL into the URL field and click OK.
- Access web content Dialog Box will appear.

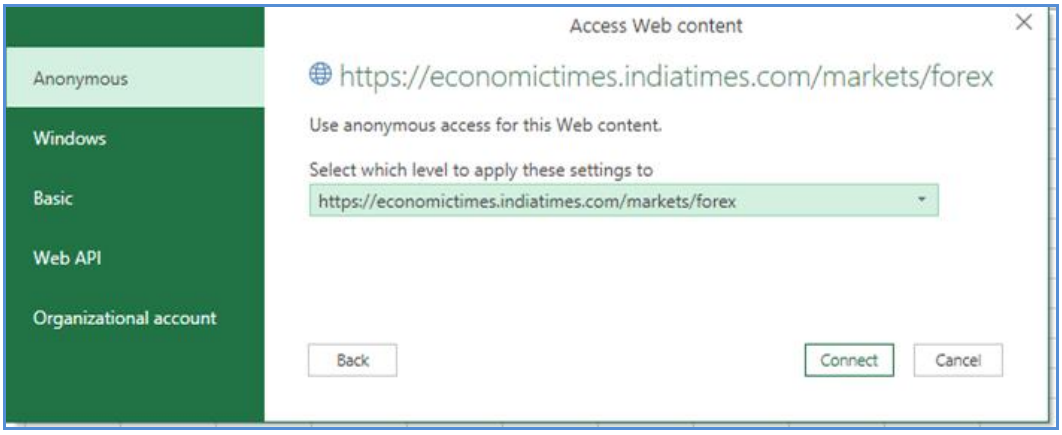
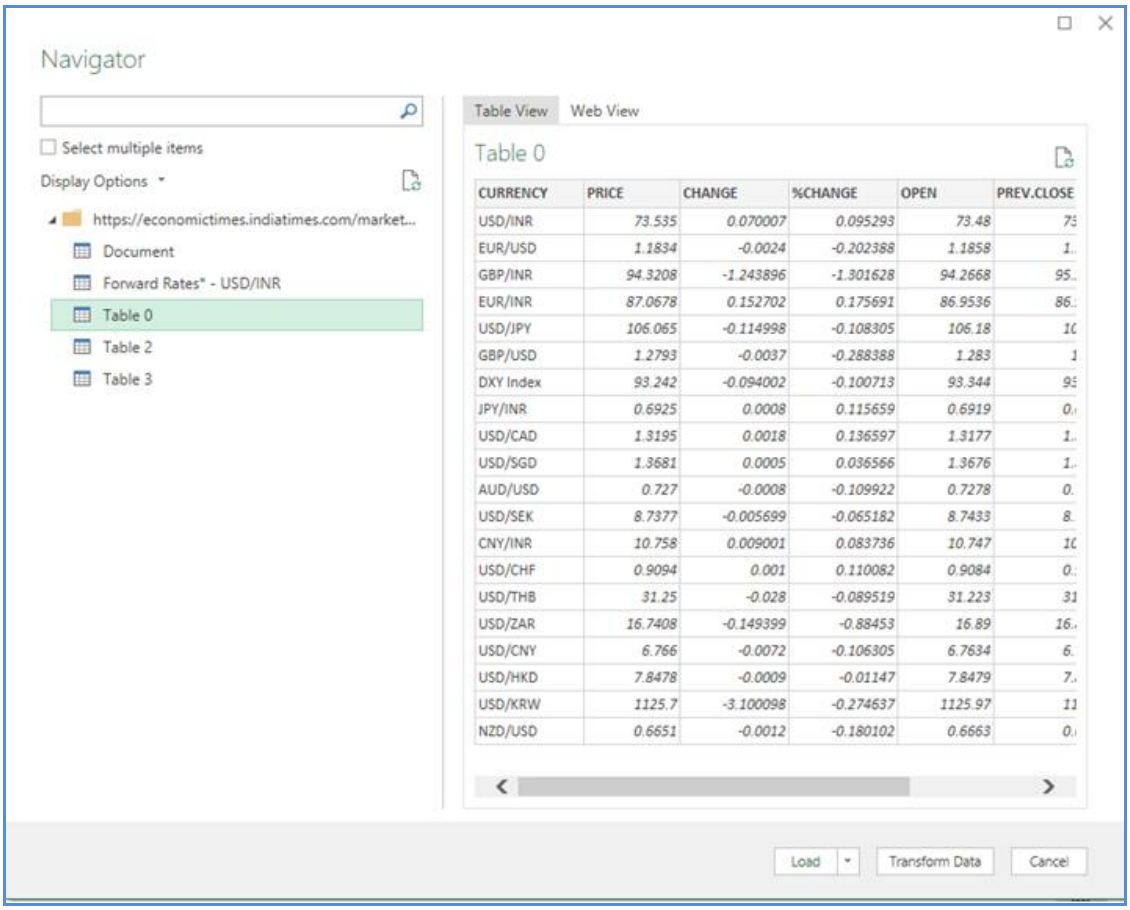


Figure 7.3.16 Access web content Dialog Box

- The **Navigator window** displays the components of the webpage in the left panel.
- Click on the desired Table, **preview** of each item in the right-side preview panel, is available.



CURRENCY	PRICE	CHANGE	%CHANGE	OPEN	PREV.CLOSE
USD/INR	73.535	0.070007	0.095293	73.48	73
EUR/USD	1.1834	-0.0024	-0.202388	1.1858	1.
GBP/INR	94.3208	-1.243896	-1.301628	94.2668	95.
EUR/INR	87.0678	0.152702	0.175691	86.9536	86.
USD/JPY	106.065	-0.114998	-0.108305	106.18	10
GBP/USD	1.2793	-0.0037	-0.288388	1.283	1
DXY Index	93.242	-0.094002	-0.100713	93.344	93
JPY/INR	0.6925	0.0008	0.115659	0.6919	0.
USD/CAD	1.3195	0.0018	0.136597	1.3177	1.
USD/SGD	1.3681	0.0005	0.036566	1.3676	1.
AUD/USD	0.727	-0.0008	-0.109922	0.7278	0.
USD/SEK	8.7377	-0.005699	-0.065182	8.7433	8.
CNY/INR	10.758	0.009001	0.083736	10.747	10
USD/CHF	0.9094	0.001	0.110082	0.9084	0.
USD/THB	31.25	-0.028	-0.089519	31.223	31
USD/ZAR	16.7408	-0.149399	-0.88453	16.89	16.
USD/CNY	6.766	-0.0072	-0.106305	6.7634	6.
USD/HKD	7.8478	-0.0009	-0.01147	7.8479	7.
USD/KRW	1125.7	-3.100098	-0.274637	1125.97	11
NZD/USD	0.6651	-0.0012	-0.180102	0.6663	0.

Figure 7.3.17 Navigator Window

- If the data does not require any further transformations, you can click **Load/Load To.**
- In the **Import Data** dialog box, select "**Existing worksheet**" and point to a cell on your desired destination worksheet (like cell A1 on "Sheet1").

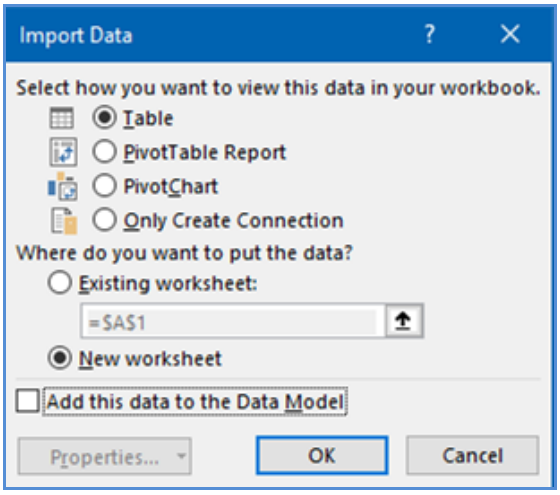


Figure 7.3.18 Import data Dialog Box

- The result is a table that is connected to a query.

	A	B	C	D	E	F	G
1	CURRENCY	PRICE	CHANGE	%CHANGE	OPEN	PREV.CLOSE	DAY's LOW-HIGH
2	USD/INR	73.535	0.070007	0.095293	73.48	73.465	73.4075- 73.62
3	EUR/USD	1.1835	-0.0023	-0.193954	1.1858	1.1858	1.1819- 1.1874
4	GBP/INR	94.3208	-1.243896	-1.301628	94.2668	95.5647	94.1111- 94.5372
5	EUR/INR	87.0678	0.152702	0.175691	86.9536	86.9151	86.8565- 87.151
6	USD/JPY	106.057	-0.123001	-0.115842	106.18	106.18	106.057- 106.266
7	GBP/USD	1.2793	-0.0037	-0.288388	1.283	1.283	1.2763- 1.2866
8	DXY Index	93.242	-0.094002	-0.100713	93.344	93.336	93.162- 93.352
9	JPY/INR	0.6925	0.0008	0.115659	0.6919	0.6917	0.6913- 0.6931
10	USD/CAD	1.3195	0.0018	0.136597	1.3177	1.3177	1.3151- 1.3208
11	USD/SGD	1.3681	0.0005	0.036566	1.3676	1.3676	1.3657- 1.37
12	AUD/USD	0.7269	-0.0009	-0.123664	0.7278	0.7278	0.7256- 0.7306
13	USD/SEK	8.7377	-0.005699	-0.065182	8.7433	8.7434	8.7335- 8.7805
14	CNY/INR	10.758	0.009001	0.083736	10.747	10.749	10.734- 10.769
15	USD/CHF	0.9093	0.0009	0.099079	0.9084	0.9084	0.908- 0.9106
16	USD/THB	31.25	-0.028	-0.089519	31.223	31.278	31.188- 31.266
17	USD/ZAR	16.7408	-0.149399	-0.88453	16.89	16.8902	16.7384- 16.9342
18	USD/CNY	6.766	-0.0072	-0.106305	6.7634	6.7732	6.754- 6.7698
19	USD/HKD	7.8478	-0.0009	-0.01147	7.8479	7.8487	7.8472- 7.8484
20	USD/KRW	1125.7	-3.100098	-0.274637	1125.97	1128.8	1123.25- 1126.95
21	NZD/USD	0.6652	-0.0011	-0.165091	0.6663	0.6663	0.6642- 0.6695

Figure 7.3.19 Excel connected to Web

7.4 DATA TOOLS

7.4.1 TEXT TO COLUMNS

Generally, when you import data from some external application, as an example a banking application, you might get all information in one column whereas you want it in separate columns. There could also be a case where you may want to split the data into the cells in a Column for better filtering, sorting, or a detailed analysis.

In all these cases **Text to Columns** feature under **Data Tools** in **Data Tab** comes in very handy. It facilitates Splitting text by a specified Delimiter like Tab, Space Comma, etc or even by Fixed width

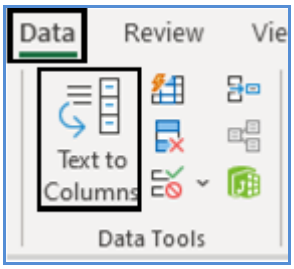


Figure 7.4.1 Text to Column

Tab or some other special character delimited files.

Comma-separated value. A comma (,) usually separates each column eg data from a CSV file

Fixed-length or space-delimited, where each column has a fixed length

CASE STUDY: Your principal CA Chandiwala has copied The Expense account which was in a CSV file to Excel which has come in one column as shown in Figure. He wants you to separate data in one cell into several cells such as Date, Accounting Code, Ledger Name, and Debit so that he can do some analysis on the same.

	A
1	Date,Accounting Code,Ledger name, Debit
2	02-04-2020,E0011,Godown Rent," 18,000.00 "
3	02-04-2020,E0020,Printing & Stationery," 2,500.00 "
4	04-04-2020,E0001,Bank Charges, 500.00
5	05-04-2020,E0020,Printing & Stationery," 5,400.00 "
6	05-04-2020,E0021,Staff Welfare Expenses, 250.00
7	10-04-2020,E0015,Membership Fees," 20,000.00 "
8	16-04-2020,E0005,Courier Charges," 1,500.00 "
9	16-04-2020,E0013,Interest Account," 9,778.96 "
10	20-04-2020,E0021,Staff Welfare Expenses, 300.00
11	25-04-2020,E0013,Interest Account," 3,691.27 "
12	28-04-2020,E0021,Staff Welfare Expenses, 325.00
13	30-04-2020,E0002,Basic Pay," 46,923.00 "

Figure 7.4.2 Comma Separated values

Strategy:

You can use the Text to Column feature for that.

1. Select the cells you want to Split.
2. Navigate to the **Data tab > Data Tools** group,
3. Click the **Text to Columns** button.

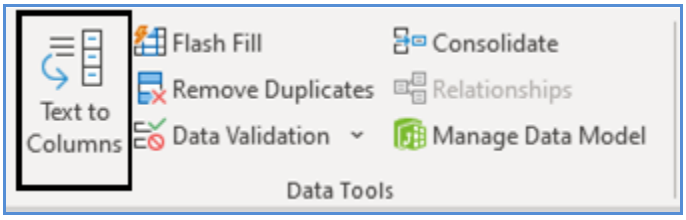


Figure 7.4.3 Text to Columns

4. A 3 Step **Text to Columns Wizard** will appear
5. In the first step of the **Convert Text to Columns wizard**, you choose how to split cells - by delimiter or width.
6. In your case, the cell contents are separated with delimiter which is a comma,
7. Select Delimited, and click Next.

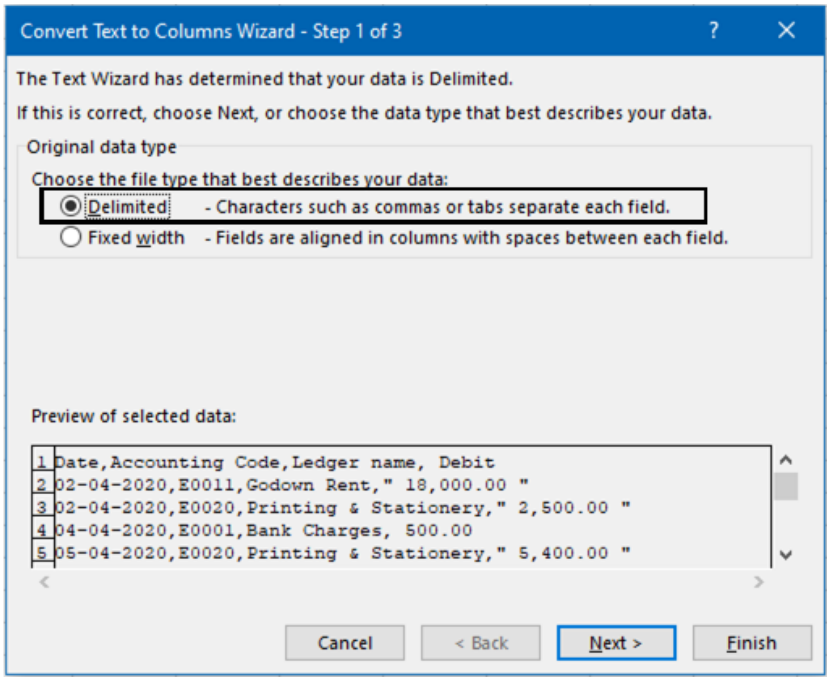


Figure 7.4.4 Text to Column Wizard

8. In Step 2, you specify the delimiters and, optionally, text qualifier.

9. You can choose one or more predefined delimiters as well as type your own one in the Other box. In this case, you select Comma.

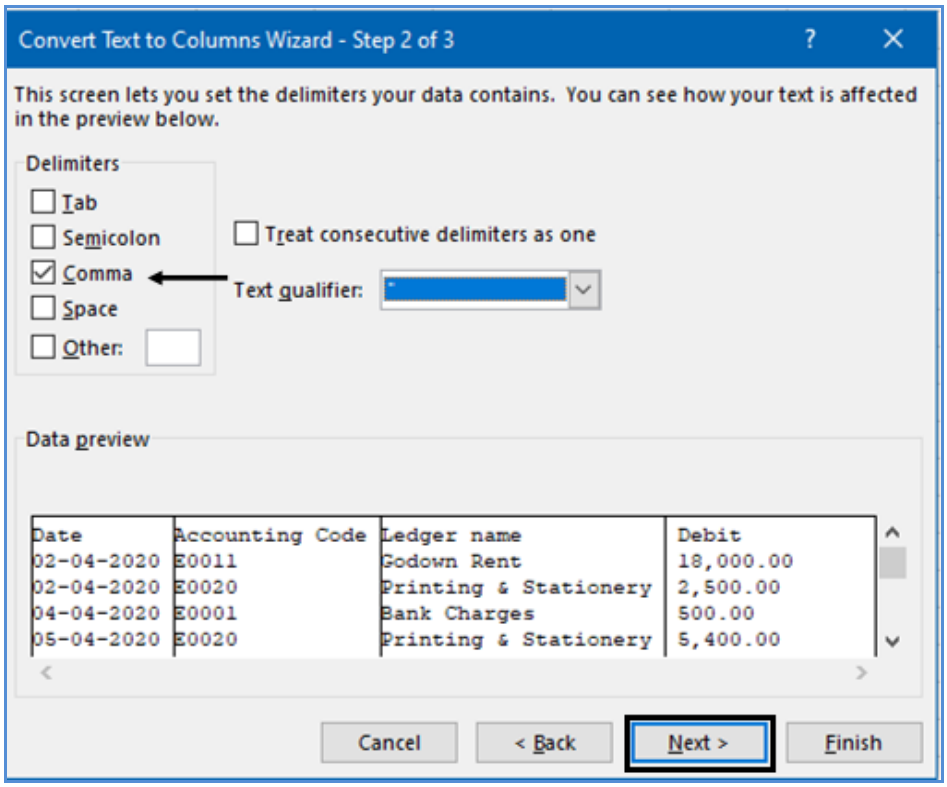


Figure 7.4.5 Specify Delimiter

- ***Treat consecutive delimiters as one*.** select this option when your data may contain two or more delimiters in a row, e.g. a comma and space.
- ***Specifying the text qualifier***. Use this option when some text is enclosed in single or double quotes, and you'd like such portions of text to be inseparable. For example, if you choose a comma (,) as the delimiter and a quotation mark (") as the text qualifier, then any words enclosed in double-quotes, e.g. "18,000.00", will be put into one cell as 18,000.00. If you select {none} as the text qualifier, then "18 will be distributed into one cell (together with an opening quotation mark) and 000.00" into another (together with a closing mark).
- ***Data preview***. You can scroll through the Data preview section to make sure Excel has split all cell's contents correctly.

10. In Step 3 you choose the data format and specify where you want to paste split data:

- **Data format.** By default, the General format is set for all columns, which OK in most cases. In your example, you need the Data format for the dates with a DMY date mask. You can change the data format for a particular column, click on that column under Data preview to select it, and then choose one of the formats under Column data format.

- **Destination**. You can specify where you want to output the separated data, select $A$1

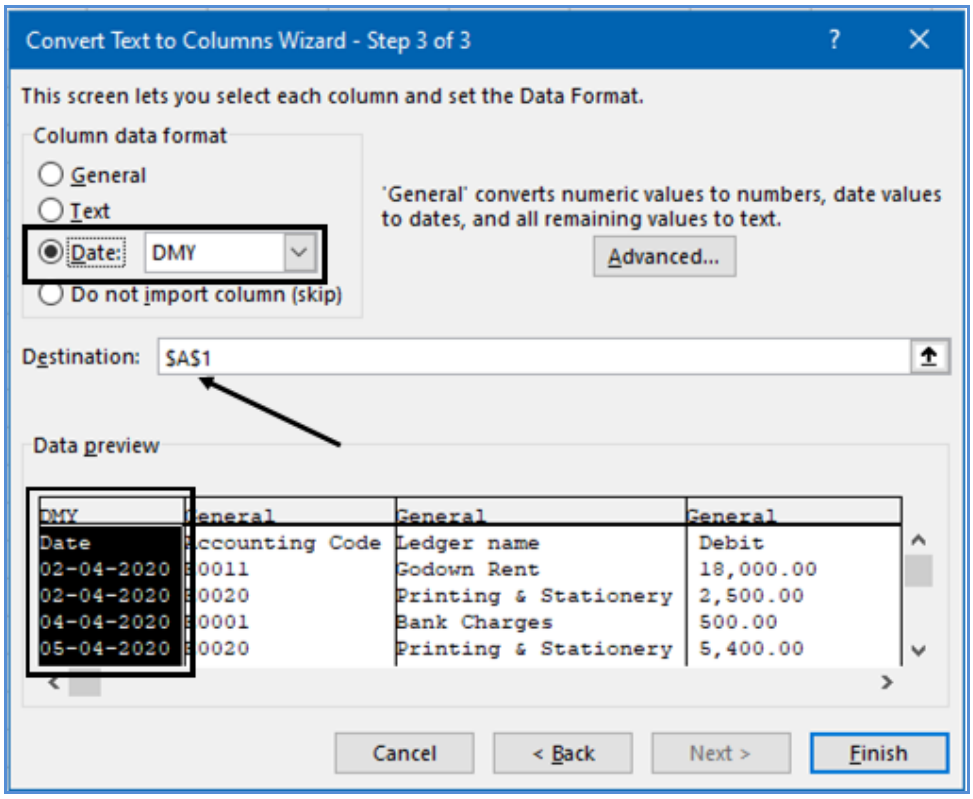


Figure 7.4.6 Specify Format and Destination

11. Click the Finish

12. Excel has perfectly split the contents of one cell into several cells as shown in Figure

	A	B	C	D
1	Date	Accounting Code	Ledger name	Debit
2	02-04-2020	E0011	Godown Rent	18,000.00
3	02-04-2020	E0020	Printing & Stationery	2,500.00
4	04-04-2020	E0001	Bank Charges	500.00
5	05-04-2020	E0020	Printing & Stationery	5,400.00
6	05-04-2020	E0021	Staff Welfare Expenses	250.00
7	10-04-2020	E0015	Membership Fees	20,000.00
8	16-04-2020	E0005	Courier Charges	1,500.00

Figure 7.4.7 Final Data

Gist: You have used Text to Columns to split your data into separate Columns

Commands Learnt: Data Tools> Text to Columns

7.4.2 DATA CONSOLIDATION

You have data on separate worksheets you can check them individually in separate worksheets, but would be great if you can get summarised results into one worksheet. Excel allows this through the **CONSOLIDATE** feature under **Data Tab** thereby lets you pull-each record from the separate worksheet, consolidating data from into a single master sheet. Consolidation is used for budgets, inventory requirements, business forecasts, surveys, experimental results, and a lot more

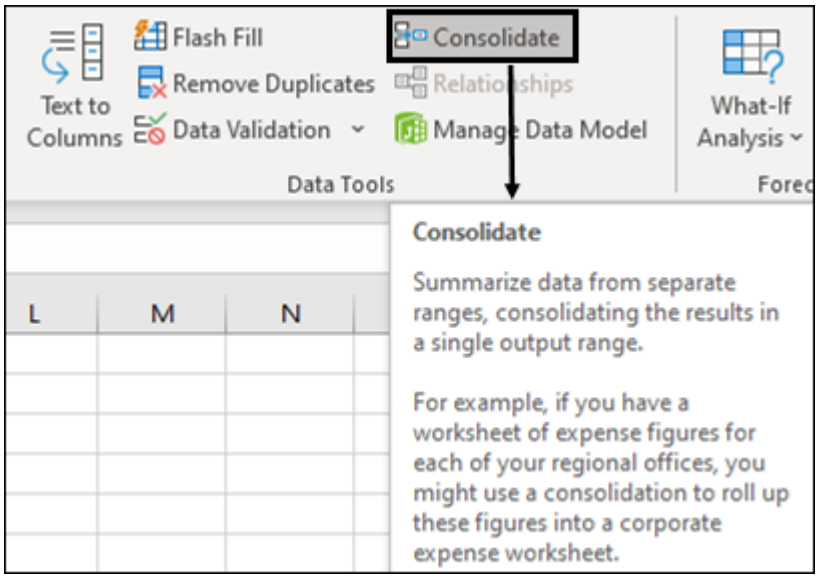


Figure 7.4.8 Consolidate in Data Tools

Consolidation is the process of combining values from several ranges of data either from within the same or different workbooks. It can be used to summarize data from different worksheets into the master worksheet and create a report using a variety of calculations

The worksheets you consolidate can be in the **same workbook** as the master worksheet or in **multiple workbooks**.

Benefits of consolidation of data

- Easy updation
- Aggregation in one window on a regular or ad-hoc basis.

Data can be consolidated differently:

- **Consolidated by Position** when all the referring data is in the same location and order,
- **Consolidate by Category** when location and order are not the same.

Consolidate by Position

Useful, when worksheet data is **identical in order and location**. This Consolidation works like layering ie. if you open the workbook you wish to consolidate, you specify the ranges to consolidate, then the values in one

worksheet are layered on those of another worksheet until all the worksheets are overlaid. It is then that you can select the calculation to be performed on the data. You could use Sum, Count, Average, Max, Min, Product, STDdev, StdevP, Var, etc You can also, by checking the create links to the source data box, have the consolidation to update automatically when the source data changes.

Consolidate by Category

Useful, when data is organized **differently but has identical row and column labels**. This Consolidation works in nearly the same way but in this case, while selecting data to consolidate, you have to include row and/or column headings. Excel will then examine the row and/or column headings and will be able to plot the layout of the Worksheets and Consolidate data by examining the contents of the ranges to be used.

CASE STUDY: Your principal CA Chandiwala has the expenditure of 4 quarters in 4 worksheets Q1, Q2, Q3 & Q4 and they are having the same layout but in each of the sheets order of expense head is different, consolidating them manually is a challenge and he wants to consolidate them into annual expenditure.

	A	B
1	**Head**	**Expenditure**
2	Salaries & Wages	6,20,600
3	Bank Charges	6,090
4	Depreciation	54,406
5	Professional Fee	15,600
6	Insurance Expenses	5,440
7	Travelling Expenses	17,898
8	License & Fees	40,500
9	Dues & Subscriptions	13,620
10	Staff Welfare	411
11	Advertising Expense	281

Q1 Q2 Q3 Q4 Y1

Figure 7.4.9 Quarterly Expenditure

Strategy:

You can help him consolidate the same using excel as follows:

Consolidate by position

In each worksheet that contains the data that you want to consolidate, set up the data by doing the following:

1. Ensure that range of data is in list format with Labels/Headers in the first row
2. There are no blank rows & columns within the list
3. Each range which you want to consolidate is on a different worksheet

4. Each range has the same layout.
5. In your case, all the worksheets are on the same layout.
6. You can also give each quarter results, a **Name range**.
7. Go to the first sheet by the name of *Q1,* and select the data, from the top-left corner of the sheet, enter *Quarter_1,* as shown in **Figure** below.

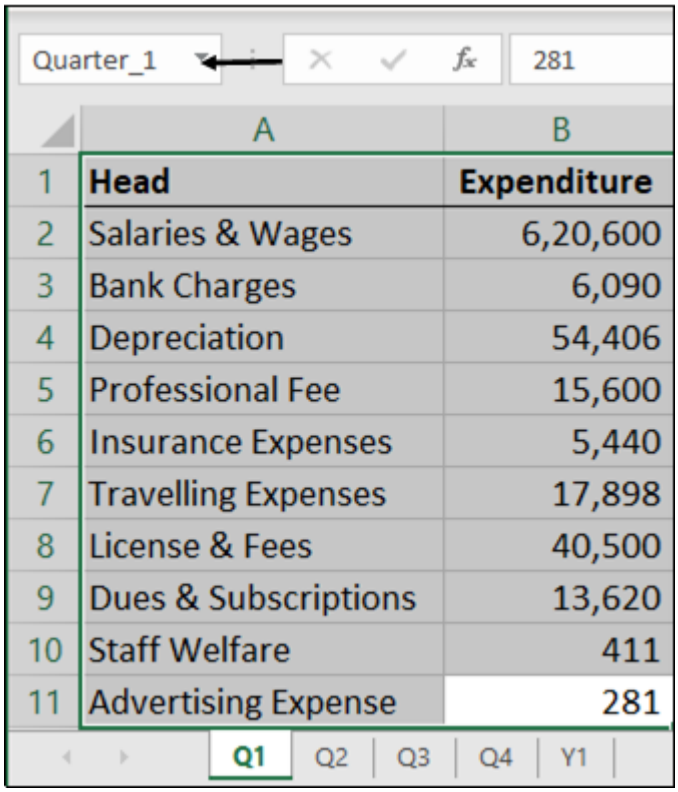


Figure 7.4.10 Naming a range in Worksheet

Alternate strategy:

8. Since the layout of all worksheets is the same you can select A1:B11 also
9. Repeat the same procedure in other sheets Q2, Q3 & Q4 to assign them name range *Quarter_2, Quarter_3, and Quarter_4* respectively.
10. Now click the upper-left cell o*f the destination area for the consolidated data (in your case "Y1")*
11. *This is where the consolidated* data will begin.
12. *From* the **Data** menu, under **Data tools** choose **Consolidate**.

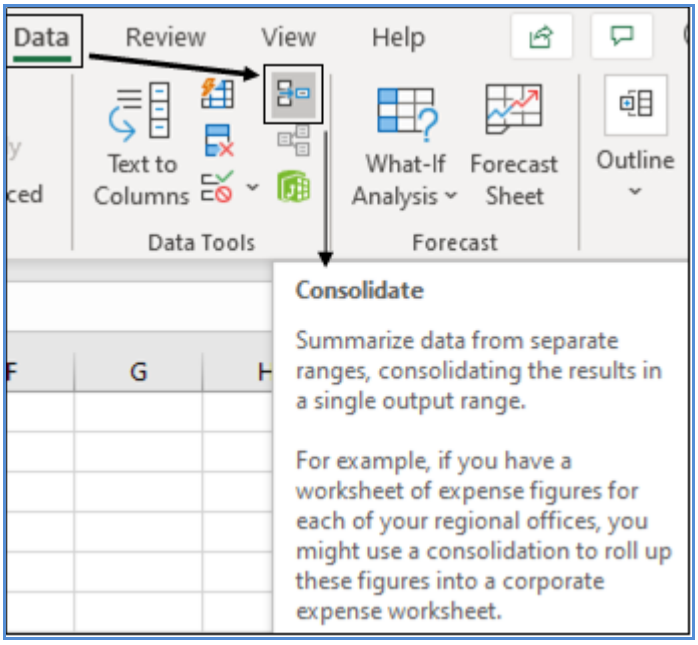


Figure 7.4.11 Consolidate is in the Data tab under Data tools

13. **Consolidate** Dialog Box appears as shown in Figure

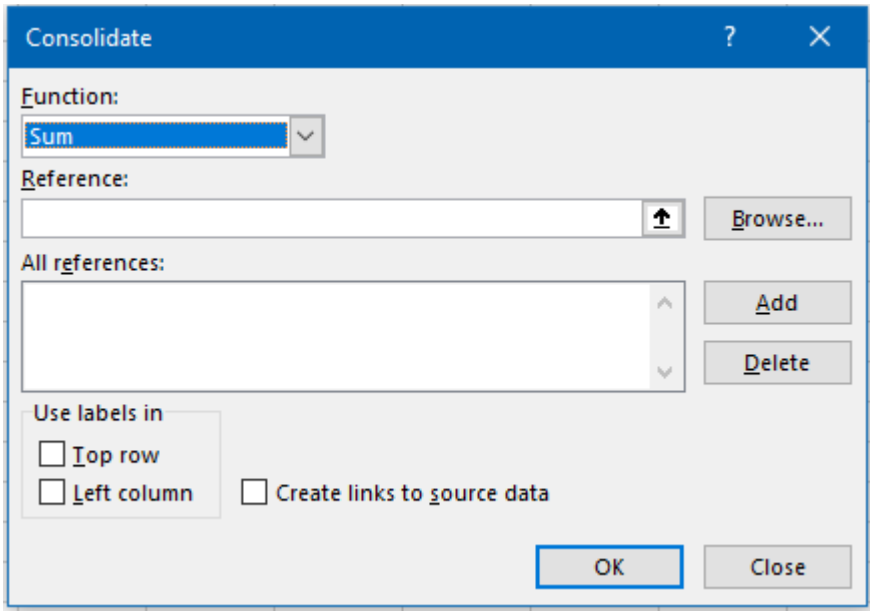


Figure 7.4.12 Use Sum under function and add the reference & other options

14. In the **Function** box, click the operation you want to use to consolidate the data. Since you want to Annual results you will use **SUM** (there are however 11 operations possible)

    - Sum, Count, Average,

- max, Min, Product,
- Count Nums, STDdev,
- StdevP, Var, VarP

15. In the **Reference** box, enter a source area you want to consolidate. If the worksheet that contains the data that you want to consolidate is in another workbook, click **Browse** to locate that workbook, and then click **OK** to close the Browse dialog box.

    In your case since you have defined **Name range** and sheets are in the same workbook, So **press F3** on the keyboard to open name ranges in **Paste Name** dialog and select *Quarter_1*, rather than selecting sheets manually, which could be hectic if you are dealing with a huge spreadsheet.

16. Click **Add**.

17. Repeat steps 14 and 15 for each source area to consolidate ie *Quarter_2, Quarter_3, and Quarter_4* respectively.

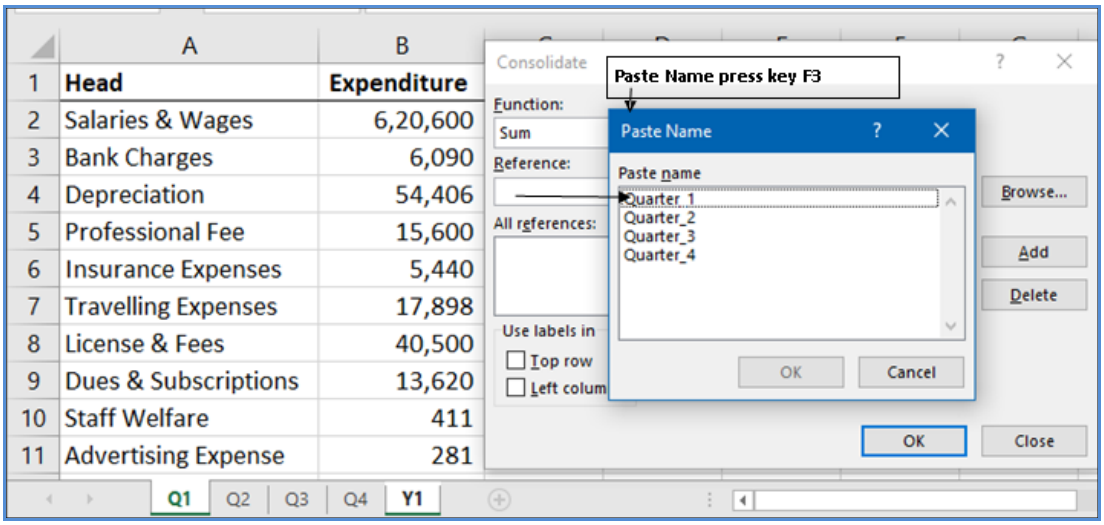


Figure 7.4.13 Press F3 to open name Ranges in paste Name dialog box

18. To update the consolidation table automatically when the source data changes, select the **Create links to source data** checkbox.

    To create links, the source and destination areas must be on different worksheets. Once you create links, you cannot add new source areas or change the source areas that are included in the consolidation.

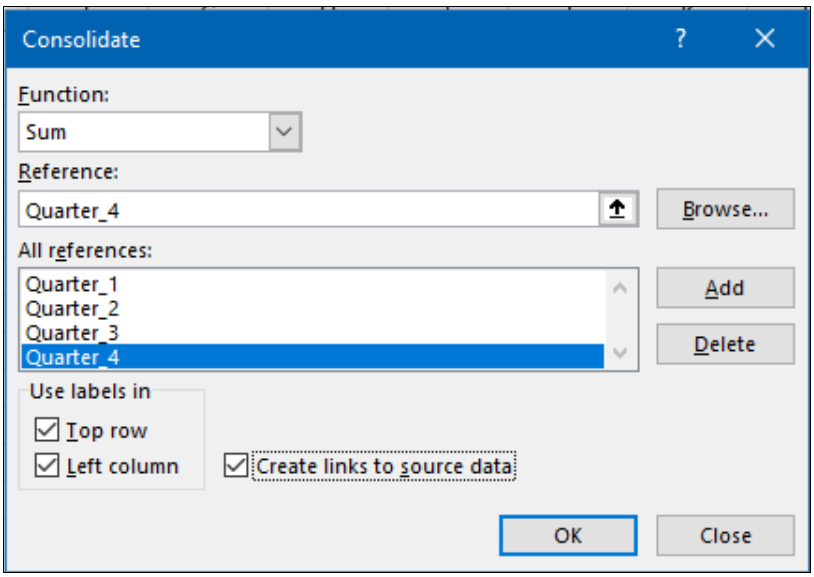


Figure 7.4.14 Adding all the references, enable create links to source data

19. Upon click OK, you will notice that the values from the sheets have added-up into the *final* sheet, as shown in the Figure below.

		A	B
	1		Expenditure
+	6	Salaries & Wages	27,99,924
+	11	Bank Charges	27,476
+	16	Depreciation	2,45,462
+	21	Professional Fee	70,382
+	26	Insurance Expenses	24,544
+	31	Travelling Expenses	80,750
+	36	License & Fees	1,82,721
+	41	Dues & Subscriptions	61,449
+	46	Staff Welfare	1,854
+	51	Advertising Expense	1,268

Figure 7.4.15 Final result after the consolidation of data

20. To set up the consolidation so that you can update the consolidation manually by changing the included cells and ranges, clear the **Create links to source data** checkbox.

21. When you consolidate by position, MS Excel does not copy the category labels in the source areas to the destination area. If you want labels for the destination worksheet, copy them or enter them manually.

Gist: You have consolidated Quarterly Expenditure to Yearly Expenditure.

Commands discussed: Data > Consolidate, Naming ranges

7.5 WHAT IF ANALYSIS

You can perform a sensitivity analysis or What-if Analysis in Excel. Something to help assess a business's performance on both a historical and projected basis is the financial model. It provides a way for the analyst to organize a business's operations and analyse the results in both vis-à-vis itself over some time ie this year vs the previous year, this quarter vs previous quarter, or maybe analyse performance against benchmarks or other industry peers.

Once you have input both historical financial results and assumptions about future performance, you can then calculate and interpret various ratio analysis, scenarios, and other operational performance metrics such as profit margins, inventory turnover, cash collections, leverage, and interest coverage ratios, among numerous others.

What-if analysis has powerful Excel tools that allow you to see what the desired result of the financial model would be under different circumstances. It allows you to select two variables, or assumptions, in the model and to see how the desired output, such as earnings per share or profit would change based on the new assumptions. It includes, **Goal seek**, **Scenario Manager**, and **Data Table**. It gives you lots of flexibility when it comes to analysis and presentation of data.

Scenarios and **Data Tables** take sets of input values and determine possible results. A data table works only with one or two variables, but it can accept many different values for those variables. A scenario can have multiple variables, but it can accommodate only up to 32 values. **Goal Seek** works differently from scenarios and data tables in that it takes a result and determines possible input values that produce that result.

7.5.1 GOAL SEEK

After every election, you are glued to the news where psephologists are analyzing elections and predicting so many votes needed to achieve the goal of getting elected. You want to achieve the target sales or you want to achieve the goal of profit. In all these cases you want to achieve a certain goal and you want to analyse inputs required for the same.

The Goal Seek feature in Excel is a what-if analysis tool that enables you to find the input values needed to achieve a goal or objective.

This **back-solves** the problem and finds the input value that satisfies your requested output value.

You can think of goals seek as the opposite of formulas. Formulas tell you what is the output of some inputs connected with an equation. Goal seek tells you what inputs you need to give in order to get a certain output.

Goal Seek requires the following Information.

- **Set cell** – The output cell
- **To value** – the target value of the output cell
- **By changing cell** - The cell that should change

Two important criteria for using Goal seek

- **The set cell must contain a formula**

- **By changing cell should never contain a formula**

CASE STUDY: As a Mortgage Analyst in a Bank you propose to give a Housing Loan to a prospective client on the following terms

- ***Loan amount Rs. 40,00,000.***
- ***ROI Annual 12%***
- ***Term 240 Months***

You calculate the EMI but your client says he cannot spare more than 42000 every month. You want to reset the term to achieve a targeted EMI of 42000 pm.

Strategy:

You first set up Formula to calculate EMI in Excel =-PMT(B4/12,B5,B3) as shown in Fig.1.10.1which is based on syntax =***PMT(rate,nper,pv,fv,type)***

- **Rate** is the interest rate for the loan. In your case, it is 12 % ie cell B4. The rate of interest should be in accordance with periods. Since the period is in months use monthly interest ie B4/12
- **Nper** is the total number of payments for the loan. In your case, it is 240 months ie cell B5
- **Pv** is the present value, or the total amount that a series of future payments is worth now; also known as the principal. In your case, it is 40,00,000 ie cell B3
- **Fv** is the future value, or a cash balance you want to attain after the last payment is made. If FV is omitted, it is assumed to be 0 (zero), that is, the future value of a loan is 0.
- **Type** is the number 0 or 1 and indicates when payments are due. If type is omitted, it is assumed to be 0 which represents at the end of the period. If payments are due at the beginning of the period, the type should be 1.

	A	B	C	D
1	Housing loan			
2				
3	Loan Amount	40,00,000.00		
4	ROI Annual	12%		
5	Term Months	240		
6				
7	EMI	₹ 44,043.45	=-PMT(B4/12,B5,B3)	

Figure 7.5.1 Setting up formula to calculate EMI

1. Go to **Goal Seek** under **What-if Analysis** under **Data Tools** in Data Ribbon as shown in Figure

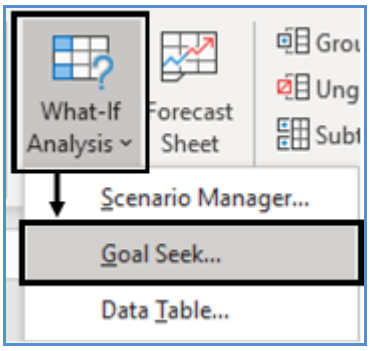


Figure 7.5.2 Goal Seek under What-if Analysis under Data Tools

2. Activate **Goal Seek** Dialog Box.
3. The "**Set cell**" must always contain a formula or a function, in this case, set it to cell B8.
4. Under "**To Value"** enter 42,000 which is the target EMI.
5. Then finally click or tab to the "**By changing cell**", which must contain a value only, and select the cell that you wish to change in this case Cell B5 which has a term for repayment as shown in Figure

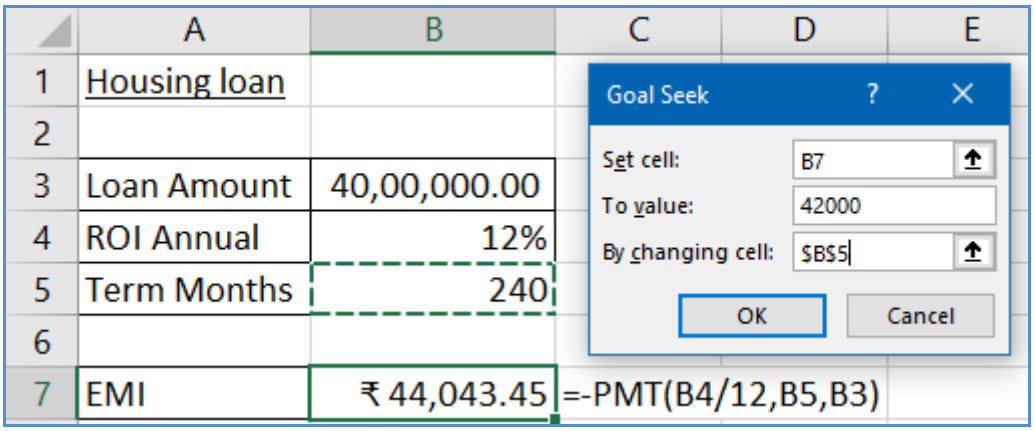


Figure 7.5.3 Enter the values in Goal Seek Dialog Box

6. Click OK
7. The Goal Seek solution in cell B5 is 306 Months as shown in Figure

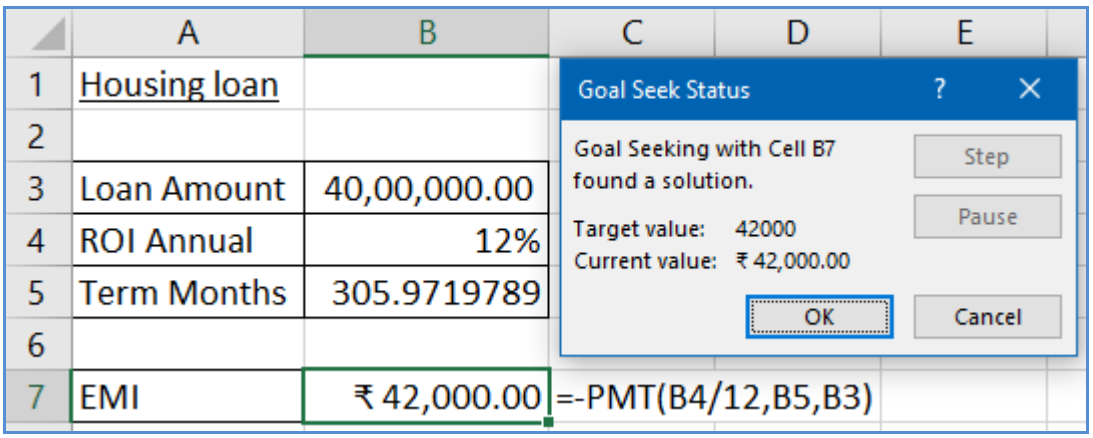


Figure 7.5.4 Goal seek gives a solution

Gist: You have created a Goal Seek solution & will propose a repayment period of 306 months.

Commands Learnt: Data > What if analysis>Goal seek, PMT Function

CASE STUDY: Your principal CA Chandiwala during a recent review, discovered that a Client has a Huge Risk which could have an impact of Rs. 1,00,00,000 and that the likelihood is currently 6%. Proposed control procedures to mitigate this risk would cost 2,00,000 and reduce the likelihood to 3%; he wants you to find the likelihood of occurrence without the control and the reduction in the expected loss if the net gain/loss is 0.

Strategy:

You can use the Goal Seek function in What-if analysis to determine the likelihood of occurrence without the control and the reduction in the expected loss if the net gain/loss is 0

1. Set up a sheet as shown in Figure

	A	B	C	D	E	F	G	H
1	Likelihood of occurrence assuming no control	Exposure	Expected Loss with no Control	Likelihood of occurrence with control	Expected Loss with Control Procedure	Reduction in Expected Loss	Cost of Control	Net Gain/Loss
2	6%	1,00,00,000	6,00,000	3%	3,00,000	3,00,000	2,00,000	1,00,000

Figure 7.5.5 Risk Assessment

2. Click on **Goal Seek** under **What-if Analysis** under **Data Tools** in Data Ribbon
3. Activate **Goal Seek** Dialog Box.
4. The "**Set cell**" must always contain a formula or a function, in this case, set it to cell $H$2.
5. Under "**To Value"** enter 0 which is the target Net Gain/Loss
6. Then finally click or tab to the "**By changing cell**", which must contain a value only, and select the cell that you wish to change in this case Cell $A$2 which has a likelihood of occurrence as shown in Figure

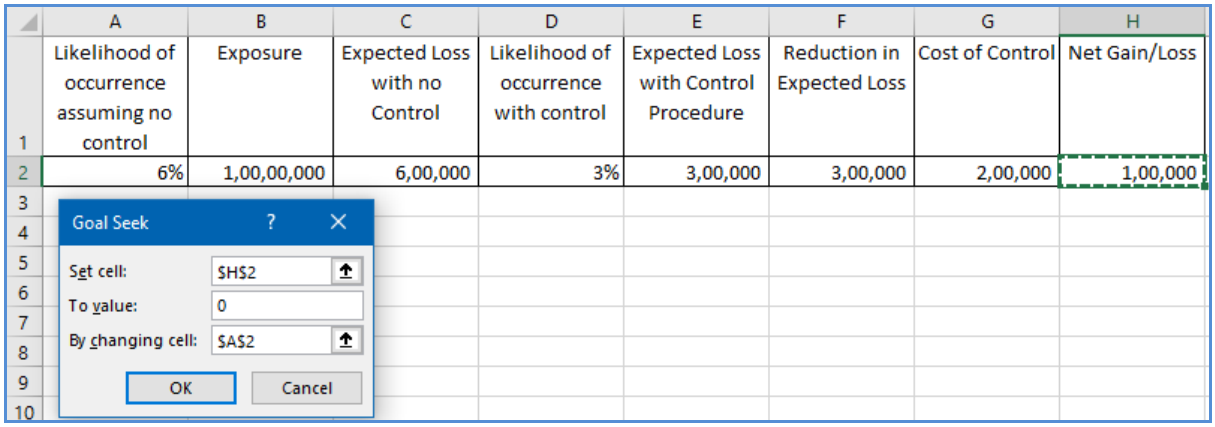


Figure 7.5.6 Set up Goal seek

7. Click OK

8. Goal seek solution is 5% as shown in Figure

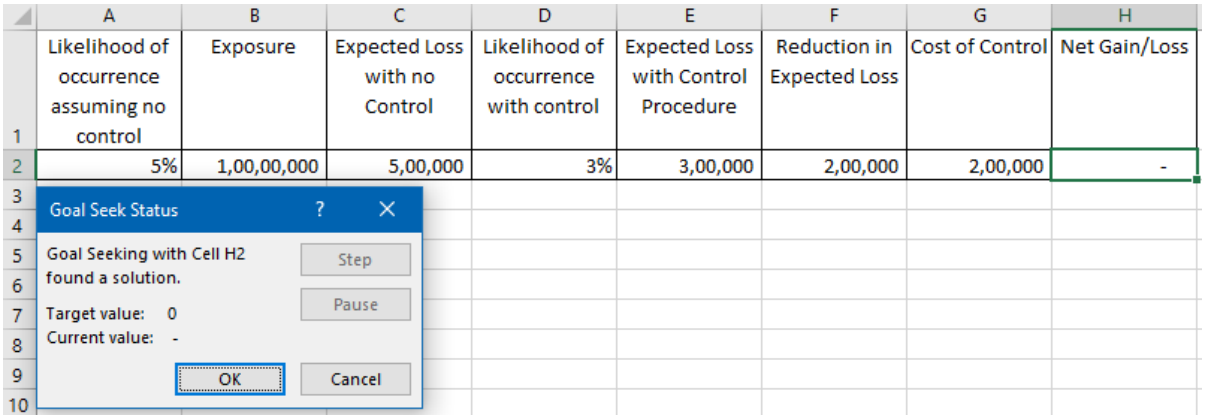


Figure 7.5.7 Result with Goal seek

Gist: You have created a Goal Seek solution where the likelihood of Occurrence is 5%.

Commands Learnt: Data > What if analysis>Goal seek

7.5.2 SCENARIO MANAGER

When preparing a budget, the Marketing and Finance departments may have different forecasts for sales. You can store each forecast as a Scenario, print them separately, or compare them side-by-side You can use Excel Scenarios to store several versions of the data in a worksheet.

As the name indicates Excel Scenario Manager enables you to create different scenarios ie different results based on different sets of input values (such as Best Case scenario, Worst Case Scenario, and Most Likely Case scenario). The key to creating the various scenarios for a table is to identify the various cells in the data whose values can vary in each scenario. You then select these cells (known as changing cells) in the worksheet before you open the Scenario Manager Dialog box.

You can use the Scenario Manager to perform analysis by varying as many as 32 input cells.

You could create lots of different scenarios for different aspects of a spreadsheet and then permutate applying them, or create more complex scenarios built from your initial ones. Scenarios can be imported and exported, so users of similar layout spreadsheets can apply scenarios someone else has created.

CASE STUDY: Your principal CA Chandiwala has created a sheet containing quarterly operating results of 4 quarters. And in the last column, he has Annual results. You have a case where the quarter to quarter growth of sales, COGS, and expenditure has been taken at 2%, 3% & 5% as shown in Figure but now he wants you to create scenarios where the growth rates may be different. So you now want to also create a Best case scenario and the worst-case scenario with assumptive growth rates 3%, 3% & 5% and 1%, 2% & 2% respectively.

	A	B	C	D	E	F	G
1		Qtr 1	Qtr 2	Qtr 3	Qtr 4	Annual	Q to Q Growth
2	Sales	20,00,000	20,40,000	20,80,800	21,22,416	82,43,216	2%
3	COGS	10,00,000	10,30,000	10,60,900	10,92,727	41,83,627	3%
4	GP	10,00,000	10,10,000	10,19,900	10,29,689	40,59,589	
5	Expenses	5,00,000	5,25,000	5,51,250	5,78,813	21,55,063	5%
6	Profit	5,00,000	4,85,000	4,68,650	4,50,876	19,04,526	

Figure 7.5.8 Quarterly Operating Results

Strategy:

You can create scenarios using the scenario manager.

1. Go to a **Scenario Manager** under **What-if Analysis** under **Data Tools** in Data Ribbon as shown in Figure

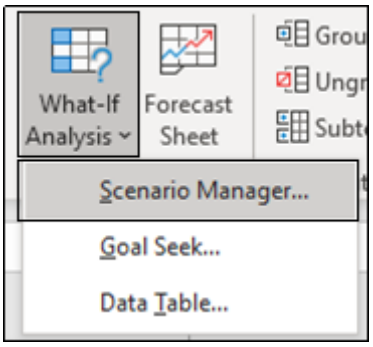
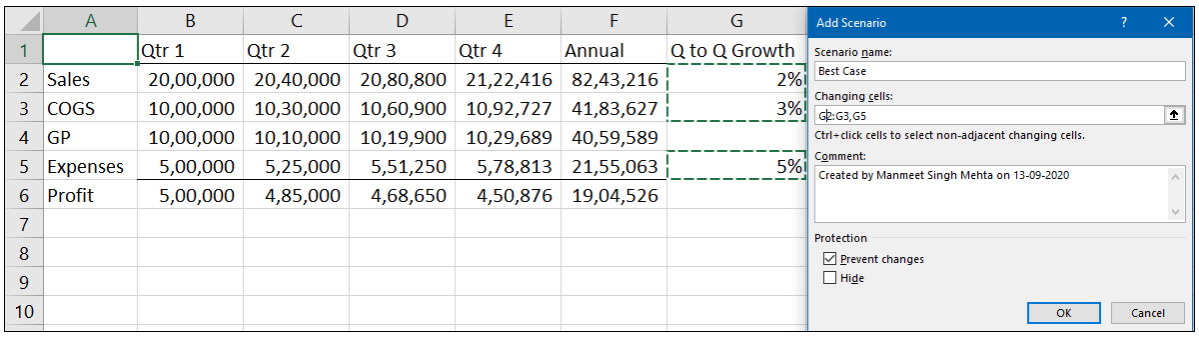


Figure 7.5.9 Scenario manager under What-if Analysis under Data Tools

2. The **Scenario Manager** dialog box appears.
3. To create a scenario, click the **Add** button.
4. The **Add Scenario dialog** box appears
5. Type the name of the scenario (Best Case, in this example) in the **Scenario Name** text box, specify the

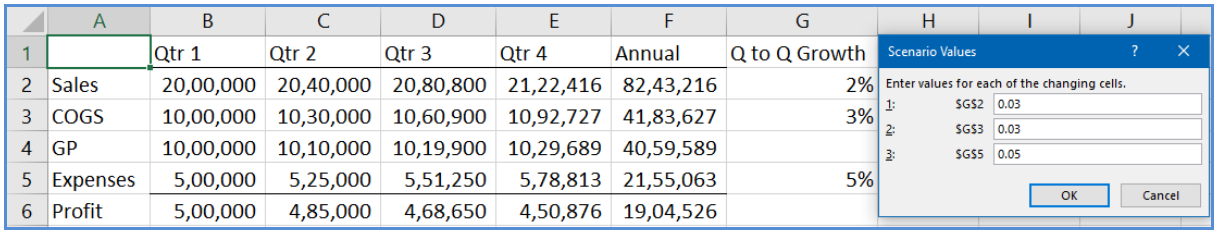
Changing Cells (if they weren't previously selected), and click **OK** as shown in Figure



	A	B	C	D	E	F	G
1		Qtr 1	Qtr 2	Qtr 3	Qtr 4	Annual	Q to Q Growth
2	Sales	20,00,000	20,40,000	20,80,800	21,22,416	82,43,216	2%
3	COGS	10,00,000	10,30,000	10,60,900	10,92,727	41,83,627	3%
4	GP	10,00,000	10,10,000	10,19,900	10,29,689	40,59,589	
5	Expenses	5,00,000	5,25,000	5,51,250	5,78,813	21,55,063	5%
6	Profit	5,00,000	4,85,000	4,68,650	4,50,876	19,04,526	
7							
8							
9							
10							

Figure 7.5.10 Enter the scenario in add scenario Dialog Box

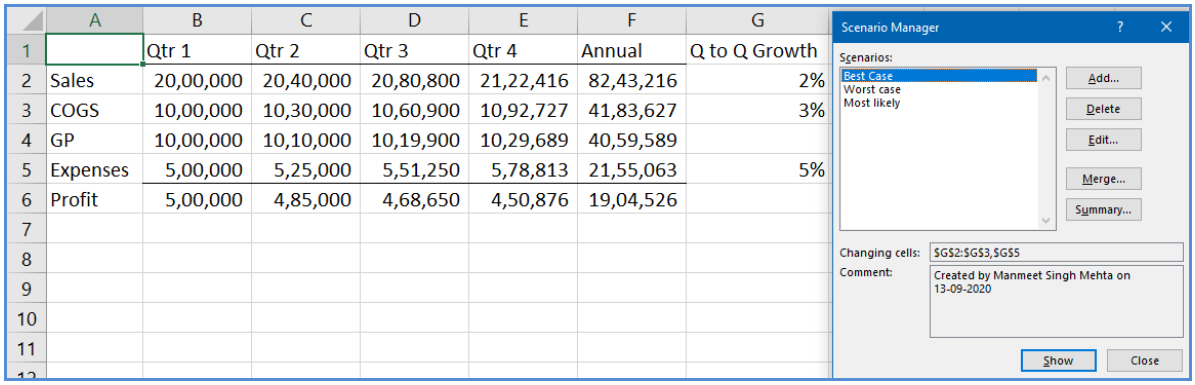
6. Excel displays the **Scenario Values** dialog box.
7. Enter the values for each of the **changing cells** in the text boxes. In this example, you would enter the following values for the Best Case scenario:
   - 0.03 in the Sales text box
   - 0.03 in the COGS text box
   - 0.05 in the Expenses text box as shown in Figure



	A	B	C	D	E	F	G
1		Qtr 1	Qtr 2	Qtr 3	Qtr 4	Annual	Q to Q Growth
2	Sales	20,00,000	20,40,000	20,80,800	21,22,416	82,43,216	2%
3	COGS	10,00,000	10,30,000	10,60,900	10,92,727	41,83,627	3%
4	GP	10,00,000	10,10,000	10,19,900	10,29,689	40,59,589	
5	Expenses	5,00,000	5,25,000	5,51,250	5,78,813	21,55,063	5%
6	Profit	5,00,000	4,85,000	4,68,650	4,50,876	19,04,526	

Figure 7.5.11 Enter values for the changing cells

7. Click the **Add** button.
8. Excel redisplays the **Add Scenario** dialog box.
9. Repeat Steps 5 through 7 to enter the other scenarios
10. When you finish entering values for the final scenario, click OK instead of Add.
11. The Scenario Manager dialog box makes another appearance, this time displaying the names of all scenarios in its Scenarios list box as shown in Figure



	A	B	C	D	E	F	G
1		Qtr 1	Qtr 2	Qtr 3	Qtr 4	Annual	Q to Q Growth
2	Sales	20,00,000	20,40,000	20,80,800	21,22,416	82,43,216	2%
3	COGS	10,00,000	10,30,000	10,60,900	10,92,727	41,83,627	3%
4	GP	10,00,000	10,10,000	10,19,900	10,29,689	40,59,589	
5	Expenses	5,00,000	5,25,000	5,51,250	5,78,813	21,55,063	5%
6	Profit	5,00,000	4,85,000	4,68,650	4,50,876	19,04,526	
7							
8							
9							
10							
11							

Figure 7.5.12 All the scenarios are displayed you can see any scenario using show

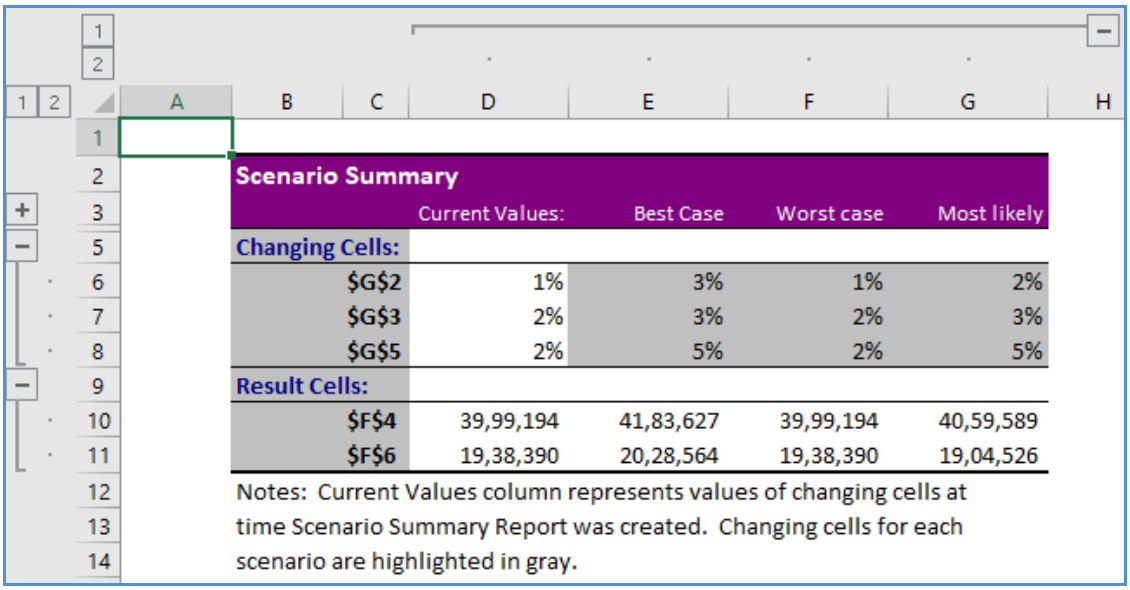
12. To have Excel plug the changing values from any scenario into the table, **click the scenario name** in the Scenarios list box and then click **Show**.

13. Click the **Close** button when you're finished with the Scenario Manager. After adding the various scenarios for a table in Your worksheet; save the workbook.

14. That way, you'll have access to the various scenarios each time you open the workbook in Excel by opening the Scenario Manager, selecting the scenario name, and clicking the Show button.

15. You can also create a summary by clicking **Summary** on **scenario Manager** Dialog Box.

16. A scenario summary dialog box would appear, by specifying the **result cells**, a **summary report** can be created as shown in Figure



Scenario Summary	Current Values:	Best Case	Worst case	Most likely
Changing Cells:				
$G$2	1%	3%	1%	2%
$G$3	2%	3%	2%	3%
$G$5	2%	5%	2%	5%
Result Cells:				
$F$4	39,99,194	41,83,627	39,99,194	40,59,589
$F$6	19,38,390	20,28,564	19,38,390	19,04,526

Notes: Current Values column represents values of changing cells at time Scenario Summary Report was created. Changing cells for each scenario are highlighted in gray.

Figure 7.5.13 Summary of all the scenarios

Gist: You have created Scenarios for different Growth rates.

Commands Learnt: Data > What if analysis>Scenario Manager

Food for Thought: The **Scenario PivotTable Report** option on the **Scenario Summary dialog box** presents the scenario results in a PivotTable format.

7.5.3 DATA MODELS

Often in Forecasting, you have built a complex formula that is dependent on many variables and you want to know how changing those inputs will change the results. What-if analysis data table is a tool through which you can observe all possible outcomes in one go rather than testing each variable individually.

There are two types of Data Tables – **One Way and Two Way**. A **data table** cannot accommodate more than two variables. Although it is limited to only one or two criteria (one for the row input cell and one for the column input cell), each criterion can include as many different variable values as you want. If you want to analyze more than two variables, you should use scenarios discussed earlier in 7.5.2.

Data tables are pretty useful when formula depends on many values, and you would like to experiment with different combinations of inputs and compare the results.

The data table is in the **Data tab > Forecast group**, click the **What-If Analysis** button, and then click **Data Table**.

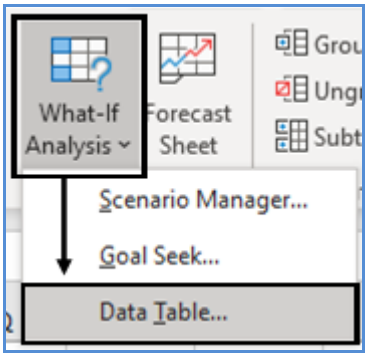


Figure 7.5.14 Data Table

CASE STUDY: Your client wants to obtain a loan of Rs. 4,00,00,000 at 8% pa interest and wants your advice on the period of the loan, His question is if by extending the period are the reduction in payments worth the additional length of the loan.

Strategy:

You can use the **Data Table** feature in Excel for this.

1. You start by creating a simple Payment function to calculate the payment amount of a loan given a loan amount, interest rate, and the number of periods.

	A	B	C
1	Loan Amount	4,00,00,000	
2	ROI	8%	Per Annum
3	Period	25	Years
4	EMI	**3,08,726.49**	=-PMT(B2/12,B3*12,B1)

Figure 7.5.15 EMI through PMT Function

2. You will now do a simple **What-If analysis** to see what the EMI will be depending on the period.
3. You are going to create a column-oriented data table
4. You will type your variable values in a column (D3:…..) and leave at least one blank column to the right for the outcomes.

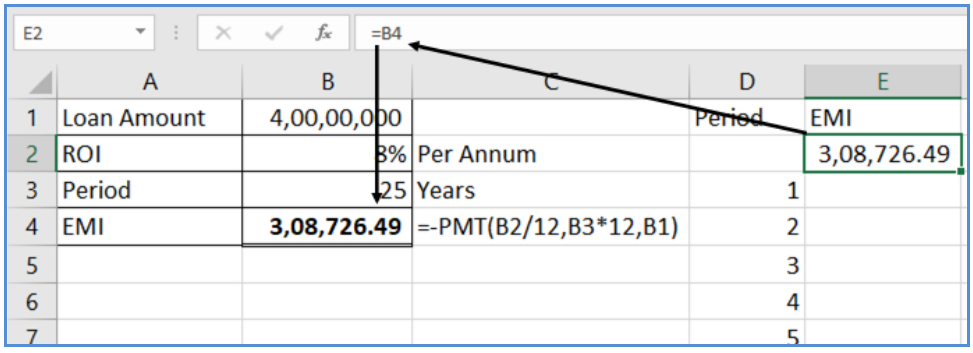


Figure 7.5.16 Set up the data variables

5. Select the data table range, including your formula, variable values cells, and empty cells for the results (D2:E27).
6. Go to the Data tab > Data Tools group, click the What-If Analysis button, and then click Data Table…

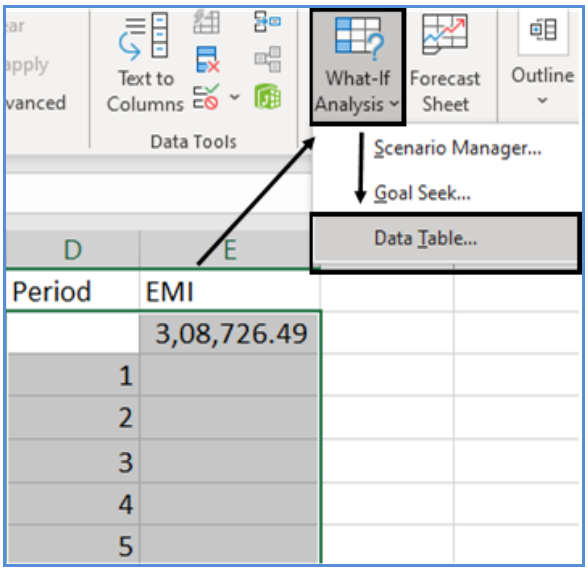


Figure 7.5.17 Data Table

7. In the **Data Table** dialog window,
8. click in the **Column Input** cell box (because your Investment values are in a column),
9. Select the variable cell referenced in your formula. In this example, you select B3 that contains the period.

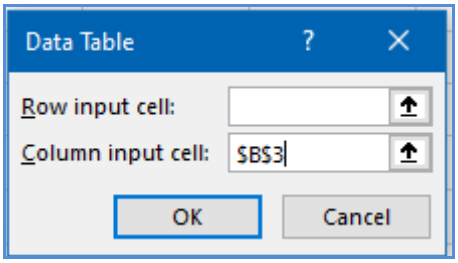


Figure 7.5.18 Data table Dialog box

10. Click OK,
11. Excel will immediately populate the empty cells with outcomes corresponding to the variable value in the same row.
12. Apply the desired number format to the results.

	A	B	C	D	E
1	Loan Amount	4,00,00,000		Period	EMI
2	ROI	8%	Per Annum		3,08,726.49
3	Period	25	Years	1	34,79,537.16
4	EMI	**3,08,726.49**	=-PMT(B2/12,B3*12,B1)	2	18,09,091.66
5				3	12,53,454.62
6				4	9,76,516.89
7				5	8,11,055.77
8				6	7,01,329.62
9				7	6,23,448.58
10				8	5,65,467.17
11				9	5,20,748.60

Figure 7.5.19 EMIs for all the periods

13. This table tells you that the same loan amount will require a monthly payment of Rs. 976516 to pay the loan off in just 4 years, or a monthly payment of Rs. 565467 to repay the loan in just 8 years.

14. For a better visualisation lets plot it on a Line Chart.

15. Select the data

16. Insert > Chart

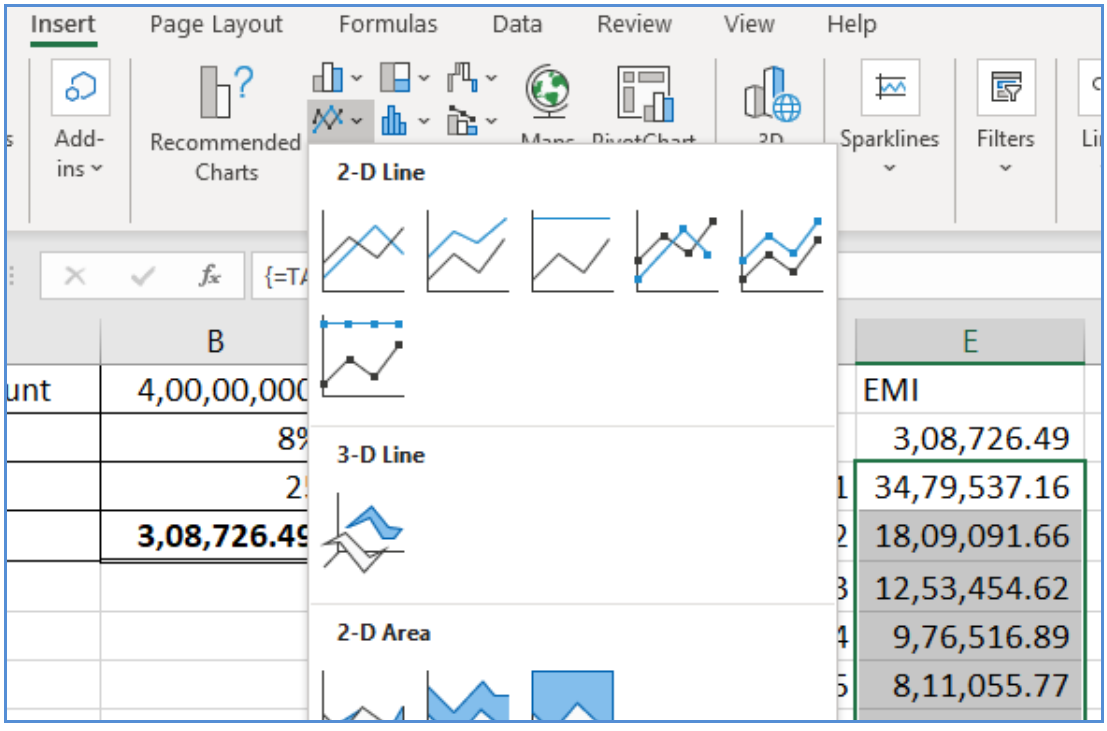


Figure 7.5.20 Insert Chart

17. The resultant Chart is as shown in Figure

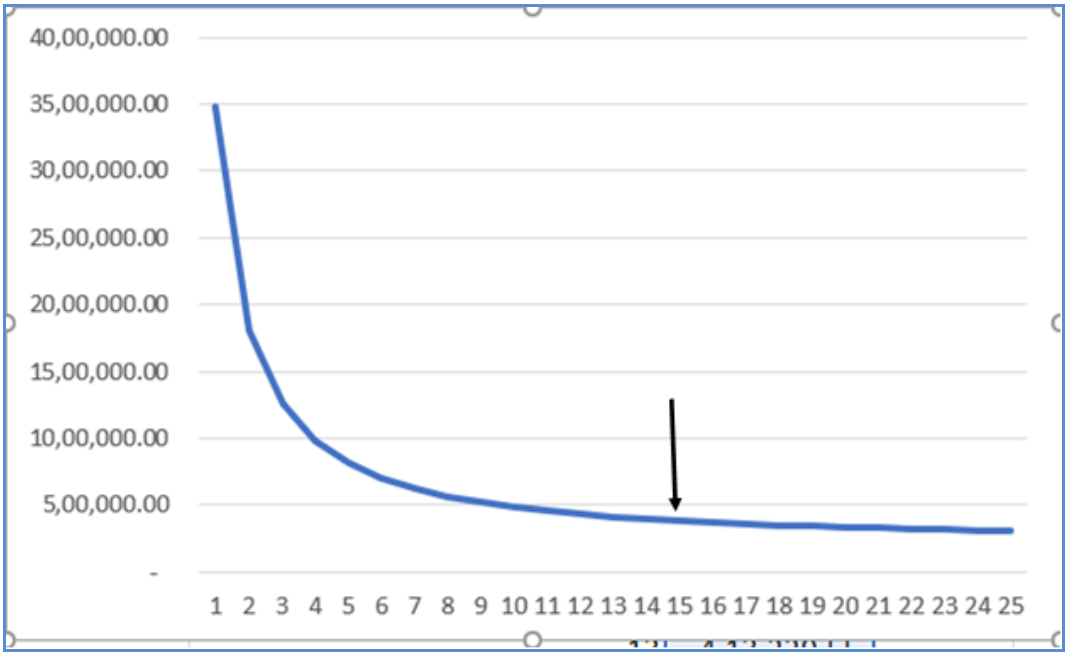


Figure 7.5.21 EMI plotted on a Chart

18. You will observe that the line flattens out at around 15 years, additional years do not have a declining impact on EMI.

19. You will therefore advise your client to keep his loan at 15 years since the reduction in EMI is not worth the additional length of the loan.

Gist: You have created a Data table for different Loan Periods.

Commands Learnt: Data > What if analysis>Data tables

7.6 SOLVER

In many situations, you want the best way to do something. Excel Solver can solve problems for you. That simple!

The solver is an Excel add-in that can solve problems by enabling a Target cell to achieve some goal. This goal may be to minimise, maximise, or achieve some target value. It solves the problem by adjusting several input cells according to a set of criteria or constraints which are defined by the user.

The solver is a planning and analysis tool that enables users to find optimal solutions for Excel models that maximize profit, or minimize cost or risk, by automatically adjusting multiple input cells. It is used in a wide range of industries, with a common thread of finding the best way to allocate scarce resources.

The solver was developed by Frontline Systems Inc. and licensed to Microsoft. Frontline offers more powerful Solvers, for use inside and outside Excel, directly to end-users via its Website http://www.solver.com.

For a given problem, an excel solver can run various permutations and combinations and find out the best possible solution for you. It is like goal seek, but better

For Goal Seek with more than one changing cell use solver. Wherever complex trial and error analysis is required solver should be used. The solver can alter a formula not just to produce a set value but also to maximize or minimize the results. Solver helps you answer optimization problems elegantly.

An optimization model has three parts:

Target cell - represents the objective or goal. Multiple Target can also be there

Changing cells-which can be changed/ adjusted to optimize Target

Constraints- Logical conditions within which a solution is desired.

The solver is available in Excel as an **Add-in**.

To install Solver, click the **File** tab, click **Options**, and then click **Add-Ins**.

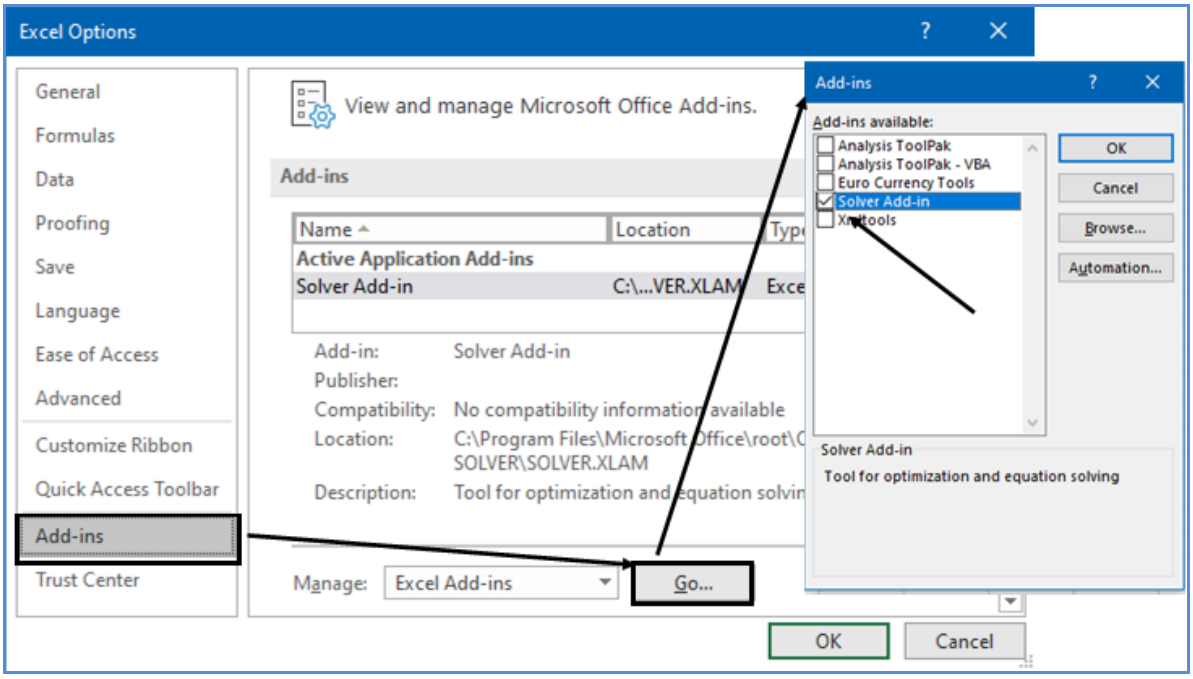


Figure 7.6.1 File Tab > excel Options>Add-ins

In the Manage box at the bottom of the dialog box, select **Excel Add-Ins**, and click **Go**. Select the **Solver Add-In** checkbox in the Add-Ins dialog box, and click OK as shown in Figure

After Solver is installed, you can run Solver by clicking **Solver** in the **Analysis** group on the **Data tab**.

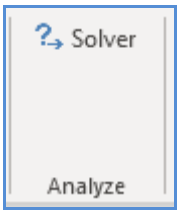


Figure 7.6.2 Solver

The Solver Parameters dialog box will appear as shown in Figure

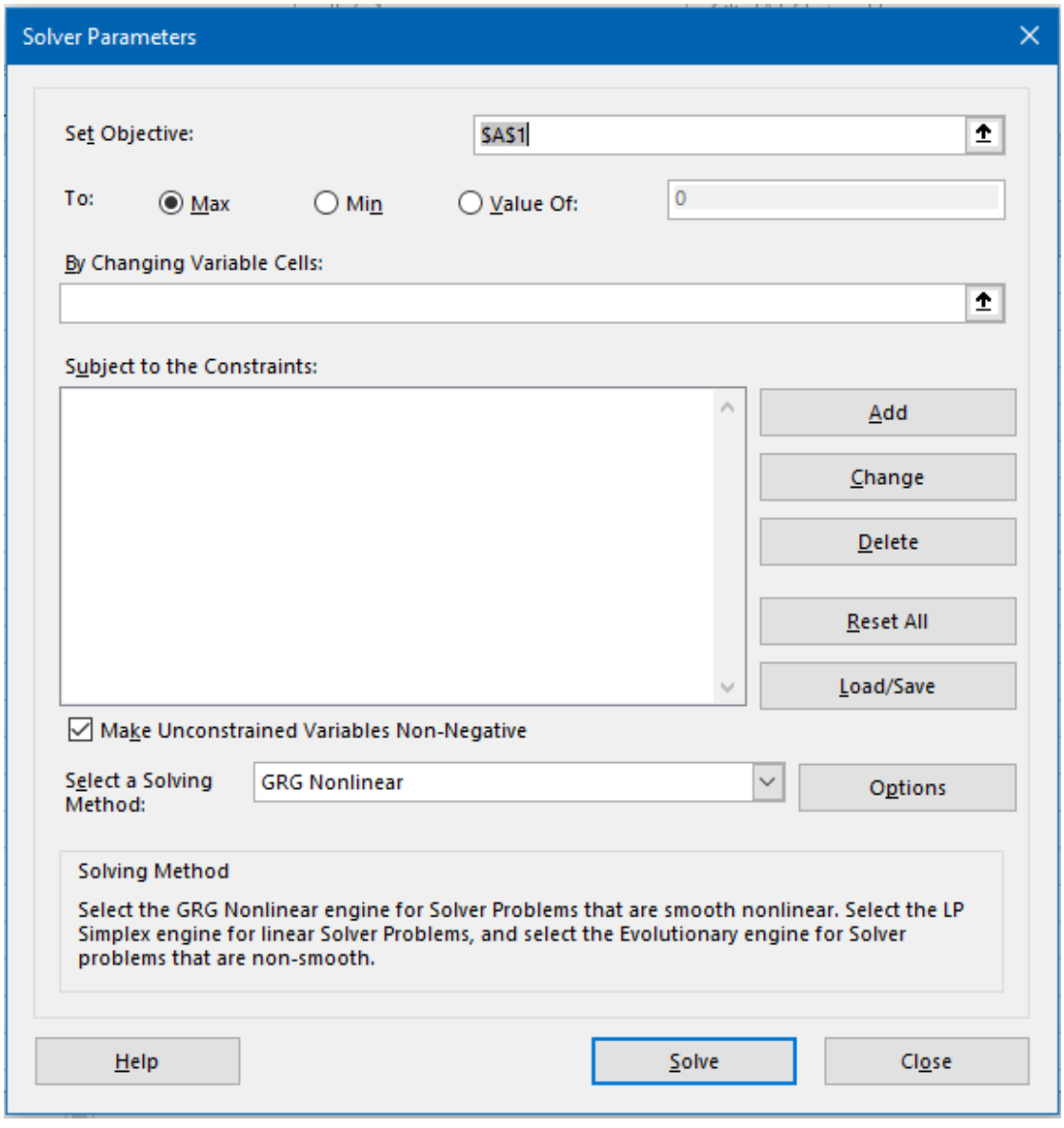


Figure 7.6.3 Solver Parameter Dialog Box

Solver provides you with the following solving methods:

The Simplex LP engine is used to solve linear optimization problems.

The GRG Nonlinear engine is used to solve optimization problems in which the target cell and/or some of the constraints are not linear

The Evolutionary engine is used when your target cell and/or constraints contain non-smooth functions that reference changing cells.

Solver searches all **feasible solutions** and finds the one that has the "best" target cell value (the largest value for maximum optimization or the smallest for minimum optimization). Such a solution is called an **optimal solution**.

CASE STUDY: As CFO for Icecreamwallah Pvt. Ltd that produces six different flavours of Ice creams at its New Delhi plant. Production of each product requires labor and Milk as its primary raw material. You set up the data as shown in Figure relating to different Favours of ice-creams. Row 4 shows the production in litres of each flavor, Row 5 shows the hours of labor needed to produce a Litre of each Flavour, and row 6 shows the Litres of Milk needed to produce a Litre of each Flavour. For example, producing a Litre of Butterscotch Ice-cream requires 5 hours of labor and 0.8 Litres of Milk. For each Flavour, the Sales price per Litre is given in row 7, the cost per Litre is given in row 8, and the profit contribution per Litre is given in row 10. The monthly demand for each icecream is given in row 9. This month, 5000 hours of labor and 2000 Litres of Milk are available. How can Icecreamwallah maximize its monthly profit?

	A	B	C	D	E	F	G
1	**Icecreamwallah P. Ltd.**						
2							
3	Available	Product	Butterscotch	Pineapple	Choclate	Choconut	Vanilla
4		Production in Ltrs	0	0	0	0	0
5	5000	Labor	5	6	3	4	2.5
6	2000	Milk Required in Ltr	0.8	0.9	0.85	0.7	0.95
7		Unit price	Rs. 10.00	Rs. 11.00	Rs. 12.00	Rs. 14.00	Rs. 7.00
8		Cost	Rs. 6.20	Rs. 6.00	Rs. 7.60	Rs. 9.80	Rs. 2.20
9		Demand	1020	952	1052	1056	1563
10		Unit profit cont.	Rs. 3.80	Rs. 5.00	Rs. 4.40	Rs. 4.20	Rs. 4.80
11							
12	0	Labor Used	0	0	0	0	0
13	0	Milk Used	0	0	0	0	0
14							
15		Profit	Rs. -				

Figure 7.6.4 Product-wise data of Icecreamwallah P. Ltd.

Strategy:

You can solve this problem using **Solver** which is under Analysis under data tab.

1. The key to the solution is to compute the **resource usage** and **profit associated with any product mix**.
2. In row 12 you set up labour for each Flavour by using the formula =C4*C5 and use Sum in A12, Similarly set up Milk Used in row 13. Alternately you could also use the Function SUMPRODUCT.

3. To calculate Total profit you can use SumProduct as =SUMPRODUCT(C10:G10,$C$4:$G$4) which simply means the formula =C10*$C$4+D10*$D$4+E10*$E$4+F10*$F$4+G10*$G$4 or which means (Flavour 1 profit per Litre)*(Flavour 1 Litres produced) + (Flavour 2 profit per Litre)*(Flavour 2 Litres produced) + ...(Flavour 6 profit per Litre)*(Flavour 6 Litres produced).

4. For considering solver the inputs are as follows:

   - **Target**- Maximise profit ie cell C15
   - **Changing Cells**- Number of Litres produced
   - **Constraints**- Following Constraints are there
   - Not to use more labour than available ie Total labour ie A12 is not more than Available Labour ie A5.
   - Not to use more Milk than available ie Total Milk ie A13 is not more than Available Milk ie A6.
   - Not to produce more than Demand
   - Production cannot be negative

5. Now **click Solver** under Analysis group under Data tab.

6. Solver parameter box will appear.

7. Click the set objective box and then select the profit cell (cell C15). Click the **By Changing Variable Cells** box, and then point to the range C4:G4, which contains the Litres produced of each Flavour. As shown in Figure

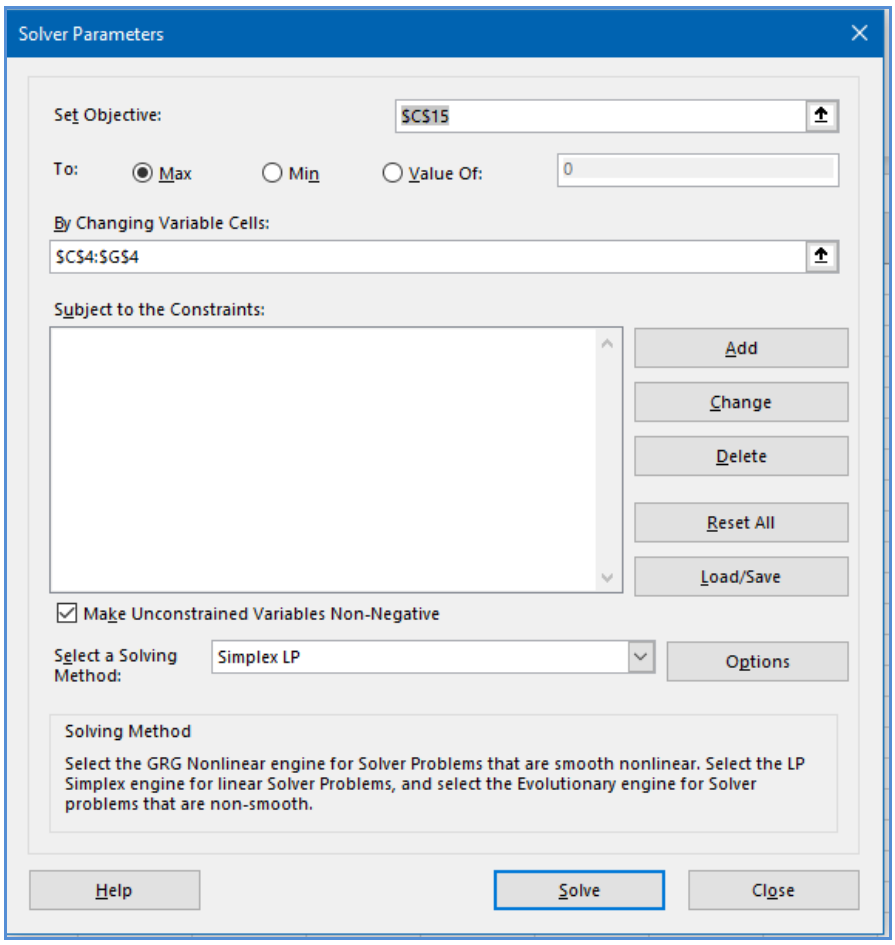


Figure 7.6.5 Solver Dialog Box

8. Now **Add Constraints** as follows **click add** and Add constraint Dialog box will appear To add the resource usage constraints, click the Cell Reference box, and then select the cell $A$12:$A$13 Select <= from the middle list. Click the Constraint **box**, and then select the cell $A$5:$A$6. The Add Constraint dialog box should now look as shown in Figure

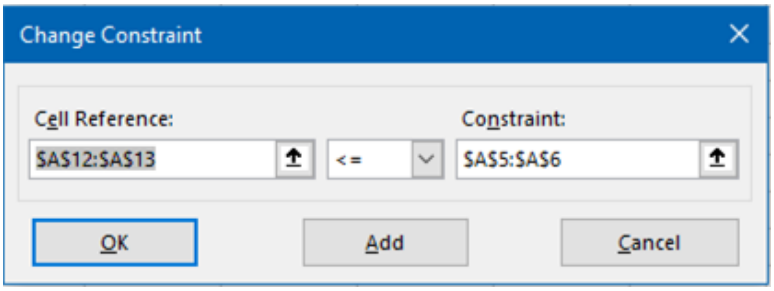


Figure 7.6.6 Add Constraint Dialog Box with resource constraints

9. Click **add** to add another constraint ie $C$4:$G$4 <= $C$9:$G$9 which ensures that production will always be less than demand.

10. Selecting the **Make Unconstrained Variables Non-Negative** checkbox ensures that all the changing cells are forced to be greater than or equal to 0.

11. Next, choose **Simplex LP** from the **Select A Solving Method** list. You choose the Simplex LP engine because the product mix problem is a special type of Solver problem called a *linear model*. Essentially, a Solver model is linear under the following conditions:

12. The target cell is computed by adding together the terms of the form *(changing cell)*(constant*).

13. Each constraint satisfies the linear model requirement. This means that each constraint is evaluated by adding together the terms of the form *(changing cell)*(constant)* and comparing the sums to a constant.

14. This problem is linear because the target cell is calculated as *(Flavour 1 profit per Litre)*(Flavour 1 Litres produced) + (Flavour 2 profit per Litre)*(Flavour 2 Litres produced) + ...* which is of the form *(changing cell)*(constant).*

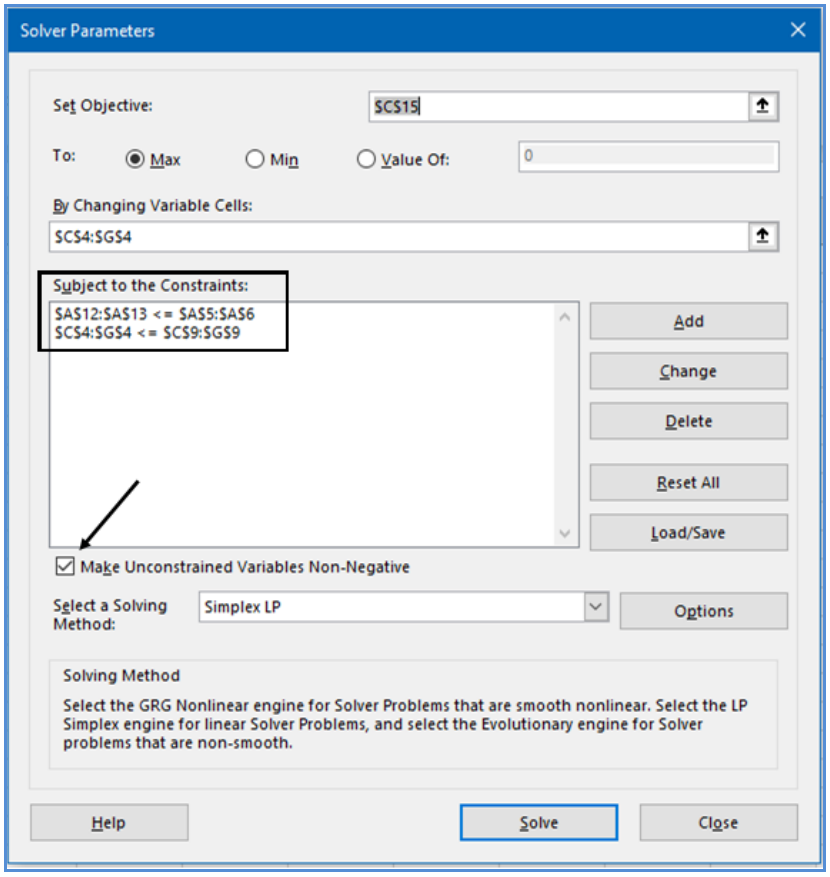


Figure 7.6.7 Solver dialog Box with constraints & other options

15. After you click **Solve**, Solver calculates an optimal solution (if one exists) for the product mix model. An optimal solution to the product mix model would be a set of changing cell values (Litres produced of each Flavour) that maximizes profit over the set of all feasible solutions.

16. The result you get is as shown in Figure.

	A	B	C	D	E	F	G
1	**Icecreamwallah P. Ltd.**						
2							
3	Available	Product	Butterscotch	Pineapple	Choclate	Choconut	Vanilla
4		Production in Ltrs	0	0	364.16667	0	1563
5	5000	Labor	5	6	3	4	2.5
6	2000	Milk Required in Ltr	0.8	0.9	0.85	0.7	0.95
7		Unit price	Rs. 10.00	Rs. 11.00	Rs. 12.00	Rs. 14.00	Rs. 7.00
8		Cost	Rs. 6.20	Rs. 6.00	Rs. 7.60	Rs. 9.80	Rs. 2.20
9		Demand	1020	952	1052	1056	1563
10		Unit profit cont.	Rs. 3.80	Rs. 5.00	Rs. 4.40	Rs. 4.20	Rs. 4.80
11							
12	5000	Labor Used	0	0	1092.5	0	3907.5
13	1794.3917	Milk Used	0	0	309.54167	0	1484.85
14							
15		Profit	Rs. 9,104.73				

Figure 7.6.8 Solution given by solver optimising the production of product mix

17. Solver also prompts you if you want to save it as a scenario.

Gist: You have used Solver for Maximising Profit within given constraints

Commands Learnt: Data > Solver

7.7 SUMMARY

Workbooks are meant to be shared but it is important that anybody using your workbook accidentally or intentionally messes with your workbook.

In this chapter, you learned to protect cells, worksheets, and workbooks so that nobody can tamper with your Excel Sheets. You also understood how you can use Get and Transform Data Functionality to connect to various data sources and also transform and clean data.

You explored how various features under what-if Analysis can be used like Goal Seek to Reverse calculations, Data Table for Sensitivity analysis, and Scenario Manager for Comparison of scenarios.

Lastly, you learned the use of Solver to find solutions to optimization problems.

REFERENCES

[1] Michael Alexander, Richard Kusleika, John Walkenbach, 'Excel 2019 Bible', Wiley,2018

[2] Greg Harvey, 'Excel 2019 All-in-One for Dummies, John Wiley & Sons,2018

[3] www.chandoo.org

[4] www.solver.com.

CHAPTER

8 GOOD TO KNOW FEATURES

LEARNING OBJECTIVES

- Understanding how to reduce the file size of excel files
- Understanding Linking of excel with Word and PowerPoint
- Gain an understanding and usage of Macros
- Understand the Form Controls in Excel
- Understanding Artificial Intelligence in Excel
- Understand Google Sheets vis-a-vis Excel

8.1 INTRODUCTION

Excel is a practical tool that is of immense utility for any Finance professional. In this chapter, you will understand various features that make Excel an indispensable application. Many of you would have encountered really large Excel files that become unwieldy to mail and operate consequently making the working very slow leading to inefficiencies, you will learn to reduce the size of large files.

These days most of the Excel files you will find the Extension .XLSX, but there are other file extensions too like .XLS, .XLSB, .XLSM etc

.XLSX -This the "default" file extension. Supports up to 2^20 (1048576) rows and 2^14 (=16384) columns, They are secure since they cannot store VBA Macros.

.XLSM - The default Excel file type for workbooks with VBA macros. Uses the open XML standard

.XLSB-An XLSB file is a Microsoft Office Excel Binary file. Information inside XLSB file is stored as binary data compared to XSLX files where data is stored as XML files which are then compressed. The size of .XLSB is significantly smaller as compared to .XLSX. therefore they open faster.

.XLS- This is the older format, They may store VBA Macros so are unsafe. Work with Excel 2003 and prior versions.

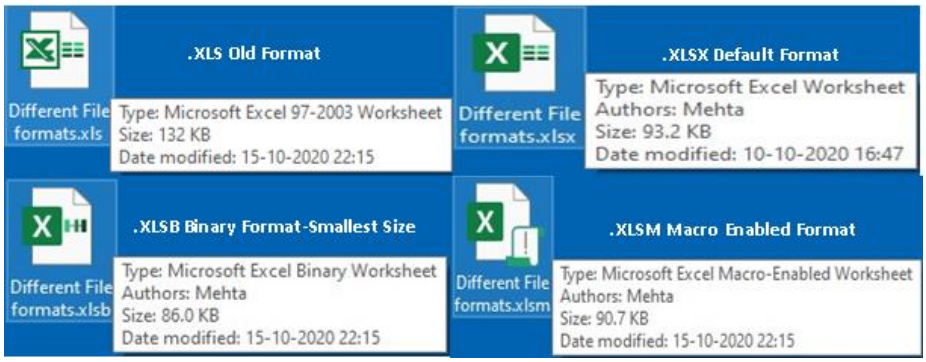


Figure 8.1.1 Different Excel File formats with Size- .XLSB is smallest

Many times you want to link the data or Charts in Excel with PowerPoint or in word documents. At the same time, you want the Data or Chart to refresh itself in PowerPoint or Word whenever Data in excel is changed. All this is possible through Paste Special with Links.

If you want to avoid the chore of doing monotonous work like data entry or some formatting which you need repeatedly. Excel offers some excellent options to automate in the form of Macros which are small programs that automatically perform repetitious steps.

Excel offers a host of facilities through which external documents can be linked and automation is possible.

Artificial intelligence is pervading every aspect of life, Excel is no exception. Today there is a multitude of Artificial Intelligence capabilities in excel like New smart data types, Flash fill, and Ideas which can provide many insights into data on the fly.

For Finance professional Spreadsheet applications are must-haves, Both Google Sheets and Microsoft Excel are well-known spreadsheet applications and have their pros and cons, which you will understand in this chapter.

8.2 REDUCE SIZE OF EXCEL FILES

You have probably felt the challenge of emailing large files in Microsoft Excel. Excel workbooks have been known to become painfully slow and heavy as more data/formulas/images are added to it.

These bloated files take too long to open and take forever to update or change formulas every time you change a value. These wastes your time and productivity deteriorated.

There is a varied range of ways to fix this and keep the file size in check., some of these ways are as discussed below.

- **Remove unnecessary sheets and data**

  Many times, you have unnecessary worksheets which are not needed, you need to delete those sheets.

- **Check for and delete unnecessary hidden sheets**

  Your file may contain hidden sheets, you can reduce the file size of your Excel workbook by unhiding all sheets, and checking if you still need that data or not.

- **Remove blank space in your sheets**

  Your files can bloat if you have unnecessary used space in your sheets. Excel has a "used range" for every sheet in your workbook. If you have a new excel file it may have the used range as only cell A1. When you keep on working on a file, it will increase to the rightmost column and lowest row that you have edited or formatted, you can check this with CTRL + END to go to the last cell used thus check and remove unnecessary blank space.

- **Do not format raw data**

  Formatting Data increases size, you may have some data in a sheet which is required only for calculation not presentation. Remove as much formatting as possible from this data.

- **Remove Conditional formatting**

  Just like regular formatting, conditional formatting also increases the file size. Use it only where required.

- **Remove unnecessary formulas**

  Too many formulas increase size, If there are unnecessary formulas convert them to values by using Paste Special as Value

- **Avoid using volatile formulas**

  There are seven volatile functions in Excel– ie, they calculate every time any cell is updated in the workbook. These functions are RAND(), NOW(), TODAY(), OFFSET(), CELL(), INDIRECT() and INFO(). You need to reduce their use to the minimum required

- **Convert the file to XLSB.**

  Convert the file to Binary format i.e. with an extension .xlsb.

  Go to **Save as** and select **XLSB** as the file type.

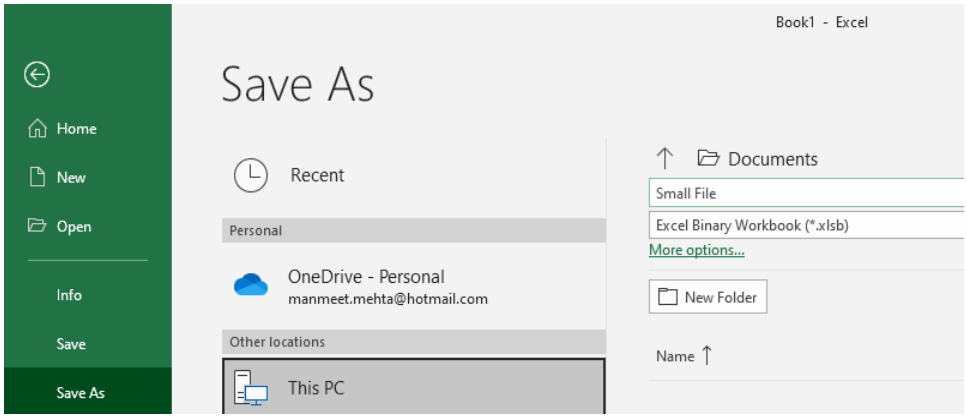


Figure 8.2.1 Save as XLSB

When you save an Excel file (.xlsx or .xlsm), it gets saved in the XML format, whereas XLSB saves it as a binary (.bin) file.

8.3 LINKING WITH OTHER FILE TYPES

8.3.1 DYNAMICALLY UPDATE PPT OR WORD WITH EXCEL FILES

You can dynamically update Charts from Excel to a PowerPoint or even a word Document. This can be a big-time saver if you regularly prepare a presentation or issue reports that need to include up-to-date dynamic charts.

***CASE STUDY*: Your principal CA Chandiwala is to give a presentation and in the PPT he wants to include a chart which is based on the monthly Revenue which is in an excel chart as shown in Figure. He also wants that the chart should dynamically update on the powerpoint presentation whenever there is any change in Excel. He also wants the Excel chart to be included in a report that is a Word Document, where again it should be dynamically linked. Your challenge is to build a Dynamically updating chart in PPT from Excel File.**

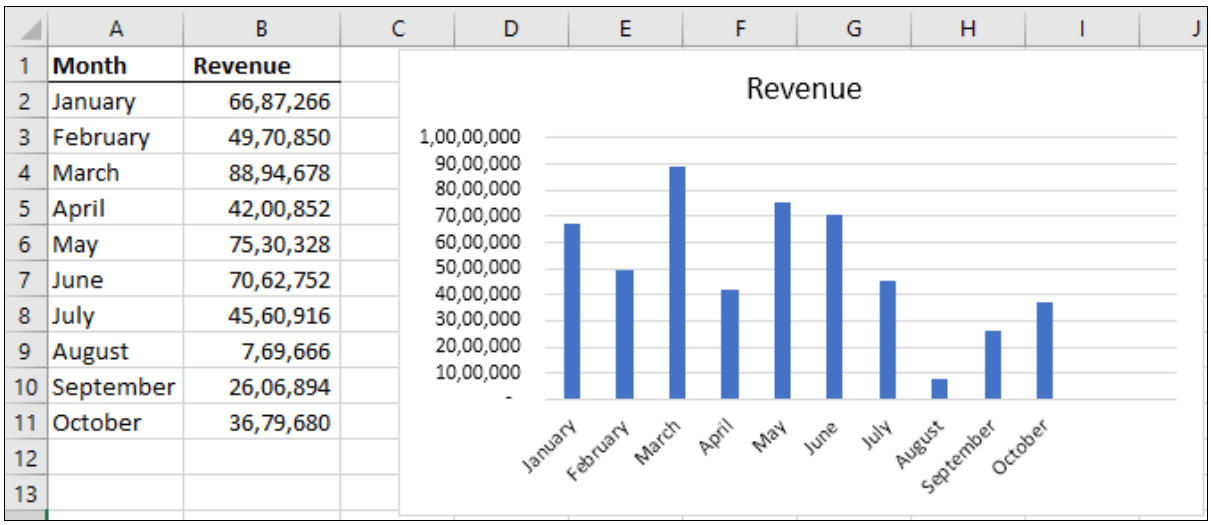
	A	B
1	**Month**	**Revenue**
2	January	66,87,266
3	February	49,70,850
4	March	88,94,678
5	April	42,00,852
6	May	75,30,328
7	June	70,62,752
8	July	45,60,916
9	August	7,69,666
10	September	26,06,894
11	October	36,79,680

Figure 8.3.1 Monthly Revenue

Strategy:

You can use the linking feature in PowerPoint to link to Excel Sheet.

1. Create a chart in Excel from the data



	A	B
1	**Month**	**Revenue**
2	January	66,87,266
3	February	49,70,850
4	March	88,94,678
5	April	42,00,852
6	May	75,30,328
7	June	70,62,752
8	July	45,60,916
9	August	7,69,666
10	September	26,06,894
11	October	36,79,680
12		
13		

Figure 8.3.2 Create Chart

2. Copy the Chart.
3. Open the PowerPoint
4. Use Paste Special option and paste the copied chart into PowerPoint with data link.

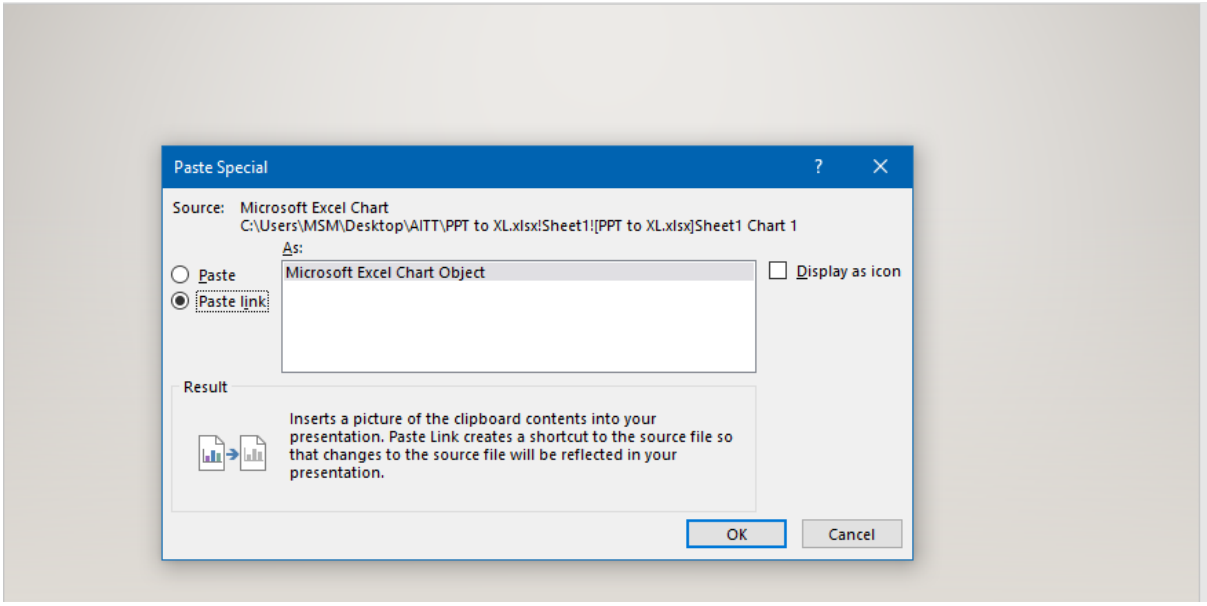


Figure 8.3.3 paste special and link data

5. Use Paste Special option and paste the copied chart into PowerPoint with data link.
6. Chart is pasted

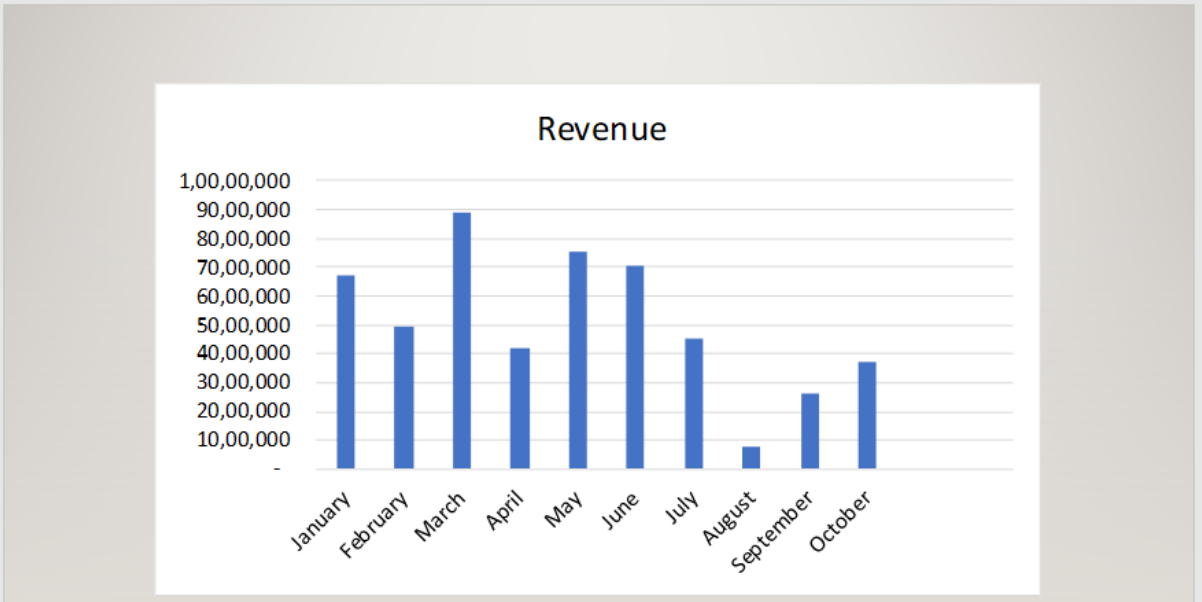


Figure 8.3.4 Chart is pasted

7. Now save the PowerPoint and Excel files and close them.
8. Re-open the Excel file add data relating to November and December and save it.

	A	B	C	D	E	F	G	H	I	J
1	**Month**	**Revenue**								
2	January	66,87,266								
3	February	49,70,850								
4	March	88,94,678								
5	April	42,00,852								
6	May	75,30,328								
7	June	70,62,752								
8	July	45,60,916								
9	August	7,69,666								
10	September	26,06,894								
11	October	36,79,680								
12	November	24,01,628								
13	December	78,26,567								

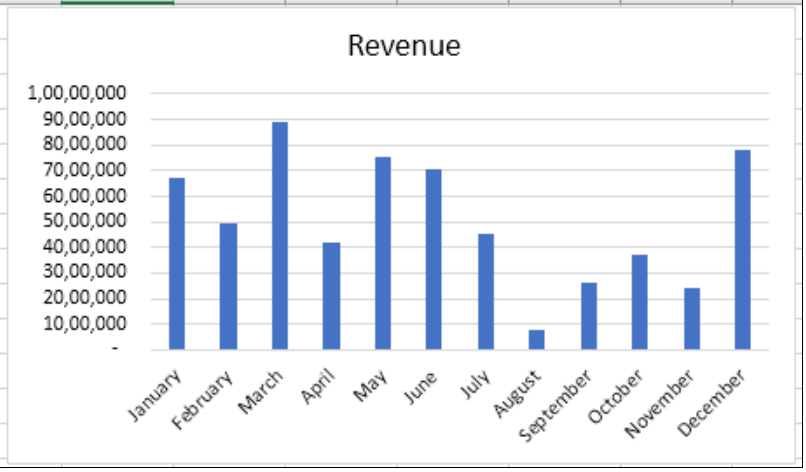


Figure 8.3.5 Update Data and Chart is updated

9. What you find is that the Chart is also updated.
10. Re-open the PowerPoint presentation and the chart is updated in PPT too.

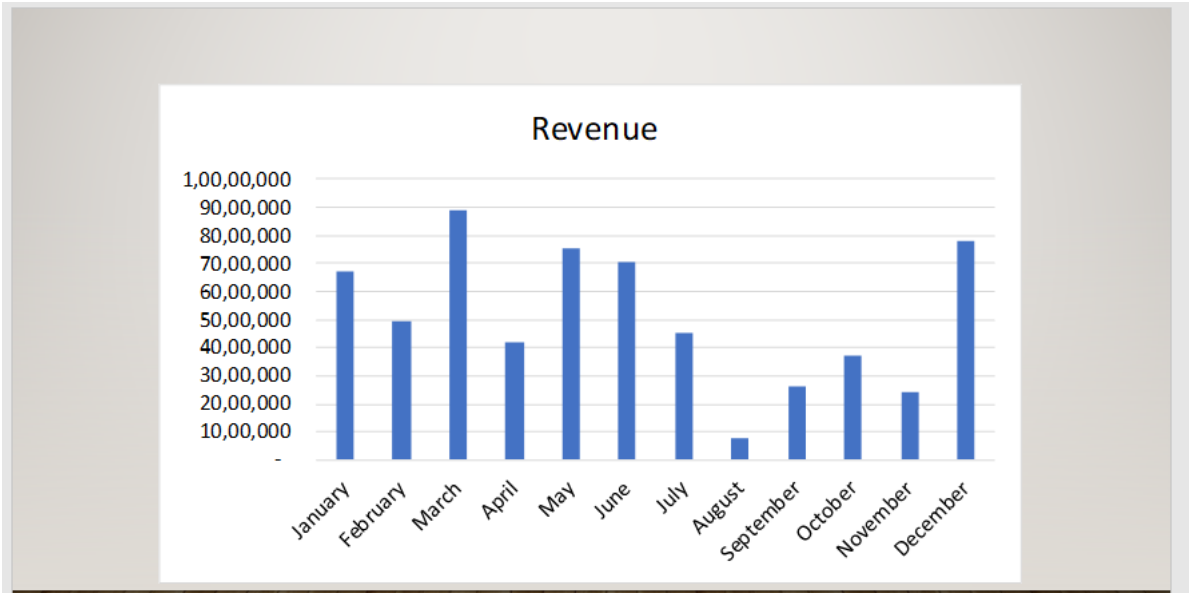


Figure 8.3.6 Updated Chart on PowerPoint

The similar procedure would be used for Word Documents also as detailed below

11. Copy the Chart from Excel.
12. Open a Word Document and Paste Special as a link as shown in Figure

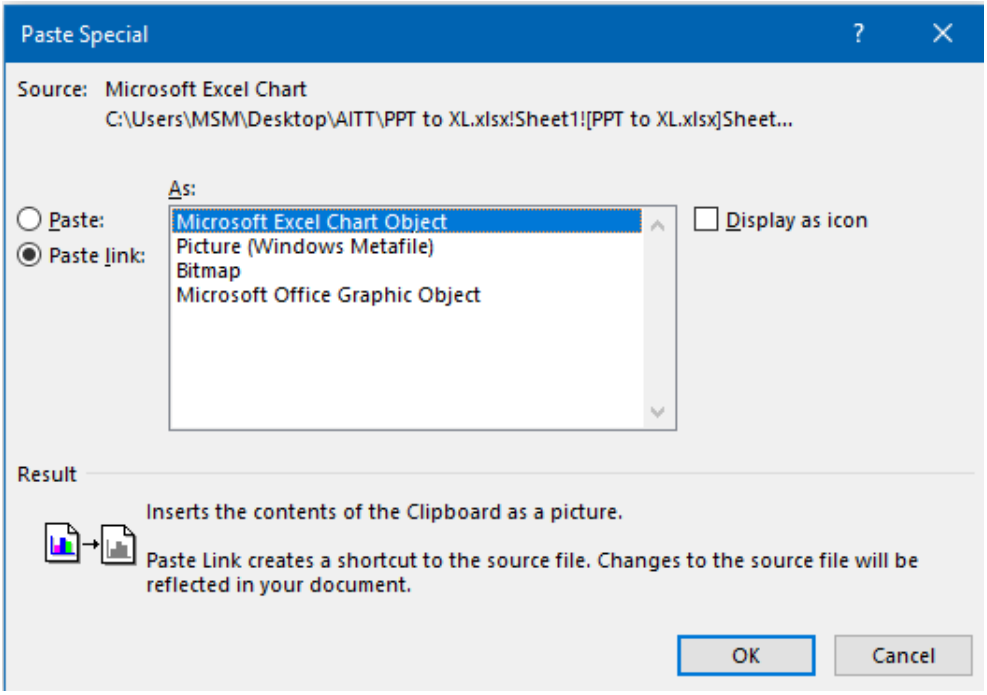


Figure 8.3.7 Paste special as a link in Word Document

13. Chart is pasted
14. Save and close the word document
15. Re-open the Excel file add data relating to November and December and save it.
16. Reopen the Word Document and the chart is updated in the word document also.

Gist: You have dynamically updated a PPT and Word Document from Excel.

Commands learnt: **Paste Special as a link**

8.3.2 MAIL MERGE

Mail Merge From Excel to Word can save you plenty of time when sending a large number of mails. You can swiftly create custom letters, emails, or mailing labels in Word by merging the information from your Excel spreadsheet.

Doing a Mail merge would involve combining data in a word document with data coming from Excel in specific fields. Inputs required are Word mail merge document with placeholders for the personalized recipient fields and Excel mail merge source file with information about the recipients, one row for each recipient, after Mail merge final Word document will be personalized letters, emails, or even printed Labels.

In a mail merge, the Excel file gets connected to the Word mail merge document, and Microsoft Word will pull the recipients' names, addresses, and other details directly from the Excel worksheet.

CASE STUDY*: *Your principal CA Chandiwala wants to send a letter to all the Debtors to confirm the balances, the address and contacts file are in an excel sheet which he wants to be linked to a word document containing the letter to be sent. Can you set up a mail merge system for him?

Strategy:

You can use the mail merge system in Microsoft Word to link to Excel Sheet.

- Set up an excel file with Data as shown in the figure.

	A	B	C	D	E	F	G	H
1	**S.No**	**Title**	**FirstName**	**LastName**	**Company**	**Address**	**City**	**Amount**
2	1	Mr	Vijay	Kumar	Earth Limited	1 Drive way	NEW DELHI	72,769
3	2	Mr	Sumit	Kumar	Galaxy Enterprises	6 Moon drive	MUMBAI	1,45,035
4	3	Mr	Anshuman	Kumar	Mars & Pluto	7 Sun View	CHENNAI	1,18,115
5	4	Mr	Ravi	Kumar	Milky Way	9 Bright Way	NEW DELHI	18,801
6	5	Mr	John	Kumar	Moon Light Enterprises	11 Milk drive	MUMBAI	1,09,736
7	6	Mr	Prabhu	Kumar	Neptune Pvt. Ltd	5 Northcross	CHENNAI	92,977
8	7	Mr	Vinayak	Kumar	Sun Moon & Star	6th Avenue	NEW DELHI	1,88,217
9	8	Mr	Janardan	Kumar	Sun Moon & Stars	5th Cross	MUMBAI	1,83,510
10	9	Ms	Vibha	Kumari	Sun Way	9 Neat way	CHENNAI	1,39,624
11	10	Ms	Chetna	Kumari	Venus Enterprises	5 Tiverton	NEW DELHI	1,48,212

Figure 8.3.8 Excel Data

- Compose the letter in a word document.
- Go to the **Mailings tab** > **Start Mail Merge** group
- Select the mail merge type - **Letters.**

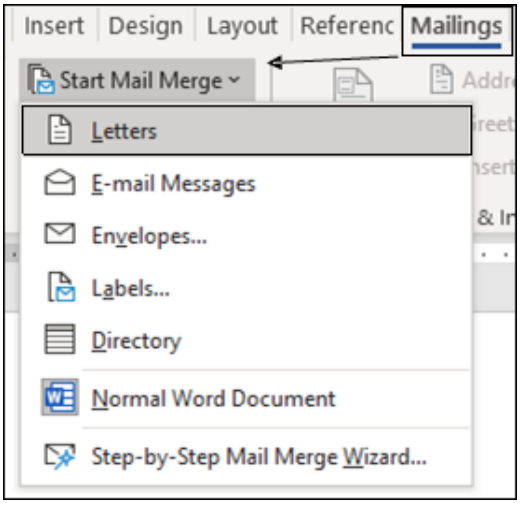


Figure 8.3.9 Start Mail Merge

- Next in the Mailings tab, click **Select Recipients > Use Existing List**

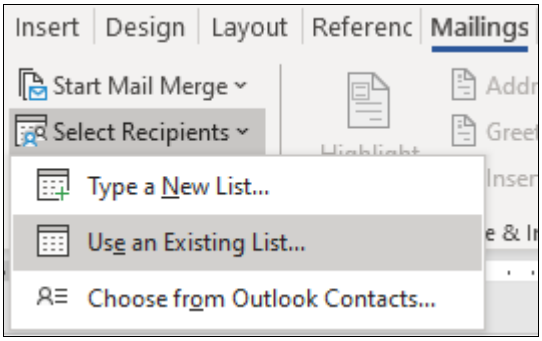


Figure 8.3.10 Select recipients

- Next, you will connect your Excel spreadsheet and Word document. By creating a connection between the Excel sheet and the Word Document you are ensuring your mail merge data source to automatically update each time you make changes to the Excel file that contains the recipients' data.
- Select the **Data Source dialog**, browse to your Excel sheet, and click **Open**.
- Go to **Select table** Dialog Box and select the table.

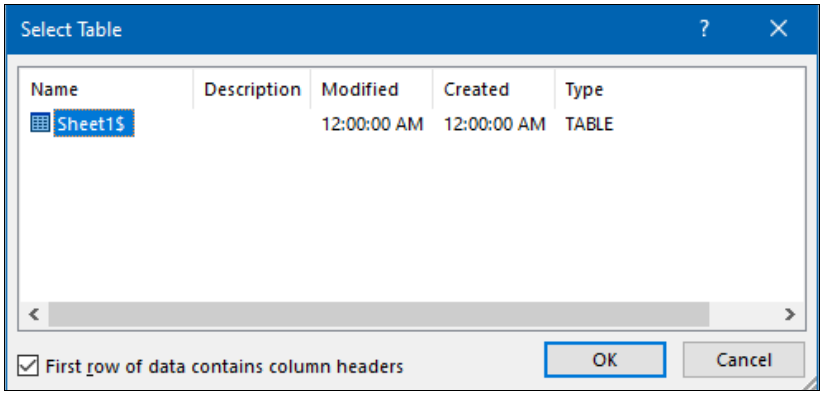


Figure 8.3.11 Select Table

- Click the **Edit Recipient** List button in the Start Mail Merge group, in case you are interested in only some of your Excel entries.

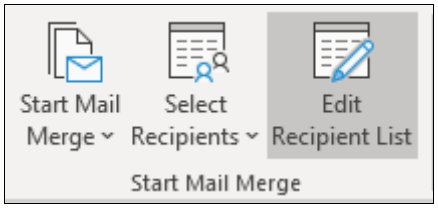


Figure 8.3.12 Edit Recipient List

- Mail Merge Recipients dialog opens, and you can check or uncheck checkboxes

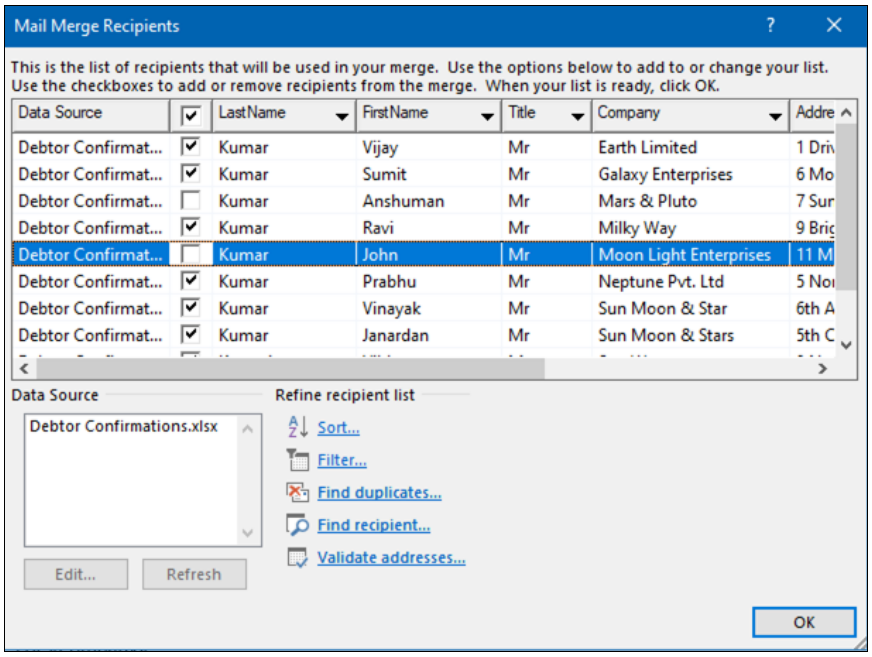


Figure 8.3.13 Mail Merge Recipients

- In the Word Document, add Placeholders.

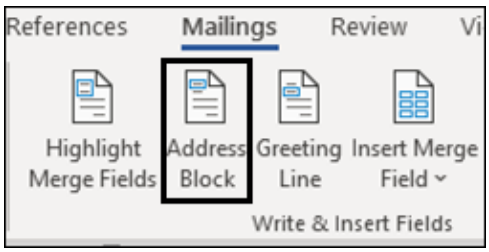


Figure 8.3.14 Add Place Holders

- Insert **Address Block**.

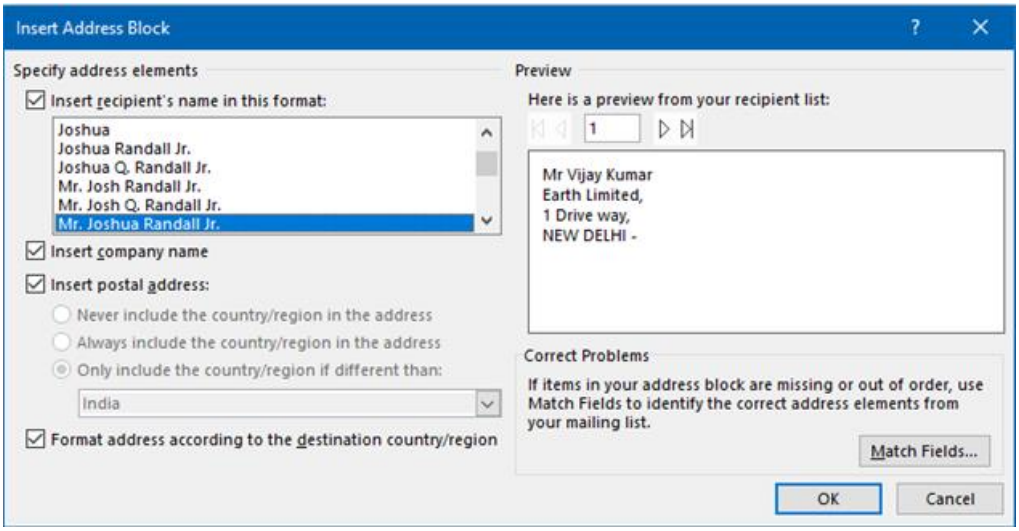


Figure 8.3.14 Insert Address Block

- Similarly, Insert Greeting Line Place Holder

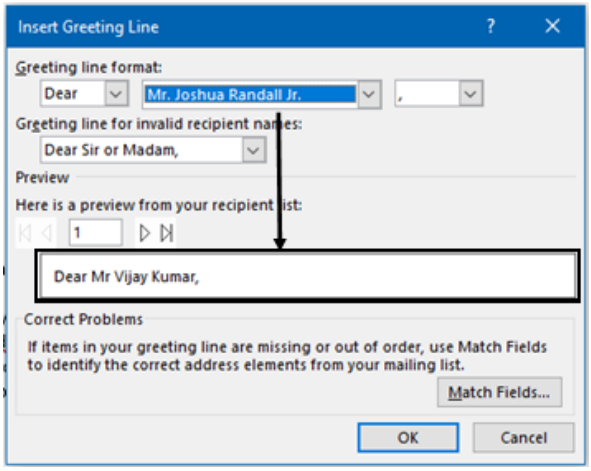


Figure 8.3.15 Insert Greeting Line

- Similarly Add other Place holders.

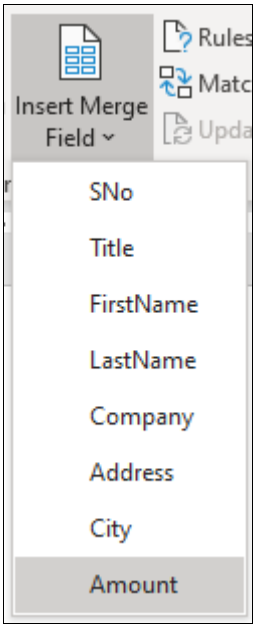


Figure 8.3.16 Add other Placeholders

- Result is as shown in Figure.

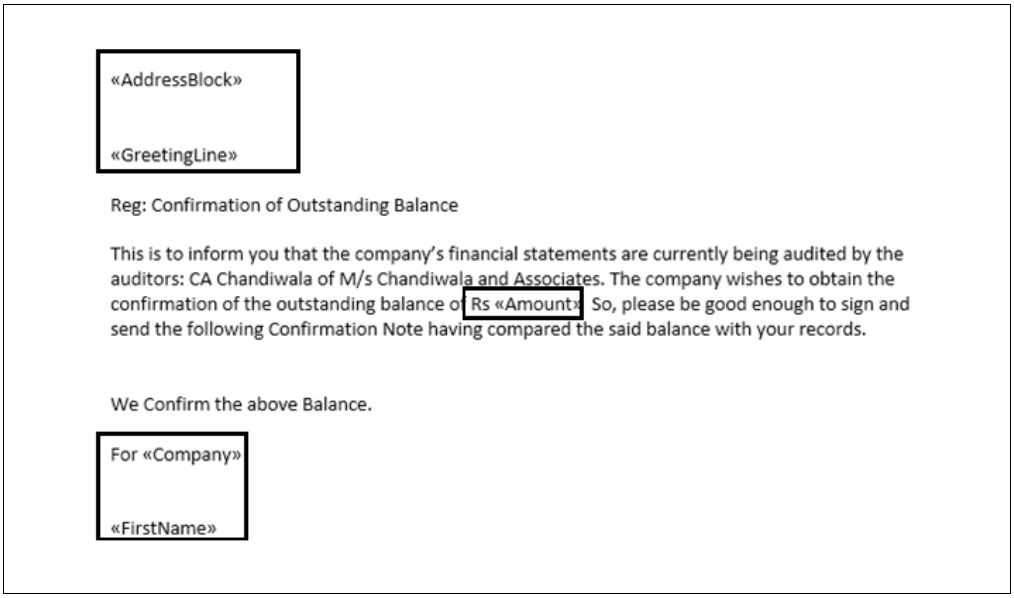


Figure 8.3.17 Document with place holders

- When the letter is printed out, all the copies will be identical except for the data of Placeholders like recipients' names and amounts etc.

- You can then preview the letter.

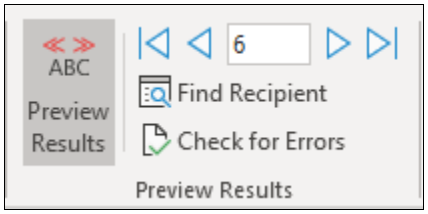


Figure 8.3.18 Preview results

- You will find a preview as shown in Figure

Mr Vijay Kumar
Earth Limited,
1 Drive way,
NEW DELHI -

Dear Mr Vijay Kumar,

Reg: Confirmation of Outstanding Balance

This is to inform you that the company's financial statements are currently being audited by the auditors: CA Chandiwala of M/s Chandiwala and Associates. The company wishes to obtain the confirmation of the outstanding balance of Rs 72769 So, please be good enough to sign and send the following Confirmation Note having compared the said balance with your records.

We Confirm the above Balance.

For Earth Limited

Vijay

Figure 8.3.19 Preview of Letter

- If you are satisfied with the preview, go to the **Finish group,** and click the **Finish & Merge button**.

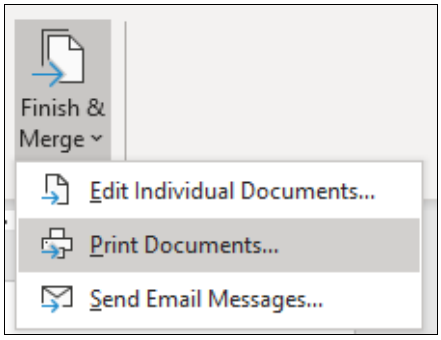


Figure 8.3.20 Print merged Letters

- You can choose to print the letters or send them as email messages.

  Gist: You have set up a mail merge to print letter with Data pulled from Excel in a Word document

  Commands learnt: **Mail merge.**

8.4 MACROS

Macros are code that automates work in a program. Wherever you need to do repetitive steps Excel gives you a facility of Macros which are small programs that automatically perform repetitious steps.

Programming of Macros is done in programming Language VBA (Visual Basic for Applications) but you can use Macros even if you do not know VBA since Excel gives you a wonderful tool in the form of Macro Recorder. A macro records your mouse clicks and keystrokes while you work and play them back later

Macros can be written in two ways

- Writing a Macro using VBA Code
- Recording a macro using Excel Macro recorder

Macros are in the **Developer Tab**, to turn on the Developer tab, execute the following steps.

- Right-click anywhere on the ribbon, and then click **Customize the Ribbon**.
- Alternatively, go to Options > Customize the Ribbon > On the right of the window, a large box lists all the tabs that are currently shown in the ribbon. Near the bottom, you see an unchecked item named Developer
- To show the **Developer tab, check this box**, and then click OK.
- Developer Tab appears on your Ribbon.

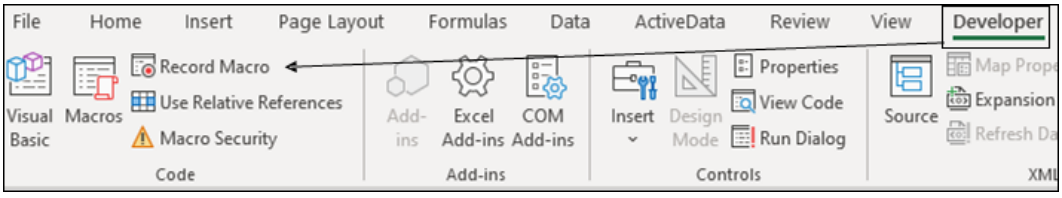


Figure 8.4.1 Developer Tab with Macros

8.4.1 RECORDING A MACRO

It is not possible to store Macros in .xlsx files. Macros can only be stored in files that have a file extension .xlsm which is macro-enabled workbooks.

Points to remember in a macro.

- Excel records every keystroke & every command which you run, so you should not do something that you do not want to record should not be done while recording Macro.
- You do not need to work fast, i.e. Macro just records your actions, so if you are just browsing, that is not recorded it is only specific actions that get recorded.

Three Methods of recording macros

Method 1

Macros are under the Developer tab as shown in Fig.8.4.1

Method 2

Recording a Macros is also available in **View> Macros** as shown in Figure

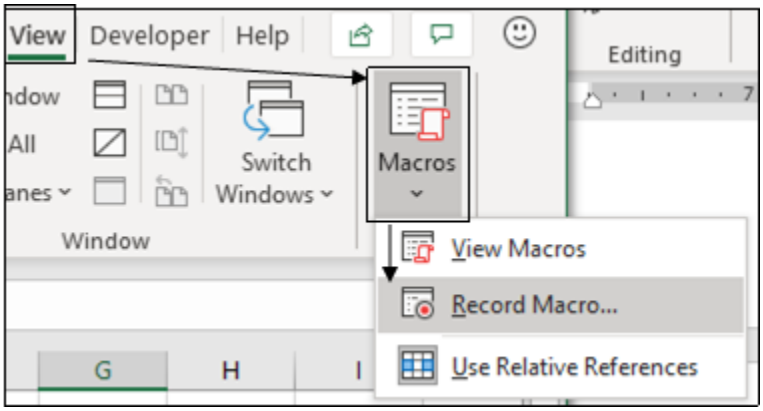


Figure 8.4.2 Macros in the View tab

Method 3

There is **one more option** to record a macro in the **status bar** as shown in Figure

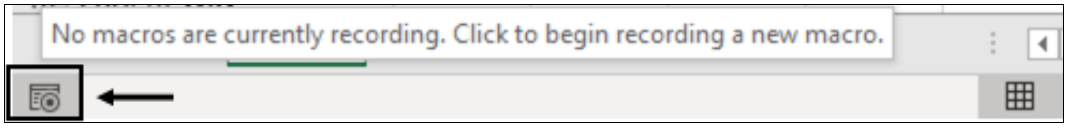


Figure 8.4.3 Macro recording in the status bar

8.4.2 ENABLING MACRO SECURITY

On the Developer tab, click **Macro Security** in the Code group.

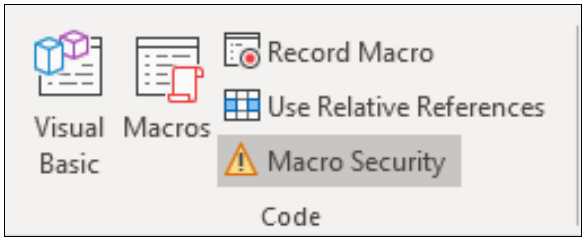


Figure 8.4.4 Macro security in Code group

- The Security dialog appears.

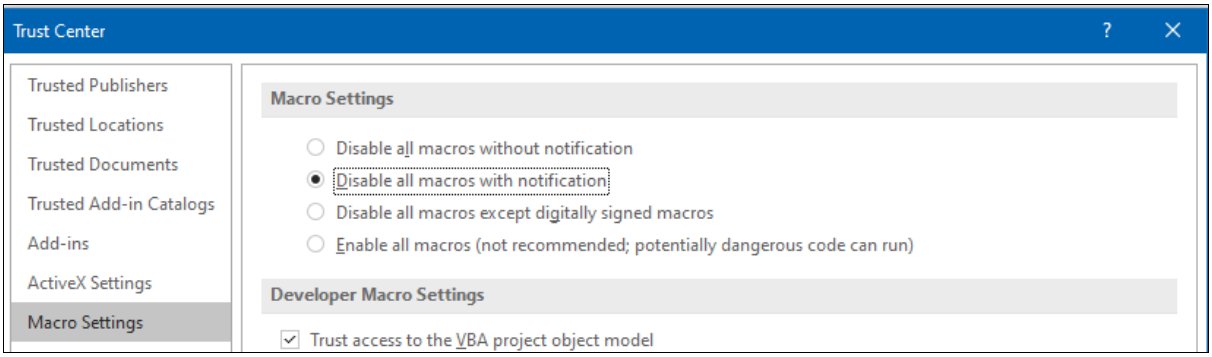


Figure 8.4.5 Enabling Macro Security

In the Security dialog, change the Macro Settings to **Disable All Macros with Notification**.

With this setting, Excel alerts whenever you will open a workbook that has macros attached.

When you open a document and get the warning that the document has macros attached, if this is a document that you wrote and you expect macros to be there, click Enable Content to enable the macros

WHERE MACROS ARE STORED

Macros can be stored in either of two locations, as follows:

- The workbook you are using, or
- Your Personal Macro Workbook (which by default is hidden from view)

If the macro applies to all workbooks, then you store it in the Personal Macro Workbook so it will always be available in all your Excel workbooks; otherwise, you store it in your current workbook.

ABSOLUTE AND RELATIVE REFERENCING

Absolute reference mode. In absolute reference mode, Excel stores the absolute references for the cells that you are modifying.

Relative reference mode. In relative reference mode, Excel tracks how far you have moved from your starting position.

RELATIVE REFERENCING

By default, the Excel Macro Recorder records your absolute steps. As an example If you move from cell A1 to B1 after performing an action, you will press the Right Arrow key, but Excel will not record this key, instead, it will only record the movement to cell B1. Now imagine you have recorded the actions and are to perform it in cell C1, when you run the Macro, a line of Macro will be executed, and then cell B1 will be selected instead of selecting cell D1(which is to the right of C1).

Excel records the movement to cell B1 instead of recording every action(keystroke). If you want Excel to record relatively, so that Excel moves to the right cell instead of selecting cell B1, you will have to enable Relative References. It can be switched on from the Macros menu just below the Start/Stop Recording option.

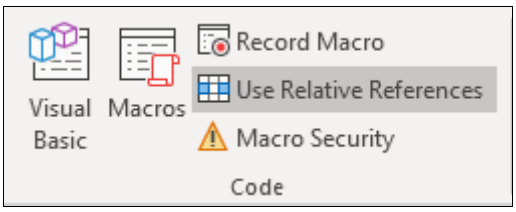


Figure 8.4.6 Use Relative Referencing

CASE STUDY Your principal CA Chandiwala has given you a dat in single column as shown in Figure. He wants you to convert it into a list with, Company, Address, City, and Amount in different columns. To convert it you will have to follow the boring routine of cut, paste & delete row for each company. You are a smart auditor and want to automate this task.

	A
1	**Company**
2	**Address**
3	**City**
4	**Amount**
5	Earth Limited
6	1 Drive way
7	NEW DELHI
8	72,769
9	Galaxy Enterprises
10	6 Moon drive
11	MUMBAI
12	1,45,035
13	Mars & Pluto
14	7 Sun View
15	CHENNAI
16	1,18,115

Figure 8.4.7 Amount List

Strategy:

You can automate this boring task using Macros in Excel

1. You will need a Columnar data with **Company**, **Address, City** and **Amount**.
2. Now before starting the recording of the above macro you must activate **"Use Relative References**" and proceeded to record the Macro.
3. Go to **Developer Tab and** Click and activate **"Use Relative References"** and you observe that the option has changed in Colour**.**

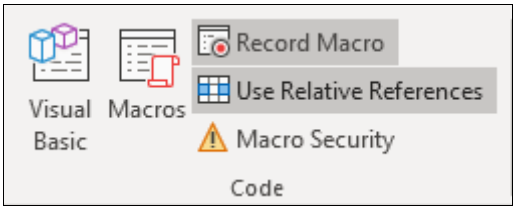


Figure 8.4.8 Record macro using Relative referencing

4. Next click on Record Macro**,** a **Record Macro Dialog Box** appears as shown in Figure.

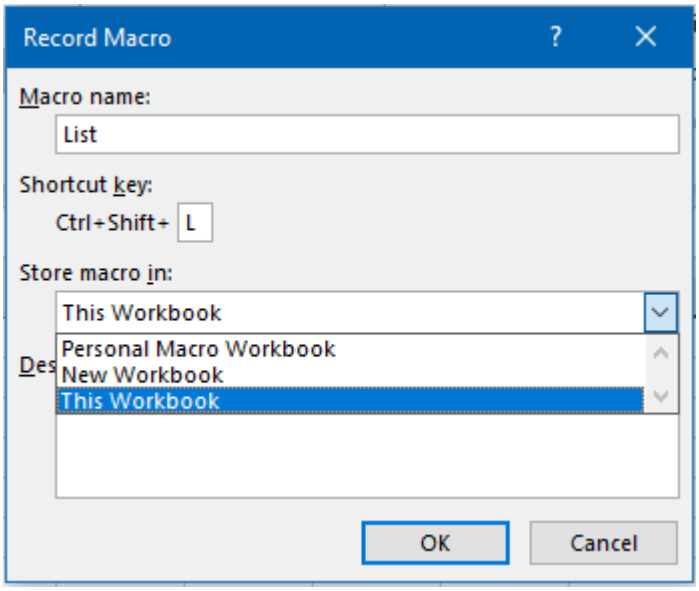


Figure 8.4.9 Record Macro Dialog Box

5. You can give the macro a name let's say *"List"*, you can also attach a shortcut key to it, since most of Ctrl + short cut keys are already reserved it is better to go for **Ctrl+Shift+.** In this case, you select "**Ctrl+Shift+L**" just a mnemonic since you want to create a list. Further Macro can be stored in:

This Workbook.

If you choose this option, Excel stores your macro in the current workbook. Remember, you need to save this workbook as a macro-enabled *.xlsm* file, or you will lose your macros.

New Workbook.

If you choose this option, Excel automatically creates a new workbook (which opens in a separate window) and stores your macro there.

Personal Macro Workbook.

If you choose this option, Excel stores your macro in a special hidden workbook named Personal.xlsb. The Personal.xlsb workbook opens automatically whenever you start Excel (although it remains hidden), so macros in this workbook are always available no matter what workbook you are using.

6. Store the macro in" this workbook "since this macro would be used by you only once in this specific workbook.
7. Click Ok
8. Now the recording starts, and all your steps would be recorded in the form of Macro.
9. As you begin recording, you see that the record macro button has changed to "**stop recording"** in both the header & status bar as shown in figure

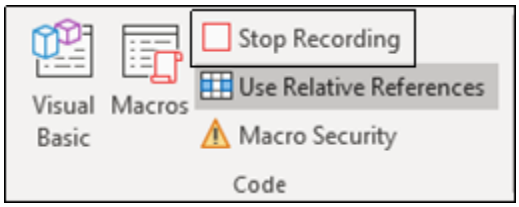


Figure 8.4.10 Stop Recording

10. You perform the steps as shown in Figure

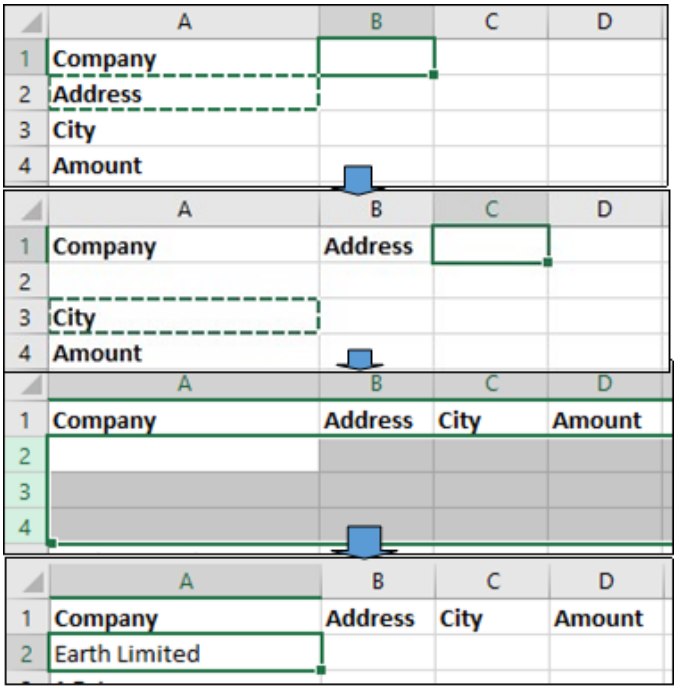


Figure 8.4.11 Steps to record

11. You start
    a. *at Cell A2*
    b. *Cut Cell A2 and paste to Cell B1.*
    c. *Cut Cell A3 and paste to Cell C1.*

d. Cut Cell A4 and paste to Cell D1.

e. Delete Rows 2 through 4.

f. Go to Cell A2

12. By **using "*Use Relative References*" you are telling Excel**

a. Move down one cell,

b. Cut the Value

c. Move one cell up, one to the right, and paste the value.

d. Move two cells down.

e. Cut the Value

f. Move two cells up, two to the right, and paste the value.

g. Move three cells down.

h. Cut the Value

i. Move three cells up, three to the right, and paste the value.

j. Delete the rows

k. Move one cell below

13. Stop Recording

14. Your Macro is now recorded.

15. You try executing it & press **Ctrl+Shift+L.**

16. The result is perfect something you desired.

17. You execute the macro several times and y**our sheet gets into the desired shape i**n a single Row as shown in Figure

	A	B	C	D
1	**Company**	**Address**	**City**	**Amount**
2	Earth Limited	1 Drive way	NEW DELHI	72,769
3	Galaxy Enterprises	6 Moon drive	MUMBAI	1,45,035
4	Mars & Pluto	7 Sun View	CHENNAI	1,18,115
5	Milky Way	9 Bright Way	NEW DELHI	18,801
6	Moon Light Enterprises	11 Milk drive	MUMBAI	1,09,736
7	Neptune Pvt. Ltd	5 Northcross	CHENNAI	92,977
8	Sun Moon & Star	6th Avenue	NEW DELHI	1,88,217
9	Sun Moon & Stars	5th Cross	MUMBAI	1,83,510
10	Sun Way	9 Neat way	CHENNAI	1,39,624
11	Venus Enterprises	5 Tiverton	NEW DELHI	1,48,212

Figure 8.4.12 Final List

Gist: You have arranged data in a single row using a Macro

Commands Learnt: Use Relative References in Macros

8.5 FORM CONTROL

In Excel Worksheet you can have Form Controls that are objects which give you the functionality to interact with your model's data.

You can implement these controls on worksheets to help select data. These could be drop-down boxes, list boxes, spinners, and scroll bars that are useful for selecting items from a list. You can also have Option Buttons and Check Boxes to allow the selection of various options. These Buttons allow execution of VBA code.

By adding a control to a worksheet and linking it to a cell, you can return a numeric value for the current position of the control.

Form Controls are located on the Developer Tab under Insert Form Control.

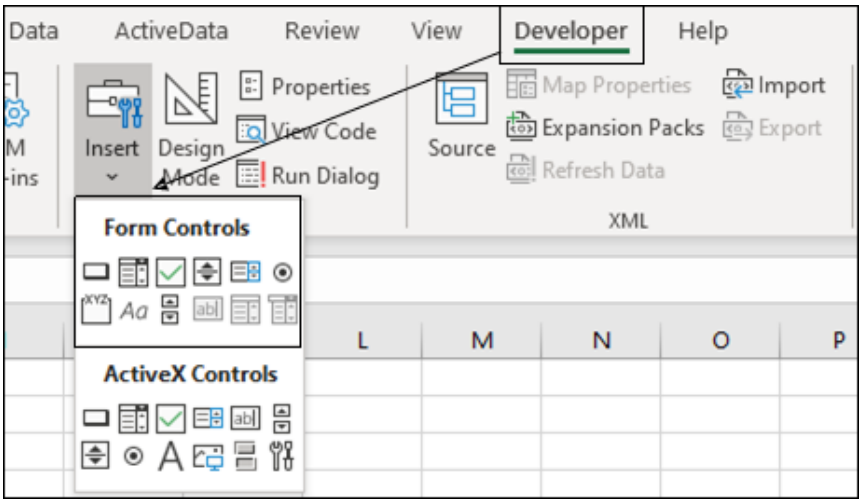


Figure 8.5.1 Form Control Under Developer Tab

There are various types of Form Controls which give interactivity ranging from a simple display to allowing multiple selections or even selection of values.

Control Name	Description	Function
Button	Push Button	Executes a macro
CheckBox	Allow selection of non-exclusive options	Multiple On/Off options
Combo Box	Drop Down selection Box	Select items from a Drop-down list
Group Box	Layout element which groups common elements	Nil
Label	A Text label	Can be static or linked to a cell
List Box	Fixed selection box	Select items from a list
Option Button	Allow selection of exclusive options	Exclusive Single On/Off option

Scroll bar	Allow Horizontal or Vertical scrolling	Increases or decreases the value of a cell by a fixed amount
Spin Button	Increment/decrement a value by a fixed amount	Increases or decreases a cell in steps by a fixed amount

BUTTON(FORM CONTROL)

The Button Form Control is a Button, Pressing the Button allows execution of a macro.

You can link Button to a macro as shown in Figure

- Right-click on the **Button** and select **Assign Macro**
- The **Assign Macro** dialog will pop up.
- **Select the macro** you want to assign to the button.

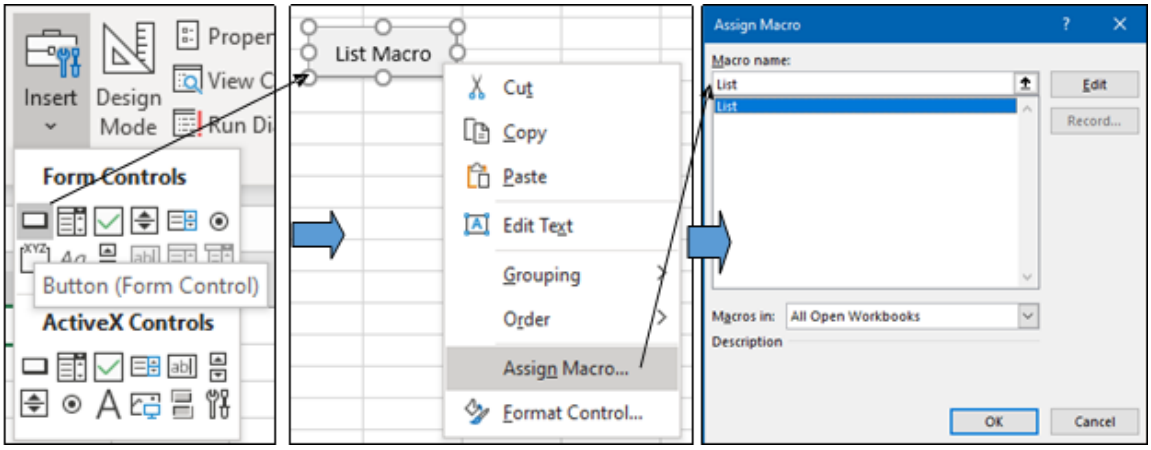


Figure 8.5.2 Button Form Control

The Button has no other controls.

COMBO BOX (FORM CONTROL)

Combo Box is like your drop-down list and it works the same as the drop-down list but the combo box excel is considered as an object.

- Select the “Combo Box” and draw anywhere on the worksheet area as shown in Figure
- Next to insert values create a list.
- Now select the “Combo Box” right click and choose “Format Control”
- Now in the “Format Control” window choose “Control”. In the “Input Range” choose the range of cells.
- Click Ok.

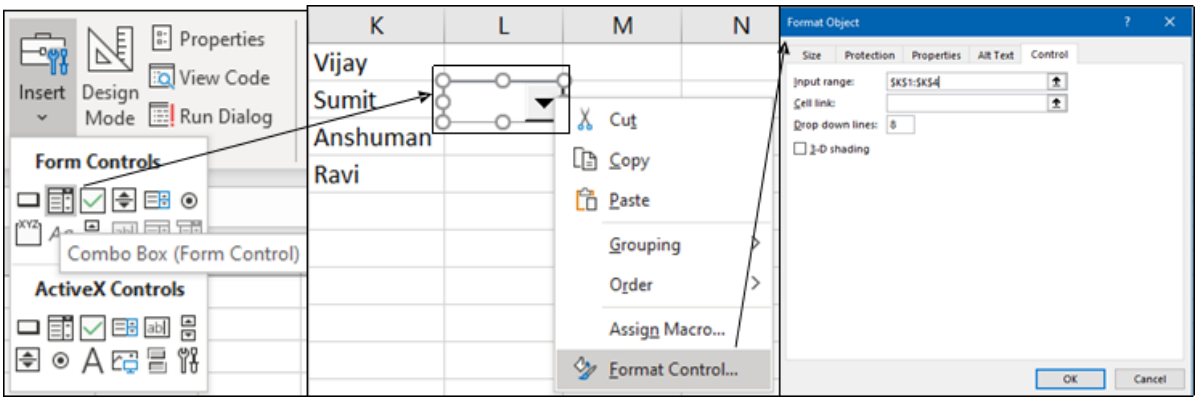


Figure 8.5.3 Combo Box

Now you can see the selected list in the combo box.

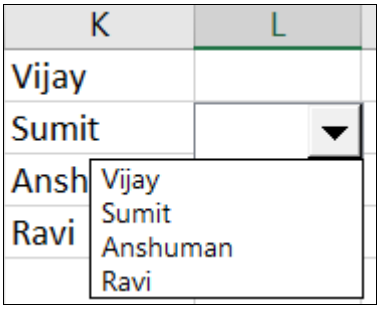


Figure 8.5.4 Combo Box Drop Down

SPIN BUTTON

Using the "Spin Button" you can increase and decrease the value in the linked cell.

- Draw the spin button to see options.
- Right-click on the button and choose "Format Control".
- Under the "Control" tab you can make all the settings.
- You can set a minimum value, a maximum value, and a current value.
- You can also configure what should be the incremental and decremented value when the spin button is clicked.
- Click ok.

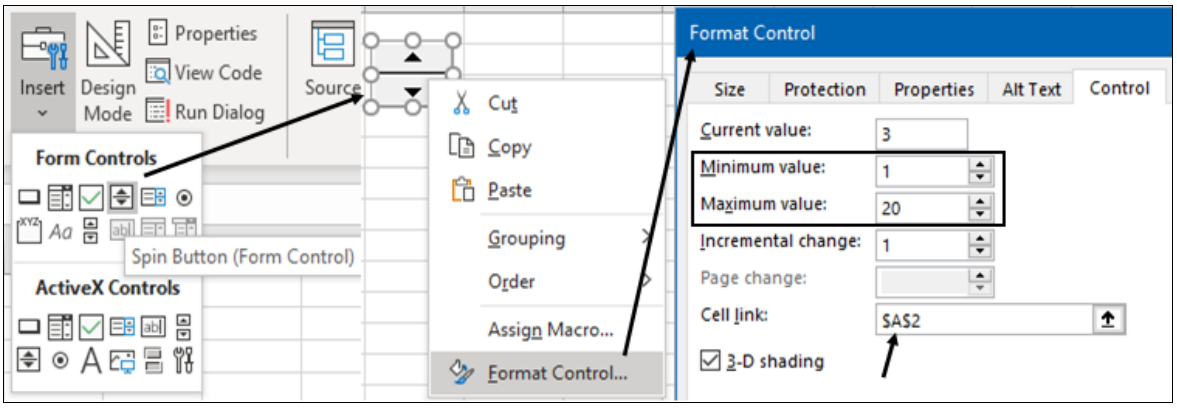


Figure 8.5.5 Spin Button

Now you click up the arrow of the spin button in cell A2 one you can see the incremental value.

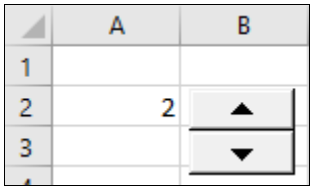


Figure 8.5.6 Number spinning

There is a maximum value of 20, so when you press up arrow it will increment by 1 until it reaches 20 after that it will not increase.

CHECK BOX (FORM CONTROL)

The CheckBox (Form Control) allows the selection of several non-exclusive options.

- There may be any number of CheckBox controls that may be implemented and they independently are on or off and have no relationship to each other.
- The CheckBox Form Control returns the value indicating its status, either **True (selected)** or **False (not selected**), to a linked cell.
- Draw the checkbox on the worksheet.
- Right-click and choose "**Format Control**".
- Under the "Control" tab choose "Unchecked" and give cell link to **C3** cell.
- Click Ok

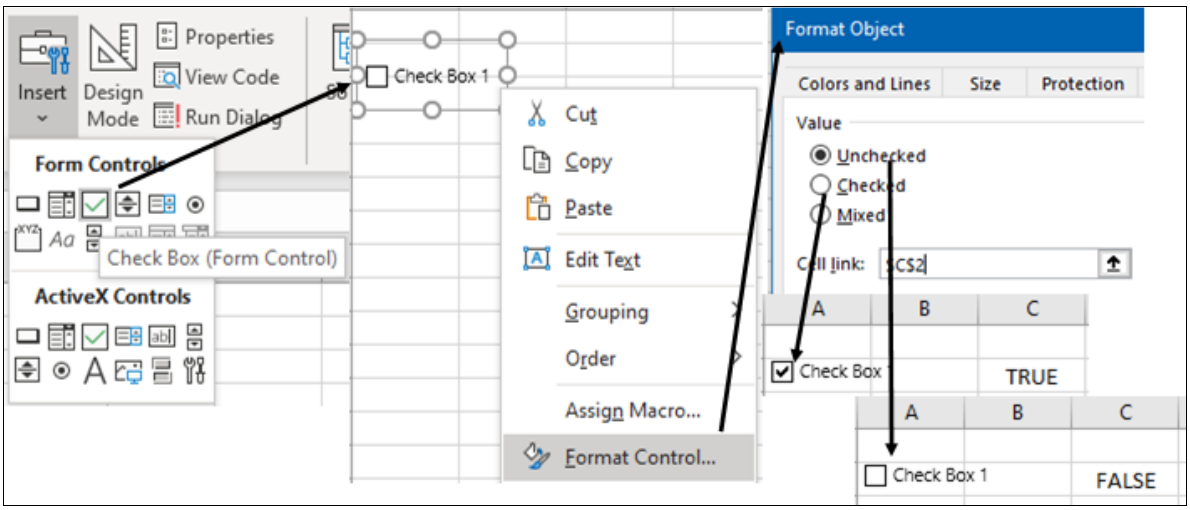


Figure 8.5.7 CheckBox

- **Check the box** to see the **TRUE** value in cell C2.
- Now uncheck the box to see the FALSE value.

The use of the information from a form control depends on your visualisation

Some uses are

- Selecting items for a chart
- Selecting data sets
- Moving data sets
- Adjusting values in a model

Form Controls do not return a value directly from a list, they all return either a number or an index number relative to the position of the item in a list.

8.6 AUTOMATION AND LINKING FILES

Excel provides many facilities for Automation and it is important so that repetitive tasks and refreshing of data should be facilitated. Excel gives a wonderful facility of linking with external files. It also provides a facility to link files with many kinds of files even text files.

You can link text files to excel so that whenever any change or expansion occurs in those data can be refreshed in excel. And results change automatically.

CASE STUDY: Your principal CA Chandiwala wants to run a regular analysis on some sales data which is in a text file. He wants you to link to an excel file so that he can refresh the same easily and the KPIs he has set on Sales data are updated automatically. You want to help him set up an excel file that can update from the text file.

Sales.txt - Notepad

File Edit Format View Help

```
DATE      INV_NO  CUSTOMER          AMOUNT
03-04-2020      1       Milky Way       " 28,667.00 "
05-04-2020      2       Sun Moon & Star " 23,589.00 "
14-04-2020      3       Galaxy Enterprises      " 8,954.00 "
16-04-2020      4       Venus Enterprises       " 17,350.00 "
19-04-2020      5       Galaxy Enterprises      " 11,416.00 "
24-04-2020      6       Sun Way " 45,460.00 "
29-04-2020      7       Galaxy Enterprises      " 11,621.00 "
04-05-2020      8       Moon Light Enterprises  " 9,864.00 "
09-05-2020      9       Galaxy Enterprises      " 13,082.00 "
14-05-2020      10      Moon Light Enterprises  " 10,618.00 "
19-05-2020      11      Earth Limited   " 59,705.00 "
20-05-2020      12      Milky Way       " 21,472.00 "
20-05-2020      13      Galaxy Enterprises      " 28,826.00 "
20-05-2020      14      Sun Way " 7,705.00 "
20-05-2020      15      Moon Light Enterprises  " 8,192.00 "
20-05-2020      16      Earth Limited   " 77,882.00 "
20-05-2020      17      Sun Way " 27,329.00 "
```

Figure 8.6.1 Sales Text File

Strategy

You can create a link to Excel with text data.

1. On the **Data** tab, go to **Get & Transform Data**, and then click **From Text**.

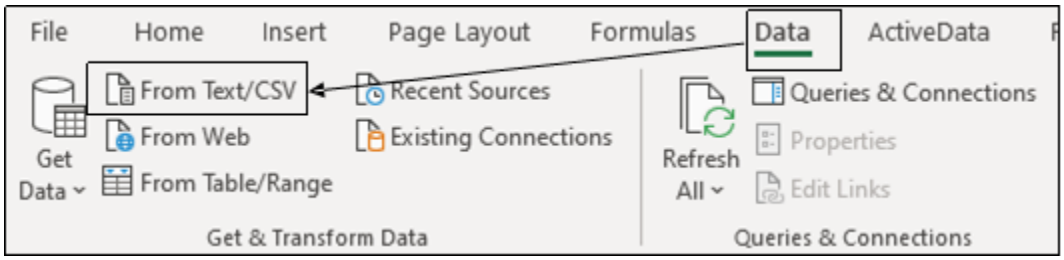


Figure 8.6.2 Get External Data from Text

2. In the **File Name** list at the bottom of the dialog box, browse to the location where you have the **Sales Link.txt** file

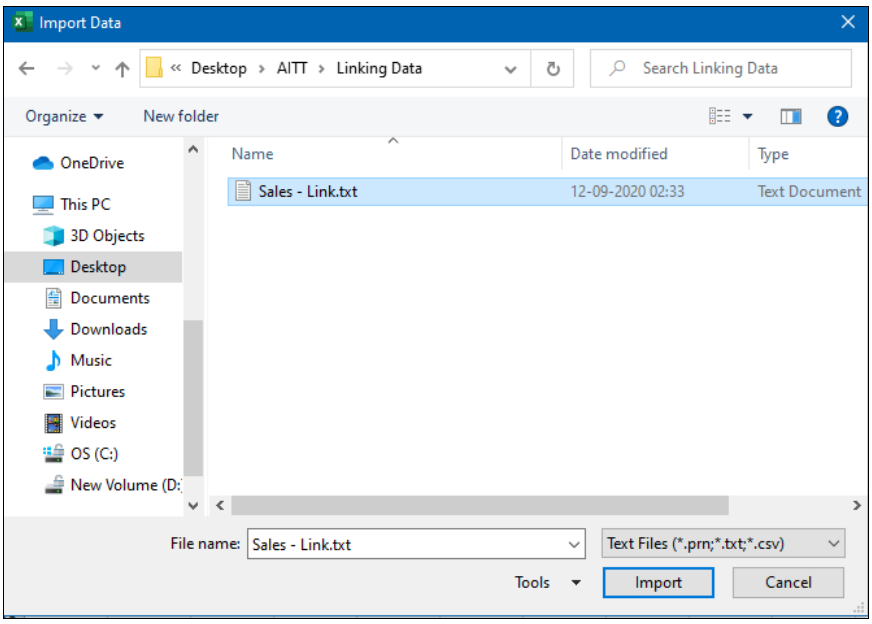


Figure 8.6.3 Import Data

3. Select the file, and then click **Import.**
4. The **Text Import Wizard** opens.
5. In step 1 of 3, you select the **Delimited** option to split the text into the proper columns.

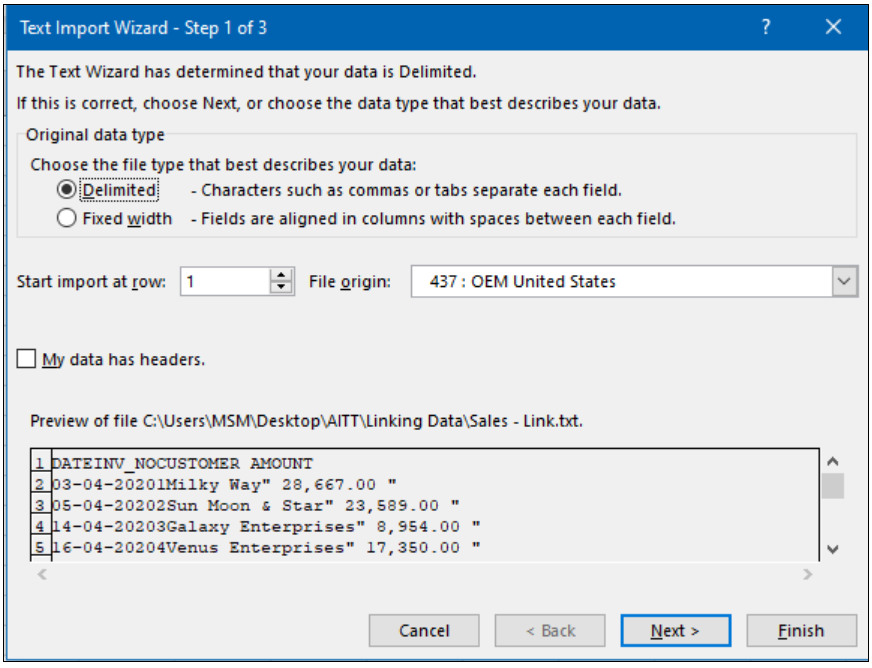


Figure 8.6.4 Step 1 of Text Import Wizard

6. Click Next.

7. In step 2 of 3 of the Text Import Wizard, you will select the delimiters contained in the text file. Since the file is tab-delimited file Tab is already selected.

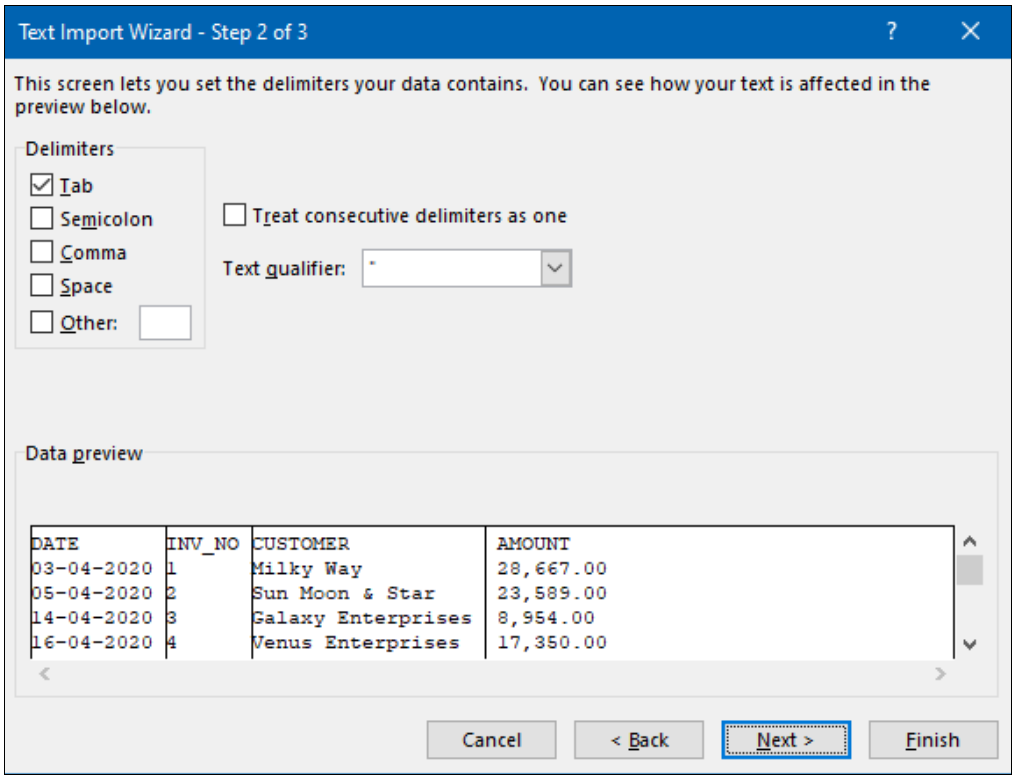


Figure 8.6.5 Step 2 of Text Import Wizard

- "**Text qualifier"-** you can select double quotation marks ("), single quotation mark ('), or none, or you can leave this option alone.
- Text qualifiers (" or ') are sometimes used in text files to indicate the beginning and end of text strings.
- You can see in the preview if qualifiers are used. For example, double quotation marks around the Sanction limit on your case this tells excel that these are text strings.
- If the file contained more than one type of delimiter (commas or semicolons, for example), you would select them all.
- Further, you could have more than one delimiter between two data, you need to select "Treat consecutive delimiters as one" checkbox.

8. You see neatly aligned Columns. Click **Next**.

9. In step 3 of 3 of the wizard, allows you to set the columns as certain data types or not import at all as shown in Figure

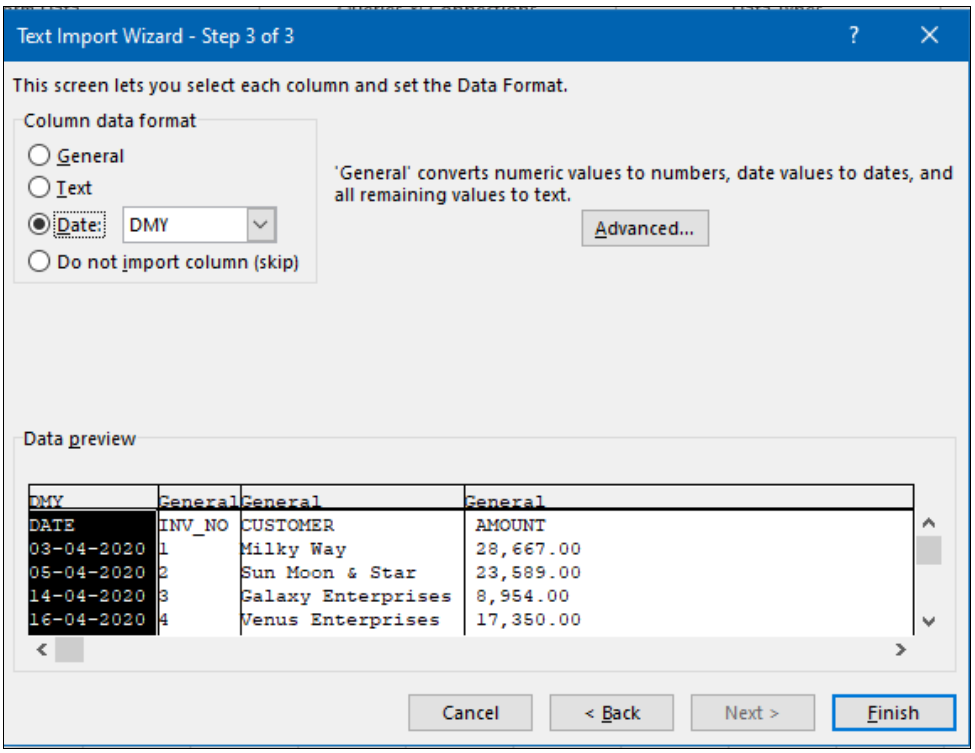


Figure 8.6.6 Step 3 of Text Import Wizard

10. You can make formatting changes if necessary. Excel automatically formats each column as **General**. Numeric values are converted to numbers, date values are converted to dates, and all remaining values are converted to text.

    Rather than accepting the default of General, you can choose from three other options:

    - **Text:** This option instructs Excel to treat a column as text, meaning leading zeros will be preserved in columns that contain numeric values, such as Account Numbers.
    - **Date**: This option instructs Excel to convert a column to date. When you choose this option, you must then specify the date format used within your text file, such as DMY. Excel will convert these dates to the standard date format for Excel spreadsheets.
    - **Do Not Import (Skip**): This option instructs Excel to in effect throw away unneeded columns within the text file. The original text file will remain intact, but any columns that you mark as Do Not Import will not appear in your spreadsheet.

11. For example, In the above case, you could select a date for column **"Date**" and "expiry date" and select an appropriate Date mask on your case DMY.

12. Click Finish.

13. After which the **Import Data** dialog box appears where you have the option of having the data in an existing worksheet or new worksheet as shown in Figure

14. Select A1

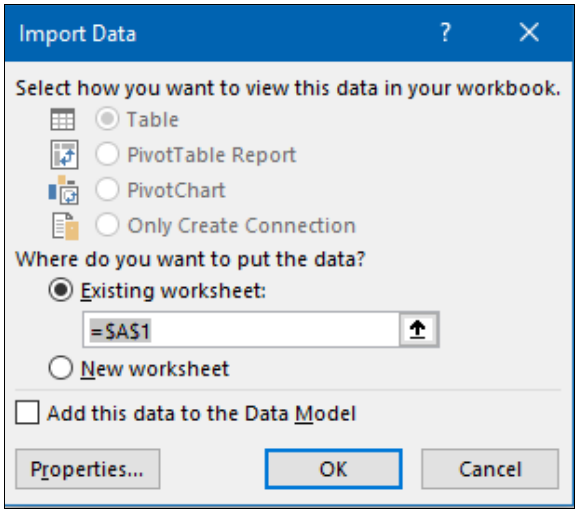


Figure 8.6.7 Import data Dialog Box

15. Under **properties** on the Import dialog box, you have various options as shown in Figure

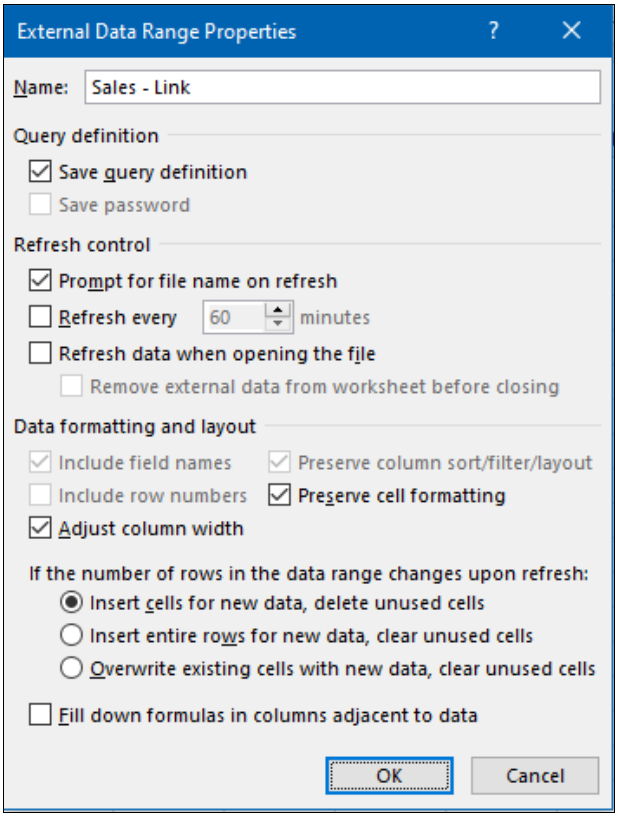


Figure 8.6.8 External data Dialog Box

16. Click OK and the **file is imported as an Excel sheet**.

	A	B	C	D
1	DATE	INV_NO	CUSTOMER	AMOUNT
2	03-04-2020	1	Milky Way	28,667.00
3	05-04-2020	2	Sun Moon & Star	23,589.00
4	14-04-2020	3	Galaxy Enterprises	8,954.00
5	16-04-2020	4	Venus Enterprises	17,350.00
6	19-04-2020	5	Galaxy Enterprises	11,416.00
7	24-04-2020	6	Sun Way	45,460.00
8	29-04-2020	7	Galaxy Enterprises	11,621.00
9	04-05-2020	8	Moon Light Enterprises	9,864.00
10	09-05-2020	9	Galaxy Enterprises	13,082.00
11	14-05-2020	10	Moon Light Enterprises	10,618.00
12	19-05-2020	11	Earth Limited	59,705.00
13	20-05-2020	12	Milky Way	21,472.00
14	20-05-2020	13	Galaxy Enterprises	28,826.00
15	20-05-2020	14	Sun Way	7,705.00
16	20-05-2020	15	Moon Light Enterprises	8,192.00
17	20-05-2020	16	Earth Limited	77,882.00
18	20-05-2020	17	Sun Way	27,329.00

Figure 8.6.9 Linked Excel sheet

17. In the **Connections** group, you can use **Refresh All** to update the data in Excel if the external data is updated.

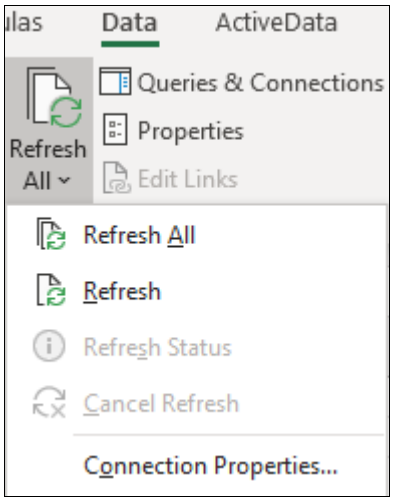


Figure 8.6.10 Connection- Refresh All

18. In cell D2, delete the figure 28,667

19. Now the external data and the data in Excel are different.
20. On the **Data tab > Connections**, click **Refresh all**.
21. It asks you to select the source, once you select "Sales-Link.txt"
22. You see that data has been refreshed from the txt file and figure 28,667 is in the cell again.

Gist: You have linked the Text file to Excel

Commands Learnt: Link connection with text data

Similarly, Excel can be linked to various other data sources for automation.

8.7 EXCEL VERSIONS

There has been a multitude of different versions of Excel, right from stretching from 1985 to the present day. These days most of you will have Excel 2016, 2019, or 365.

MICROSOFT 365 (Excel)

Excel in Microsoft 365 is slated to have an infinite life. Earlier Excel had a 3-year update cycle, 365 is continually evolving with the latest features being added as each new version is released. That is it is always latest.

Many Business users are not interested in constant change and more interested in stable releases, Microsoft introduced the concept of "update channels" for Microsoft 365. Therefore Business users could opt-in for a six-monthly update channel. This helps bring a tested stable version in January and July each year.

In Excel 365 a huge number of older Excel functions now work in an entirely new way and many old favourites have been suplemented by many modern functions, which do away with shortcomings in earlier functions. As an example VLOOKUP has been supplemented by XLOOKUP, which is much more versatile and improved the lookup functionality. and CONCATENATE has been supplemented with CONCAT which can handle ranges of cells too.

Excel 365 is always the latest and powerful Excel version you can use and it is available on a monthly subscription model rather than perpetual versions.

EXCEL 2019 AND PERPETUAL VERSIONS

Microsoft has now embraced the SaaS (software as a service) model where software is rented and not sold. At the same time, Microsoft also acknowledges that some purchasers still prefer the "buy once use forever" model. For these users, Microsoft has the Excel 2019 version which has many features but the pace at which new features are coming, Excel 2019 will always be behind Microsoft 365.

8.8 ARTIFICIAL INTELLIGENCE IN EXCEL

Excel is bringing many new features that are derived from artificial Intelligence. Excel is three things: cells of data laid out in rows and columns, a powerful calculation engine, and a set of tools for working with the data. Many new AI(Artificial Intelligence) & ML(Machine Learning) basedTools are being added to excel, one example was **Flash Fill** which has been discussed in Chapter 5, another example is **New Data Types** Like **Stock and**

Geography which have been discussed in Chapter 2.

An Excel Dashboard can be a wonderful tool when it comes to tracking KPIs, comparing data points, and getting data-inspired views and insights that can help management make effective decisions. A Dashboard is normally a one-page report which shows through Graphs, Charts, Tables, and Visualisations various metrics, which can help management get information at a glance and acquire a high-level view of the business.

KPIs (Key performance indicators) refer to a set of quantifiable measurements used to gauge an organisations' overall long-term performance. KPIs specifically help determine an organisations' strategic, financial, and operational achievements.

Dashboards are an amazing tool to monitor these KPIs immediately.

Excel can help you create powerful dashboards that can provide analysis, insight, and forewarning to management promptly.

Powered with artificial intelligence (AI) one feature that has made Excel even more powerful is **Ideas**, which helps you visualise data instantly on the fly.

Ideas is an AI-powered insights service that helps people take advantage of the full power of Office. Proactively surfacing suggestions that are tailored to the task at hand, Ideas helps users create and analyse spreadsheets in less time. In Excel, for instance, Ideas helps identify trends, patterns, and outliers in a data set—helping customers analyse and understand their data in seconds.

For the feature to work properly, your data needs to be tabular data with no blank rows or columns. Every column must have a one-row heading, and no two headings should be the same. The data should also include a date field.

Ideas is in **Home Tab**

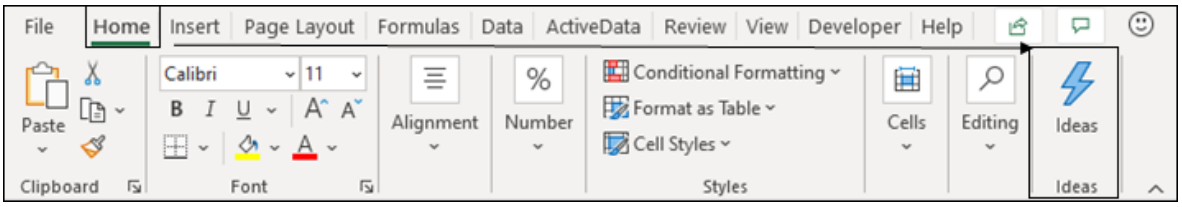


Figure 8.8.1 Ideas

It can give some phenomenal insight as shown in Figure

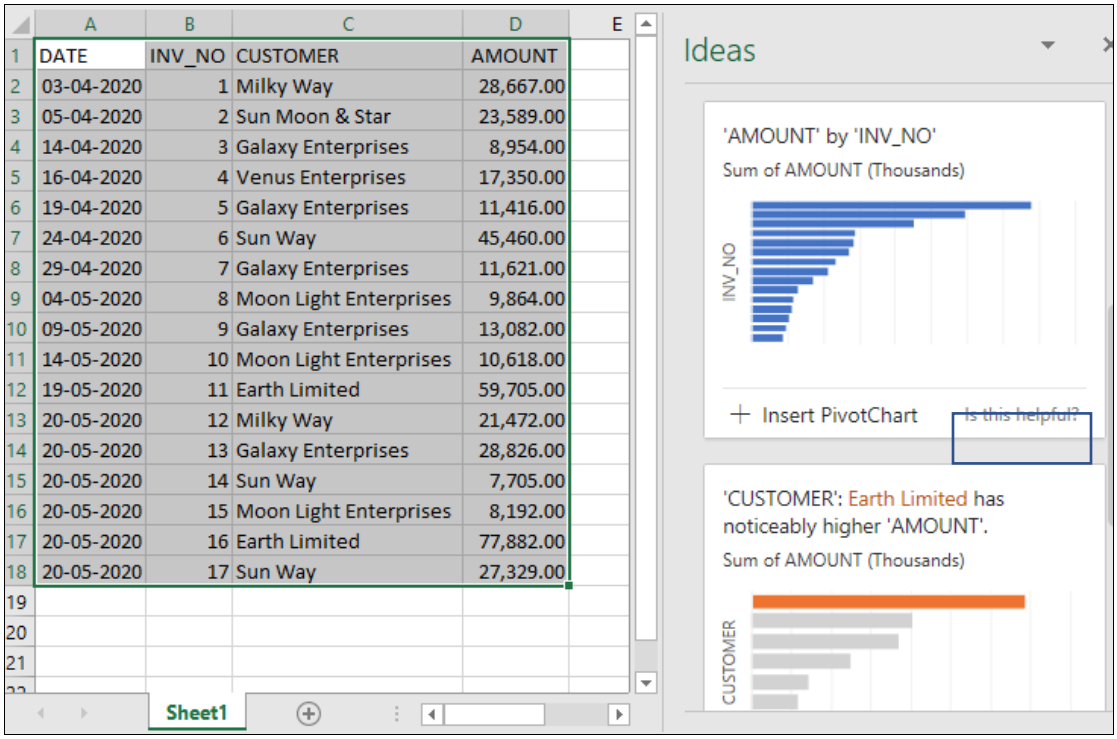


Figure 8.8.2 From List data it has generated Ideas

Ideas will return a small number of results. Each result contains a headline and a thumbnail of a chart. It can generate many charts and Pivots on its own powered by artificial intelligence.

Is This Helpful? link allows you to vote for which charts are useful. Excel will learn your preferences and return more of those types of results in the future. If you find a chart that you like, click the Insert Pivot Chart link and Excel inserts a new worksheet with a pivot table and a full-sized version of the chart.

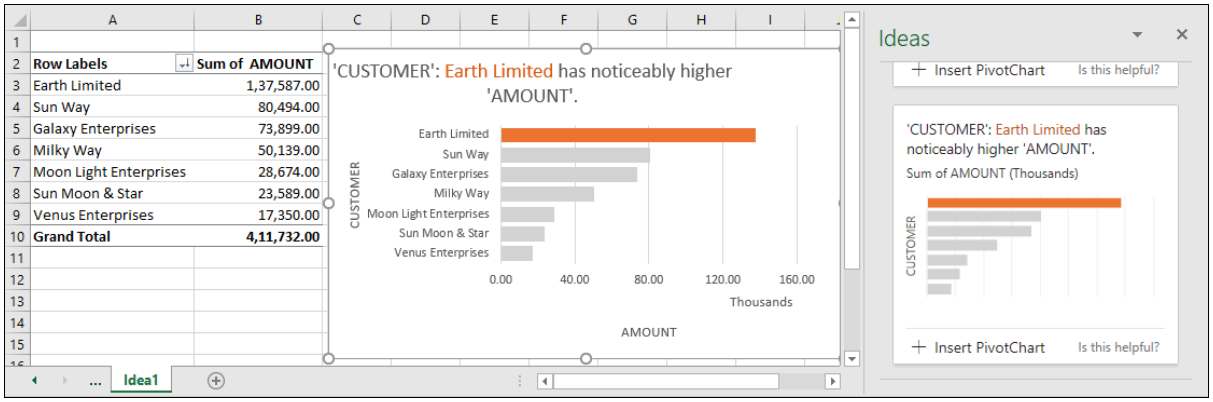


Figure 8.8.3 Generates Charts and Pivots automatically

8.9 GOOGLE SHEETS VS EXCEL

Google Sheets and **Microsoft Excel** are the two best-known spreadsheet applications available today.

Google Sheet is a spreadsheet program brought by Google, it continues to grow into a possible alternative to Excel for most spreadsheet users. One major benefit of Google Sheets is, it is free, Sheets is also better for collaboration, as the program was developed for ease of use and online sharing.

If you use spreadsheets for serious data analysis or visualization, Excel remains the better product. Excel has more built-in formulas and functions. Simple tasks like sorting and filtering are easier in Excel. It also has more charting options. Though both have their pros and cons, the gap between both Excel and Google Sheets is growing smaller day by day.

Large data sets

Excel is ideal for large data sets and such data sets can be easily manipulated using excel.

As data size increases, Google Sheets starts slowing down. It is more suitable for small data sets.

Chatting facility

The chatting facility is not available in Excel.

In Google Sheets, the Chatting facility is available in the sidebar.

Track changes

Excel has a feature of track changes. It has to be enabled to start tracking changes in data.

In Google Sheets, track changes are enabled by default. And the track is available on a real-time basis.

Online offline usability

Excel can be used in offline mode only. In online it has very limited functionality

Google sheets can be used both online as well as offline.

Cell Quantum

Excel, has around 17 billion cells.

Google Sheets has a maximum of 5 million cells.

Number of functions

Excel supports numerous functions in a varied category.

Google sheets have comparatively limited functionality, but it is catching up.

Usage in computers at a given point of time

Excel files can be used in a single computer at any point in time, and after usage needs to be shared with other users for further data editing from their computer.

Google sheets can be used in multiple computers through the internet on a real-time basis.

Economical/Expensive

Excel is more expensive as the whole MS Office pack is needed to be subscribed.

Google sheet is free, or the cost is negligible.

8.10 SUMMARY

Large Excel sheets bring in inefficiencies in work, you learned why Excel workbooks become bloated and how to keep the size in check.

In this chapter, you learned to link Excel data and charts dynamically with PowerPoint and word documents. You also understood how you can use Macros to automate repetitive tasks by recording the steps using a Macro recorder and these macros can also be activated through Form controls

You explored how linking is possible with external data for automation.

Lastly, you learned the merits and demerits of Google sheets with Excel.

REFERENCES

[1] Michael Alexander, Richard Kusleika, John Walkenbach, 'Excel 2019 Bible', Wiley,2018

[2] Greg Harvey, 'Excel 2019 All-in-One for Dummies, John Wiley & Sons,2018

[3] www.chandoo.org

[4] https://theexcelclub.com/introduction-to-excel-form-control/

CHAPTER

9 COMPUTER ASSISTED AUDIT TECHNIQUES (CAATS)

LEARNING OBJECTIVES

- Understanding need of CAAT
- Understanding structure of IDEA
- Understanding Projects in IDEA
- Understanding Import and Export of files
- Understanding formatting in IDEA

9.1 INTRODUCTION FOR CAAT

Auditors deal with information in myriad ways encompassing the areas of accounting, assurance, consulting and compliance and most of this information is now available in electronic form. This is true not only in case of large and medium enterprises but even in small enterprises. In case there are enterprises who have still not adapted the digital way, then it is an opportunity for Auditors to help such enterprises to ride the digital wave. Hence, it has become critical for Auditors to understand and use information technology as relevant for the services we provide. It is rightly said: "one cannot audit data which is flying in bits and bytes by using the ancient method of riding on a horse back". We are living in a knowledge era where the skill sets are keys to harnessing the power of technology to be effective as knowledge workers.

Computer Assisted Audit Techniques (CAATs) refers to using technology for increasing the effectiveness and efficiency of auditing. CAATs enable auditors to do more with less and add value through the assurance process which is more robust and comprehensive. This chapter provides an overview of the process, approach and techniques which could be used across various technology platforms and in diverse enterprises.

CAATs are tools for drawing inferences and gathering relevant and reliable evidence as per requirements of the assignment. CAATs provide direct access to electronic information and empower auditors not only to perform their existing audits more efficiently and effectively but also facilitate them in knowing how to create and execute new type of IT related audit assignments. CAATs provide a mechanism to gain access and to analyze data as per audit objective and report the audit findings with greater emphasis on the reliability of electronic information maintained in the computer system. There is higher reliability on the audit process as the source of the information used provides and greater assurance on audit findings and opinion.

CAATs are available in specific general audit software designed for this purpose but the techniques of CAATs can be applied even by using commonly used software such as MS Excel and by using query/reporting features of commonly used application software. CAATs can be used to perform routine functions or activities which can

be done using computers, allowing the auditors to spend more time on analysis and reporting. A good understanding of CAATs and know where and when to apply them is the key to success.

In a diverse digital world of clients' enterprises, the greatest challenges for an Auditor is to use technology to access, analyze and audit this maze of electronic data. CAATs enable auditors to move from the era of ticks of using pencil or pen to the era of clicks by using a mouse. CAATs will help auditors to change focus from time consuming manual audit procedures to intelligent analysis of data so as to provide better assurance to clients and also mange audit risks. Some of the key reasons for using CAATs are:

i) Absence of input documents or lack of a visible paper trail may require the use of CAATs in the application of compliance and substantive procedures.

ii) Need for obtaining sufficient, relevant and useful evidence from the IT applications or database as per audit objectives.

iii) Ensuring audit findings and conclusions are supported by appropriate analysis and interpretation of the evidence

iv) Need to access information from systems having different hardware and software environments, different data structure, record formats, processing functions in a commonly usable format.

v) Need to increased audit quality and comply with auditing standards.

vi) Need to identify materiality, risk and significance in an IT environment.

vii) Improving the efficiency and effectiveness of the audit process.

viii) Ensuring better audit planning and management of audit resources.

CAATs refer to using computer for auditing data as per audit objectives. This requires understanding of the IT environment and most critically the core applications and the relevant database and database structure. CAATs could be used by using the relevant functionalities available in general audit software, spreadsheet software or the business application software. However, broadly the key capabilities of CAATs could be categorized as follows:

i) **File access:** This refers to the capability of reading of different record formats and file structures. These include common formats of data such as database, text formats, excel files. This is generally done using the import/ODBC function.

ii) **File reorganization:** This refers to the features of indexing, sorting, merging, linking with other identified files. These functions provide auditor with an instant view of the data from different perspectives.

iii) **Data selection:** This involves using of global filter conditions to select required data based on specified criteria.

iv) **Statistical functions:** This refers to the features of sampling, stratification and frequency analysis. These functions enable intelligent analysis of data.

v) **Arithmetical functions:** This refers to the functions involving use of arithmetic operators. These functions enable performing re-computations and re-performance of results.

Precautions in using CAATs

CAATs have distinct advantages for Auditors and enable them to perform various types of tests. However, it is important to ensure that adequate precautions are in taken in using them. Some of the important precautions to be taken by Auditors are:

i) Correctly identify data to be audited

ii) Collect the relevant and correct data files

iii) Identify all the important fields that need to be accessed from the system

iv) State in advance the format the data can be downloaded and define the fields correctly

v) Ensure the data represent the audit universe correctly and completely.

vi) Ensure the data analysis is relevant and complete.

vii) Perform substantive testing as required.

viii) Information provided by CAATs could be only indicators of problems as relevant and perform detailed testing as required.

Need for using CAATs

As entities increase the use of information systems to record, transact and process data, the need for the auditors to utilize tools to adequately assess risk becomes an integral part of audit coverage. The use of computer-assisted audit techniques (CAATs) serves as an important tool for the auditor to evaluate the control environment in an efficient and effective manner. The use of CAATs can lead to increased audit coverage, more thorough and consistent analysis of data, and reduction in risk.

CAATs include many types of tools and techniques, such as generalized audit software, customized queries or scripts, utility software, software tracing and mapping, and audit expert systems.

CAATs may be used in performing various audit procedures including:

- Tests of details of transactions and balances
- Analytical review procedures
- Compliance tests of general controls
- Compliance tests of application controls

Traditionally, audits were performed manually, but with advancements in IT and information systems, particularly for white-collar financial crimes, audit function is currently using standardized audit software to detect fraud. This means that transparent financial statement reviews and business operations are the most critical and important steps in building confidence among the stakeholders and decreasing instances of fraud.

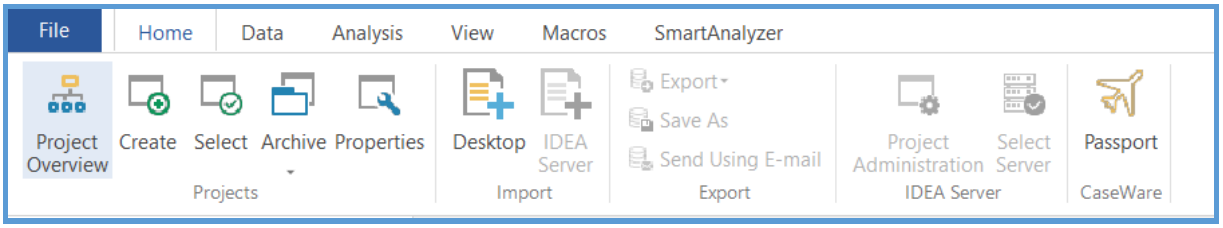
CAATs allow auditors and fraud investigators to access and analyze a large volume of data to uncover fraudulent activities while also obtaining a quick overview of business operations. When CAATs are in place, auditor can better improve the effectiveness and efficiency of the audit process.

A critical factor in the adoption of CAATs is to obtain the right data and understand what that data means. Companies have hundreds of thousands of transactions; CAATs enable auditors to quickly analyze those transactions individually and look for the anomalies. Therefore, CAATs such as IDEA help auditors to focus and put their efforts in high-risk areas that have a higher probability of misrepresentation, rather than focusing on low-risk transactions.

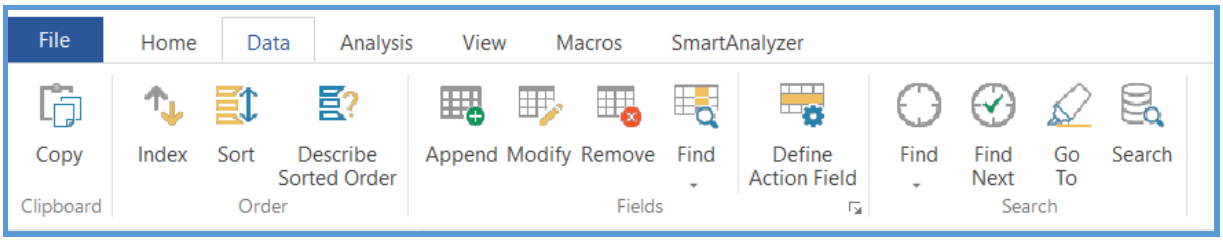
CAATs may produce a large proportion of the audit evidence developed on audits and, as a result, the auditor should carefully plan for and exhibit due professional care in the use of CAATs. The effective implementation and use of CAATT can help auditors in exercising not only a greater level of due diligence and due professional care, but more importantly, greater fraud detection during various audit activities.

The screen view is as follows:

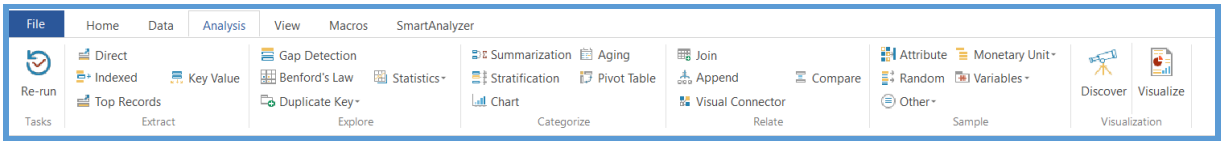
- **Home Tab**



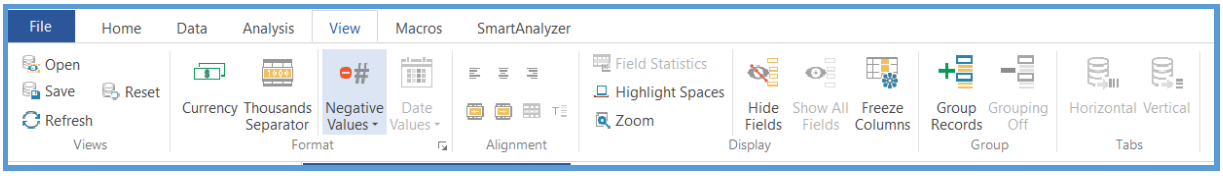
- **Data Tab**



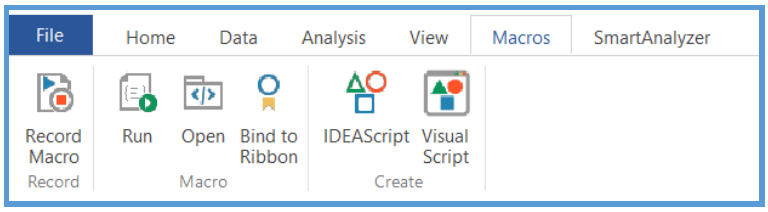
- **Analysis Tab**



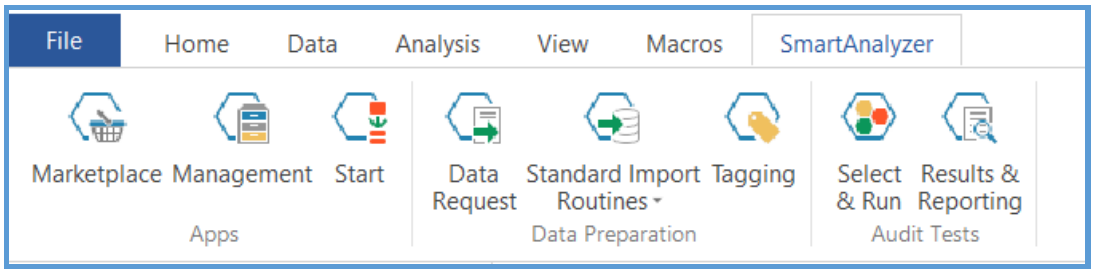
- **View Tab**



- **Macros Tab**



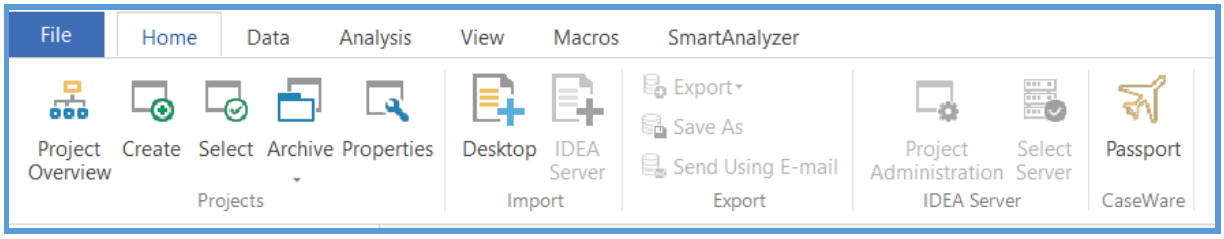
- **Smart Analyzer Tab**



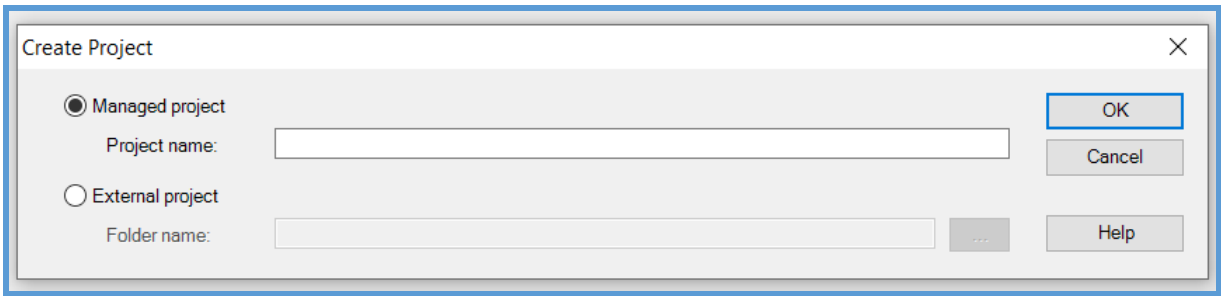
Each Tab and its features are detailed in the Chapter.

9.2 PROJECTS IN IDEA

- **Creation of New Project:** In IDEA software, we create Projects. This option is available in **Home Tab > Projects > Create**

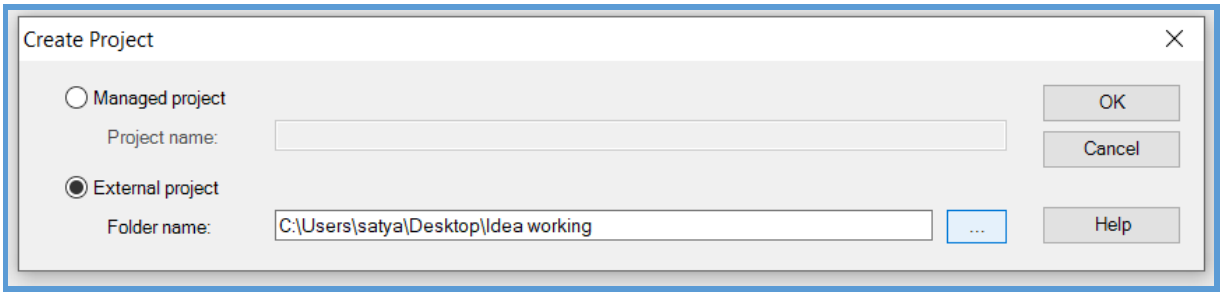


Once we click on Create option, it opens a new menu option. It contains 2 choice i.e. managed Projects and External Projects as shown below.

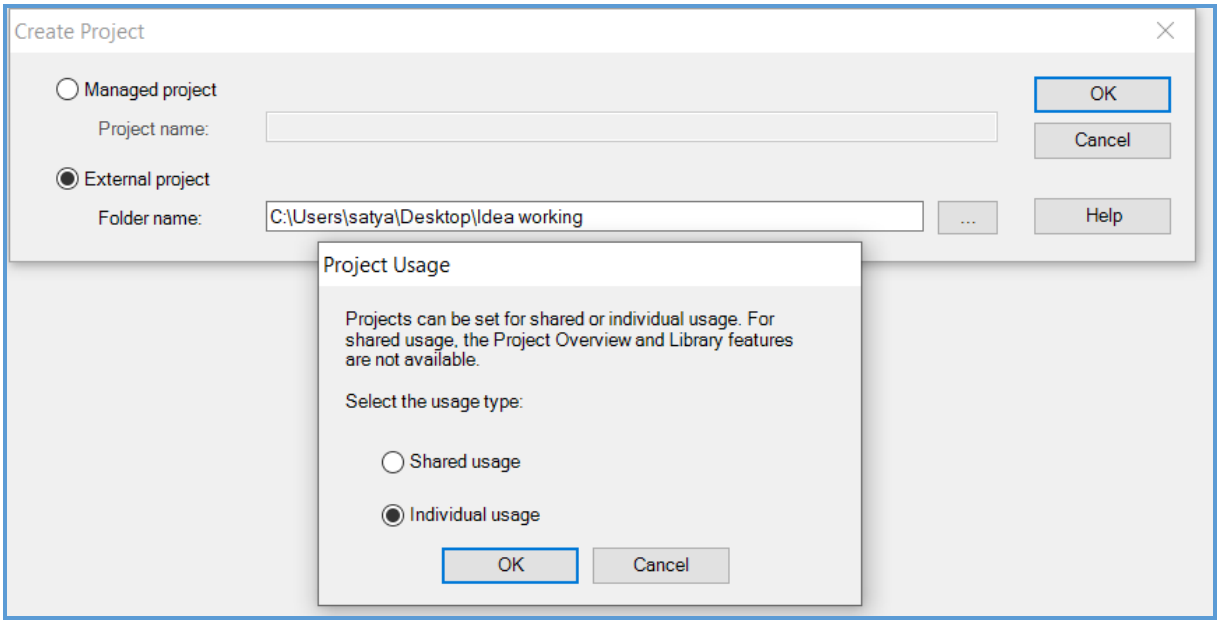


We need to select appropriate option. As we are creating a new project, we need to select the option "External Project". As shown in below image, we have selected a folder "Idea working" as out new project. Then click on button OK.

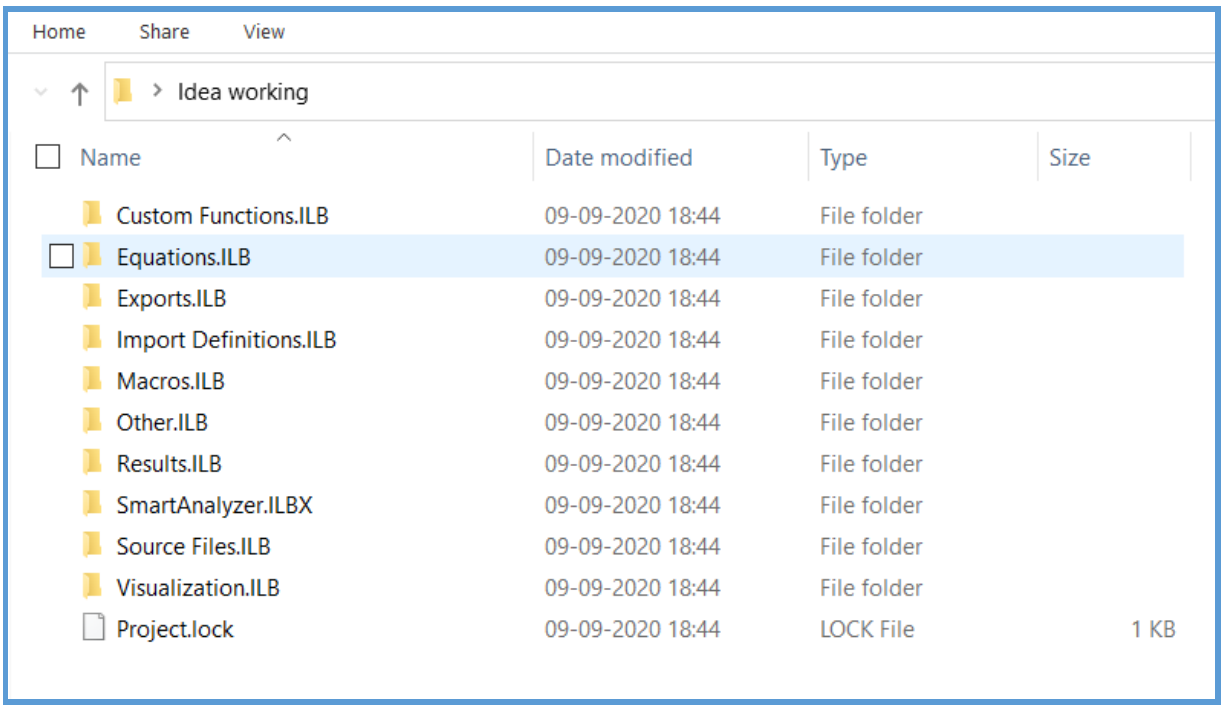
Note: Before selecting the folder, we need to create a folder "Idea working" on Dektop / any other location, where we want to store our project.



Once we click, it will ask, whether we want Shared usage or Individual usage. Depending on requirement, we need to select the option. In this scenario, we are selecting "Individual usage" option.

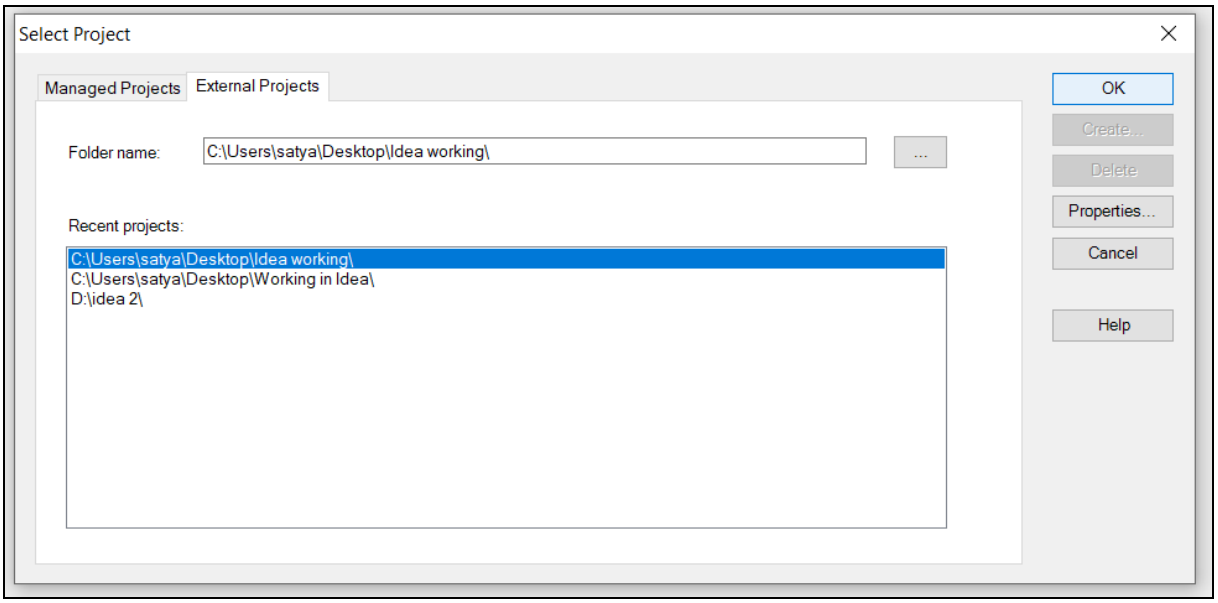


Then a new folder will be automatically created within the "Idea working" folder as shown below. These folders will be used at various stages of Idea working.

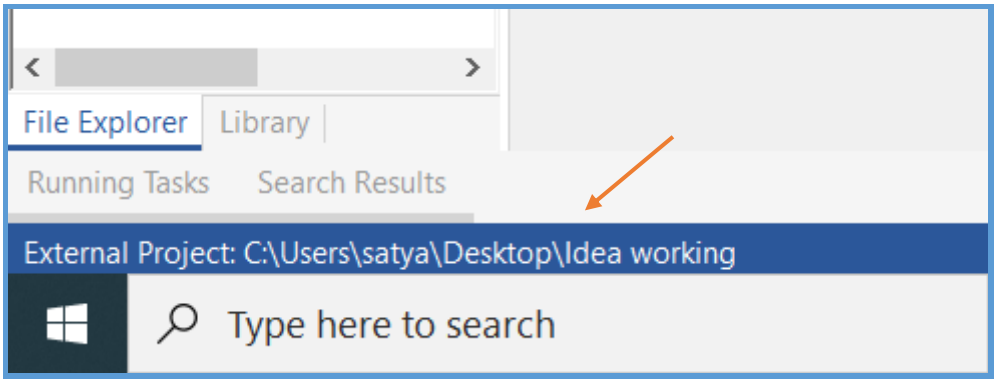


By all these steps, we have created a project i.e. “Idea working”. Using above steps, we can create more projects.

- **Select a Project:** In case, we have multiple projects and we want to select any specific project for working, then we need to go to **Home Tab > Projects >Select.** In the below image, we can see that there are 3 projects and we want to select “Idea working” project. After selecting that, we click on OK button. And that projects become active for use.



At the left side bottom of the screen, we can see the name of active project. In this case, it is Idea working.



- **Archive:** We have all come across a situation where the number of project folders in IDEA can seem daunting. The sheer volume and size of projects taking up vital hard disk space on our PCs are difficult to manage. However, there is a convenient way to tackle this hurdle: the Create Archive Project task.

  This feature in IDEA maximizes disk space on our machine by archiving Desktop projects that are no longer used. We can also use this task to share Desktop projects between IDEA users. Archived projects contain all the IDEA databases and associated Library files for a selected project. They are compressed into a zipped file (*.iarc/iarc2) and stored in a default or user-defined location. We can choose to delete the Desktop project folder after the archive process is complete to free up valuable hard disk space on PC.

  Here is information about Archive Project from IDEA Help:

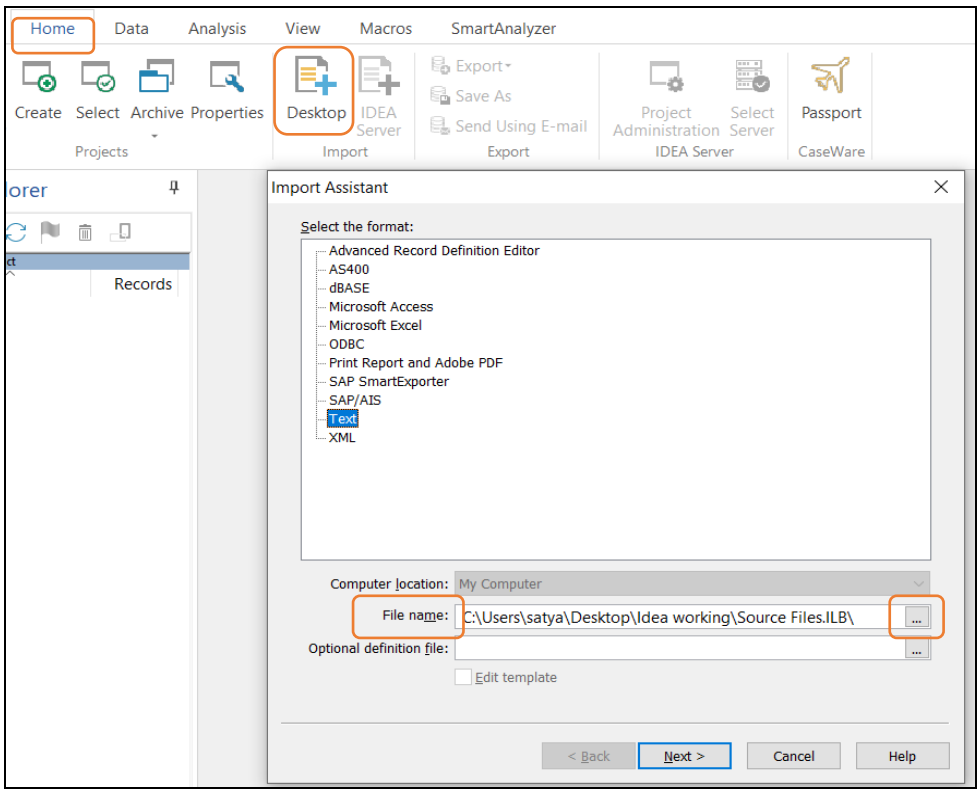
  - It is recommended to only archive Desktop projects that reside on local drive
  - Archiving Desktop projects on mapped network drives is not supported
  - Locked Desktop projects can only be archived by the user who locked the projects
  - The active Desktop project cannot be archived
  - The Archive tasks are temporarily unavailable and grayed out while other tasks are running in the active Desktop project

- **Project Overview:** This feature is discussed in Para 9.6

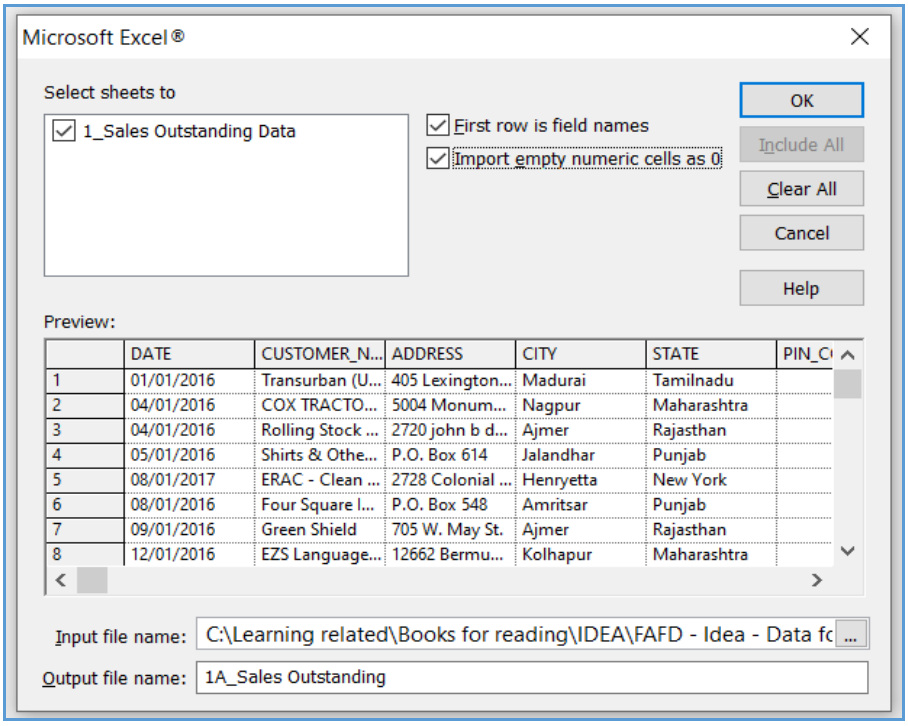
9.3 IMPORT OF FILES IN IDEA

- **Import of file using Desktop option:** Using the **Home Tab > Import > Desktop**, we can import various type of files in IDEA using the **Import Assistant**.

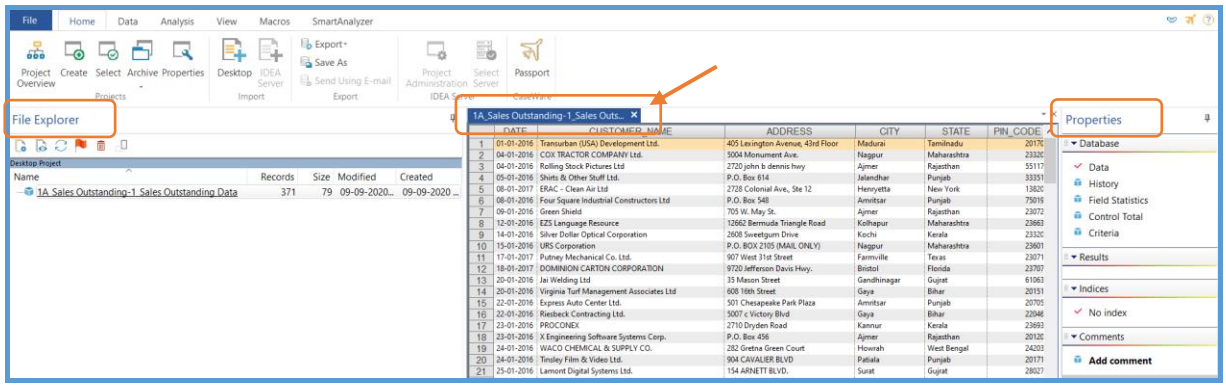
  The Import Assistant shows the various types of files, which we can open import in IDEA. For working perspective, we are going to import 3 different types of files i.e. Excel, Text and PDF.



- **Import of Excel file:** We select the file using the "File Name" option. Then we click on Next button. Then a new menu opens. It is similar to Text-to-Column wizard in Excel. We can see the list of the sheets in the current excel file. Currently, we have only one sheet naming "1_Sales Outstanding Data" We also need to tick on "First row is field names" and "Import empty numeric cells as 0" as per need. Then click on OK.



Once we click OK, the said excel file will be imported in IDEA as shown below.

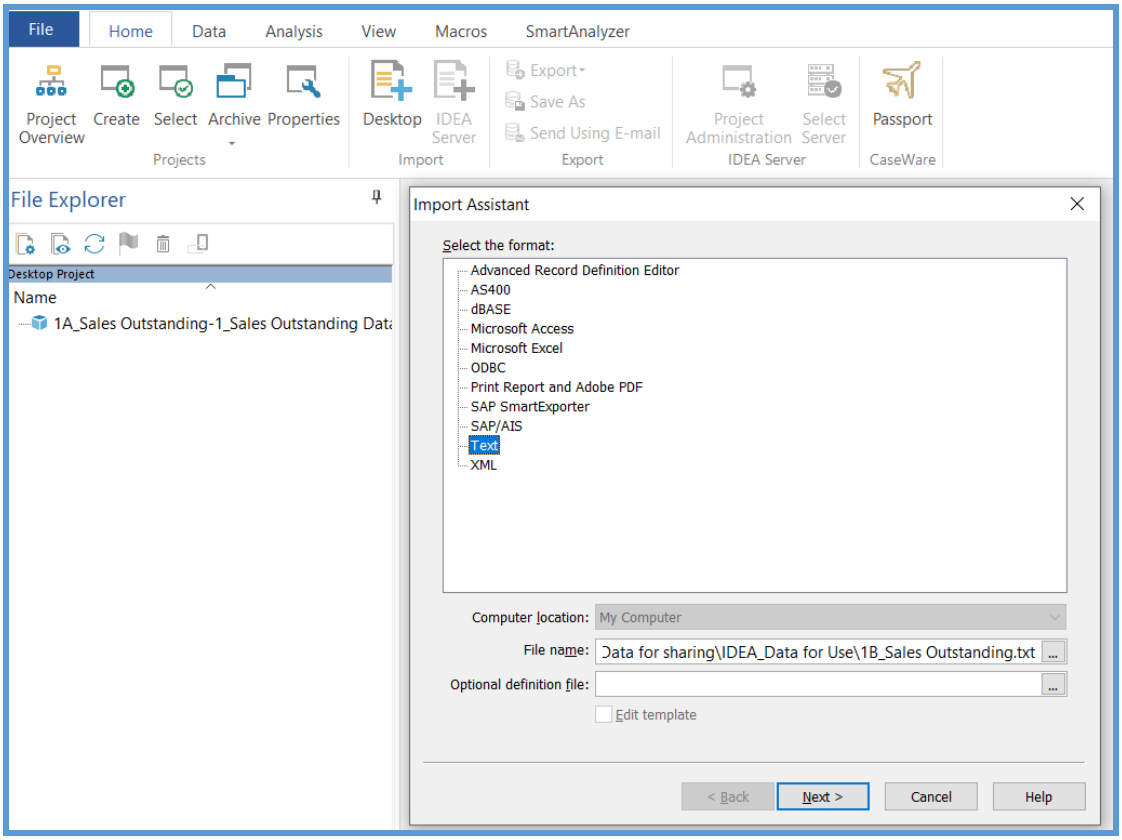


We can see following:

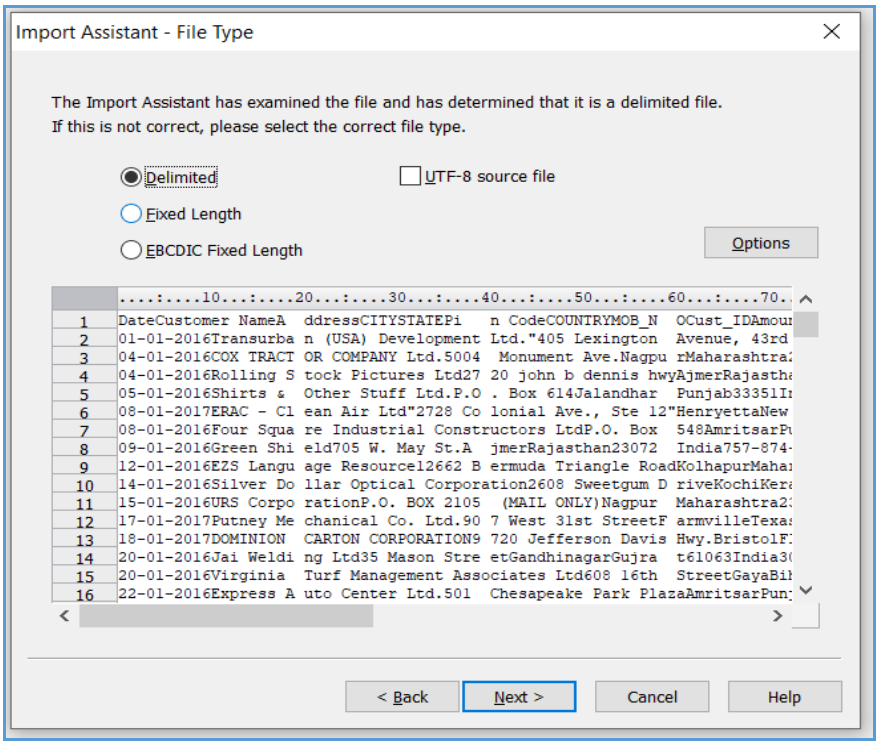
- We get File Explorer view, where we can see the file details - the file name followed by sheet name, Number of records, Size, Date modified, Date Created.
- In the centre, we will get the active sheet. Here it is "1A_Sales Outstanding Data". In case, we don't need the file, we can close it by clicking at "x" button on right side of file name. If we want to open any file, we need to double click on that file name in File Explorer.
- On the right side, we get the Properties of the current opened file.

Similar steps, we need to follow to import other types of files.

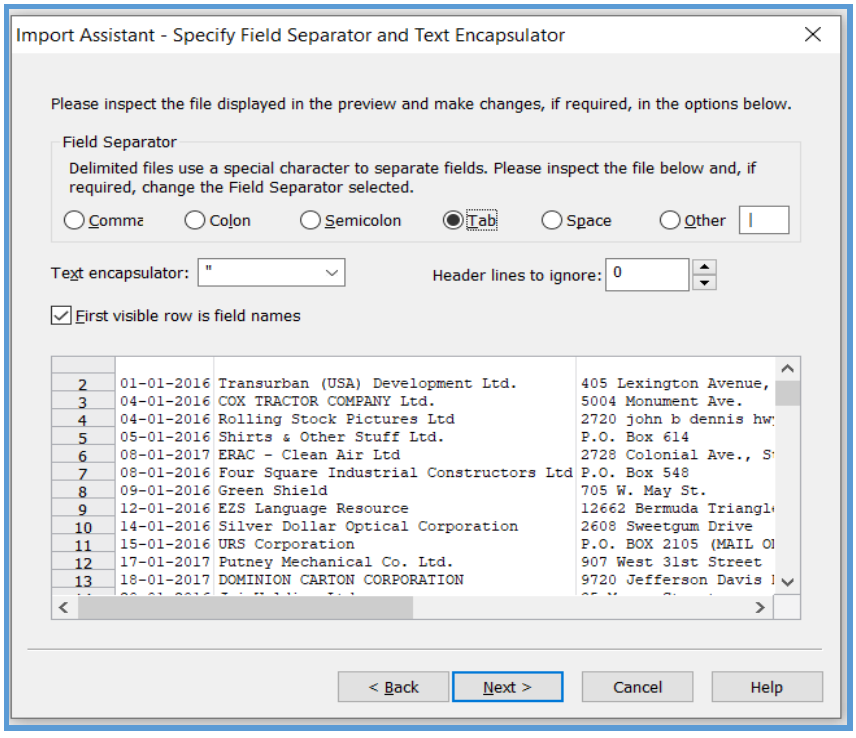
- **Import of Text file:** We click on Desktop button in Home tab and Import Assistant opens. There we select the format of file to be opened and select the file using "File Name" option.



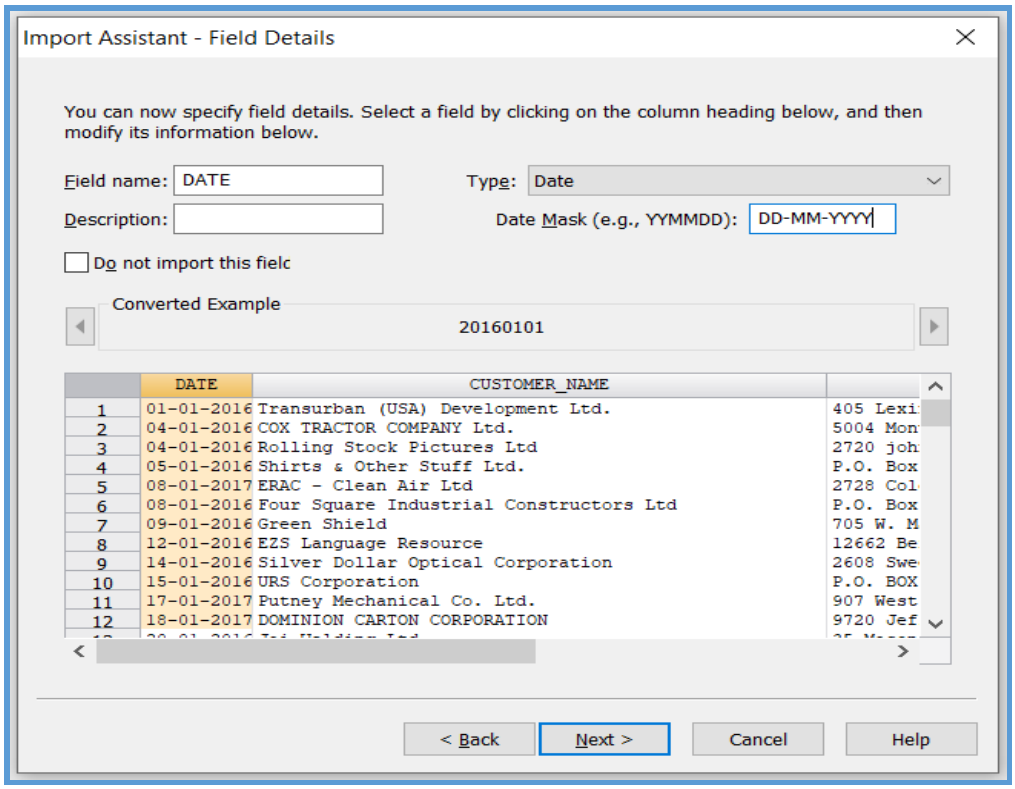
Once we click Next button, it opens "Import Assistant – File Type" option. Like Text-to-Column Import wizard in Excel, we need to select type of file.



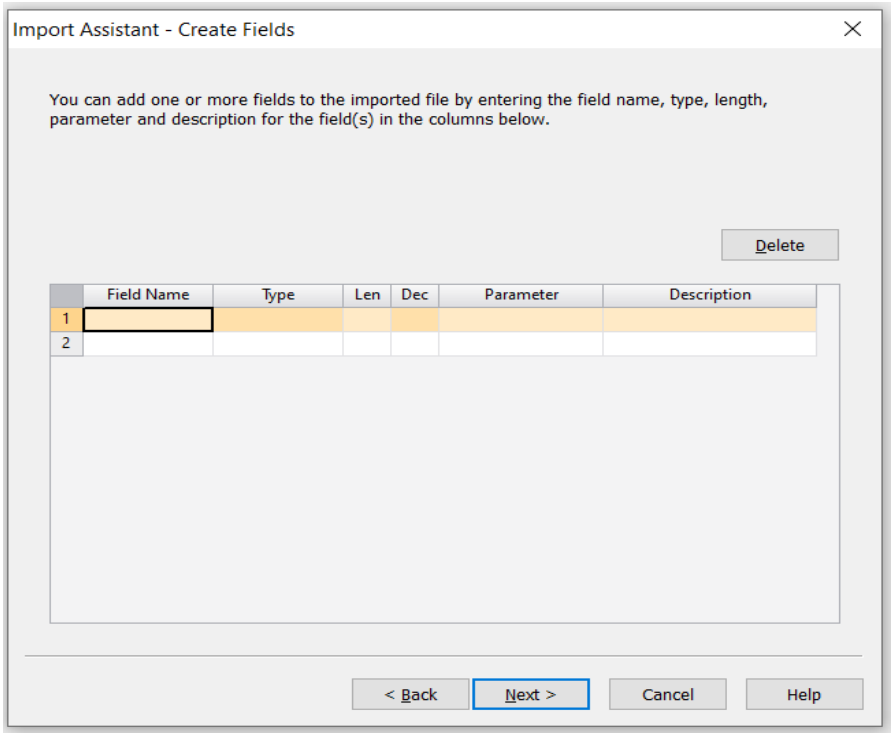
We need to select the separator and also need to click on “First visible row is field names”. Then click Next.



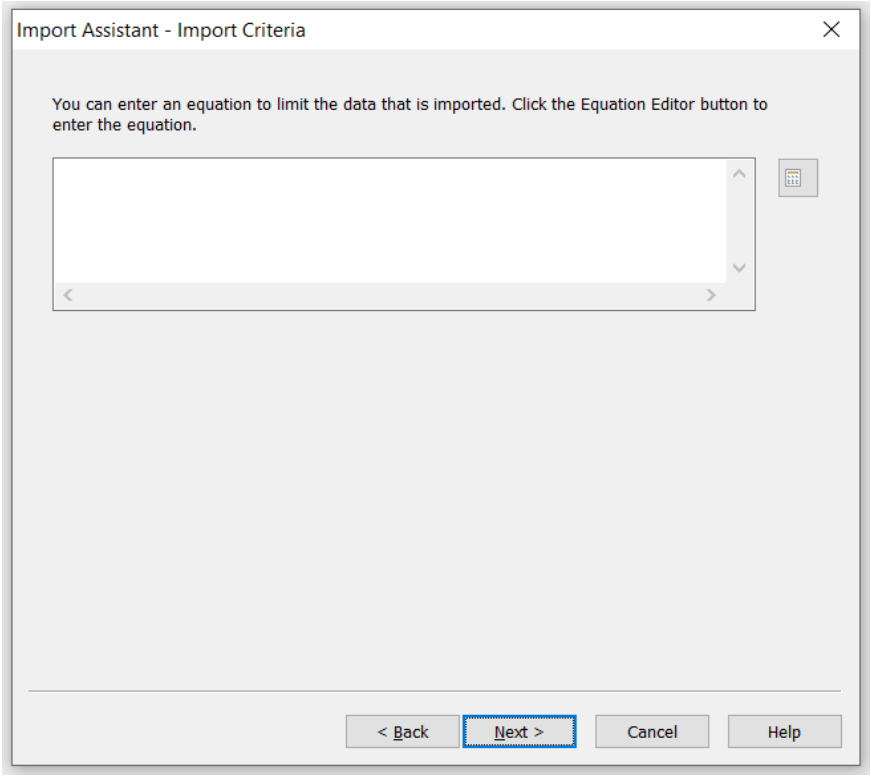
It will open on “Import Assistant – Field Details” and we need to define the **field(column)** name, Type of file (Character, Numeric, Date and Time). In case we don’t want to import any particular field/column, we can tick the option “Do not import this field”. The click Next.



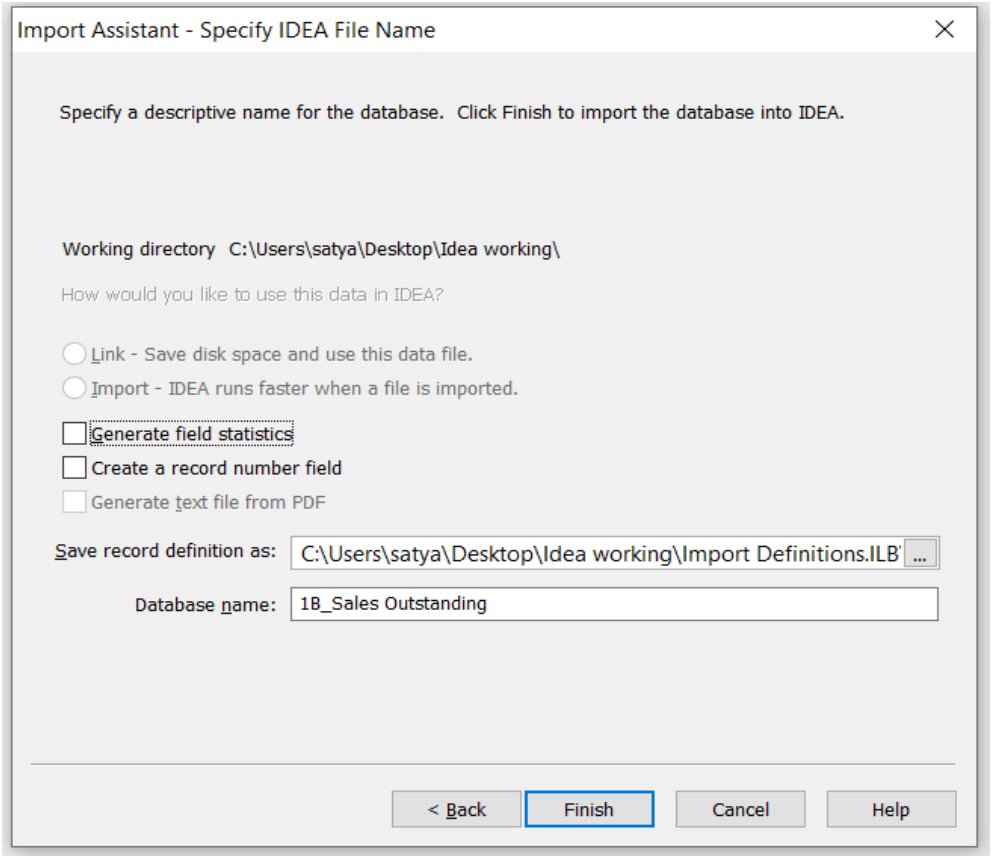
It will open the “Import Assistant – Create Fields”. In case, we want to add one or more fields, we can use this option.



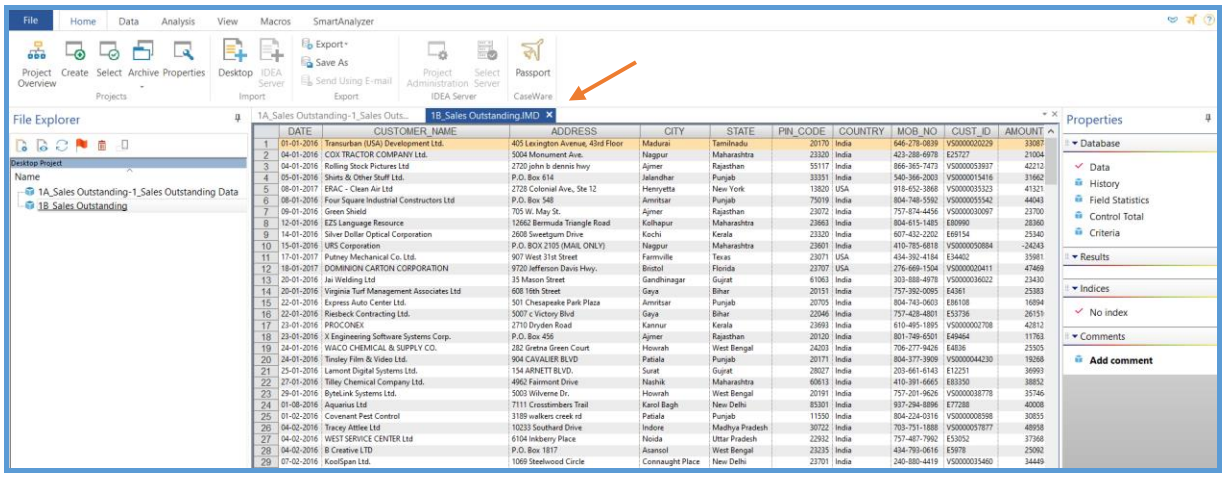
Then it will open “Import Assistant – Import Criteria” as shown below.



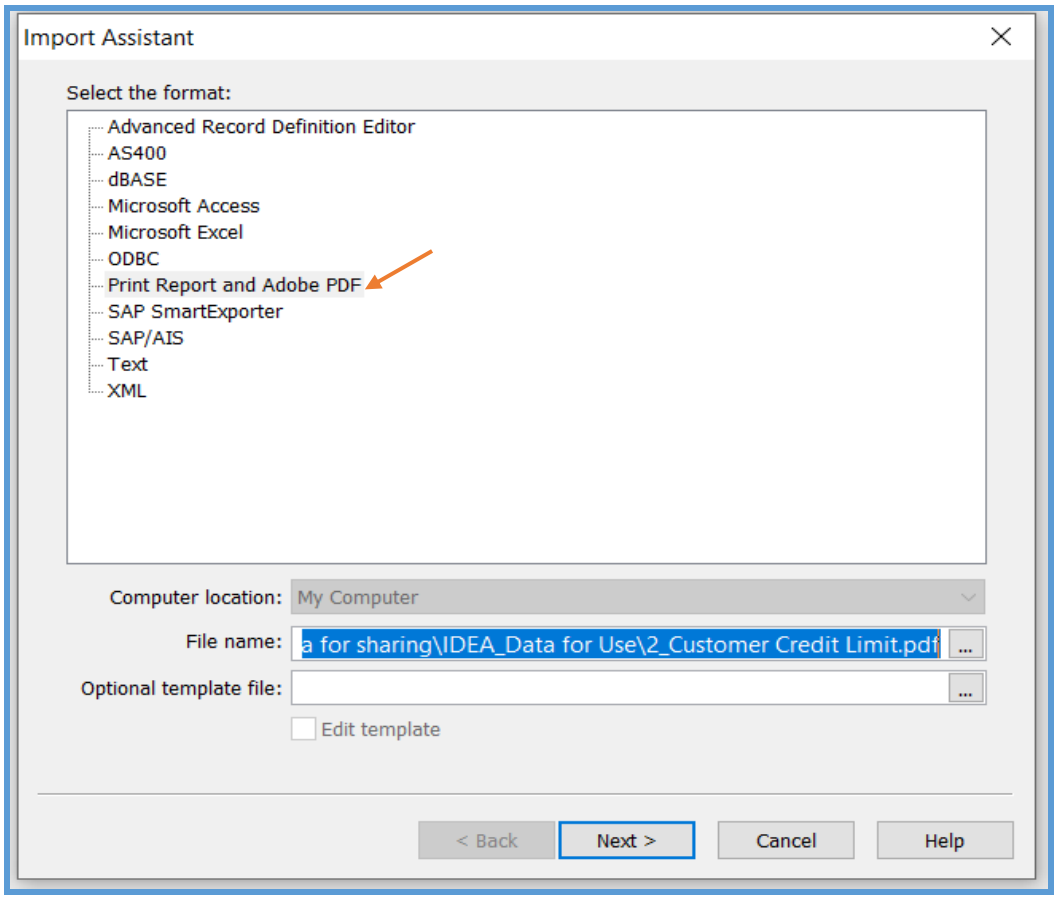
Then it will open “Import Assistant – Specify IDEA File Name”.



Once we click on Finish button, it will import the file in IDEA as shown below. On left side in File Explorer, now we have 2 files i.e1A_Sales Outstanding (Excel file) and 1B_Sales Outstanding (Text file). The active file will be highlighted by dark blue colour. On the right side, there is properties tab, which contains information about the active file.

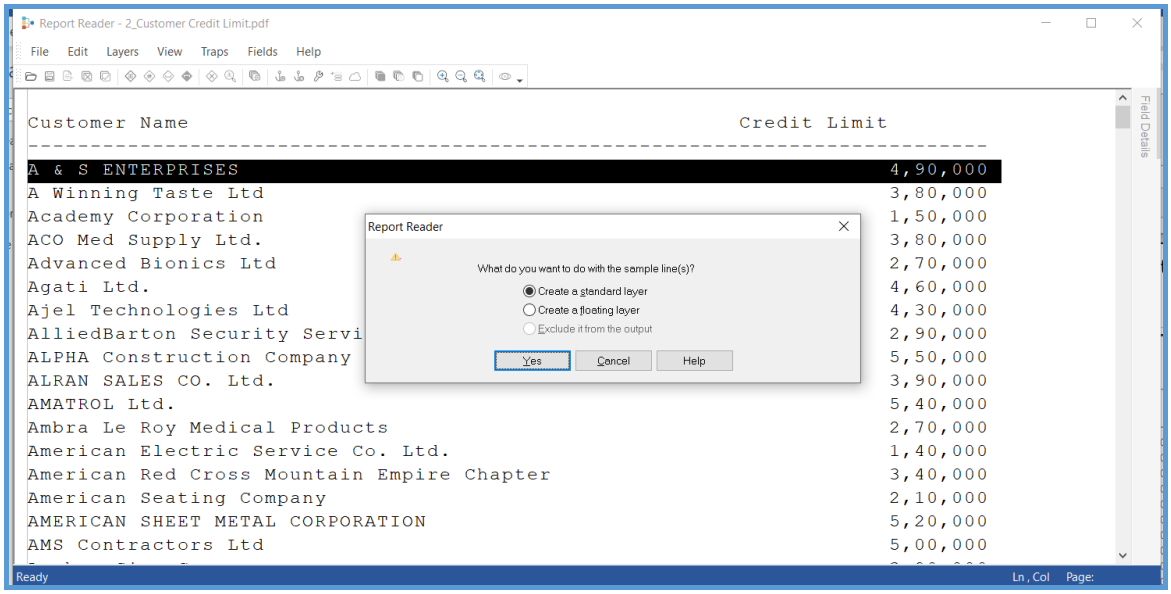


- **Import of PDF file:**We click on Desktop button in Home tab and Import Assistant opens. There we select the format of file to be opened and select the file using "File Name" option.

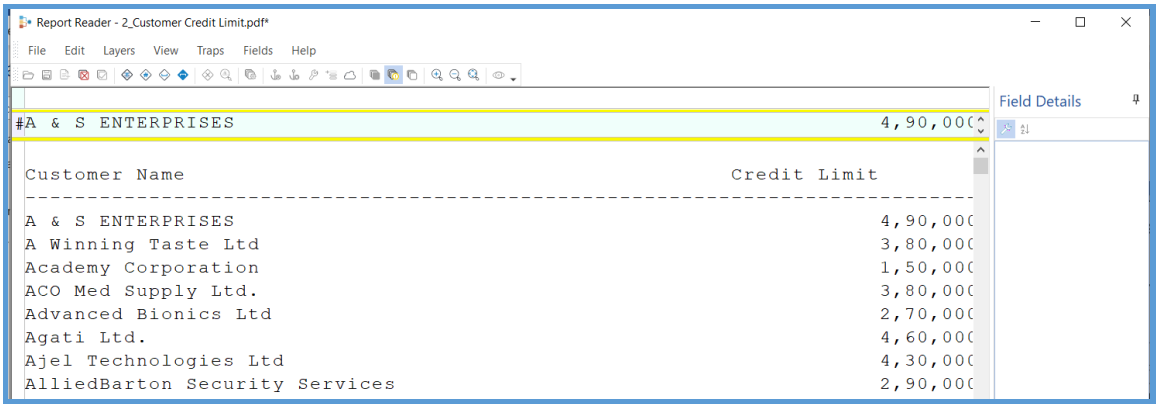


Once we click on Next button, it will open "Report Reader – 2_Customer Credit Limit.pdf". It is a 2 column file, which contains Customer Name and Credit Limit. Now we need to define the columns and its width.

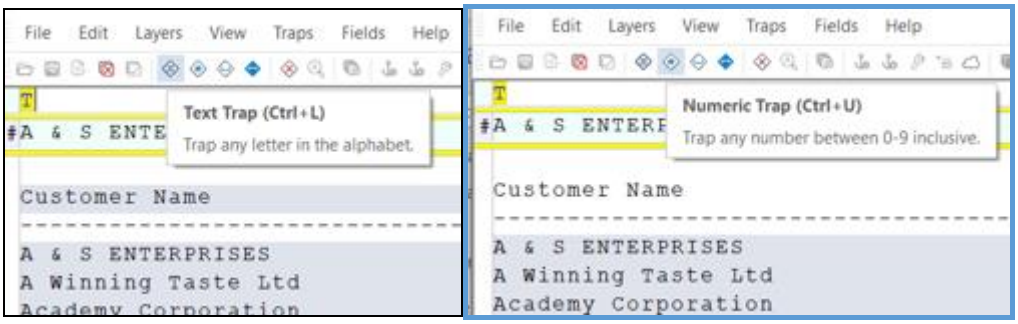
We need to select the first row (as highlighted below) form left most character. It will ask for layer information. We will select first option i.e. “Create a standard layer”.



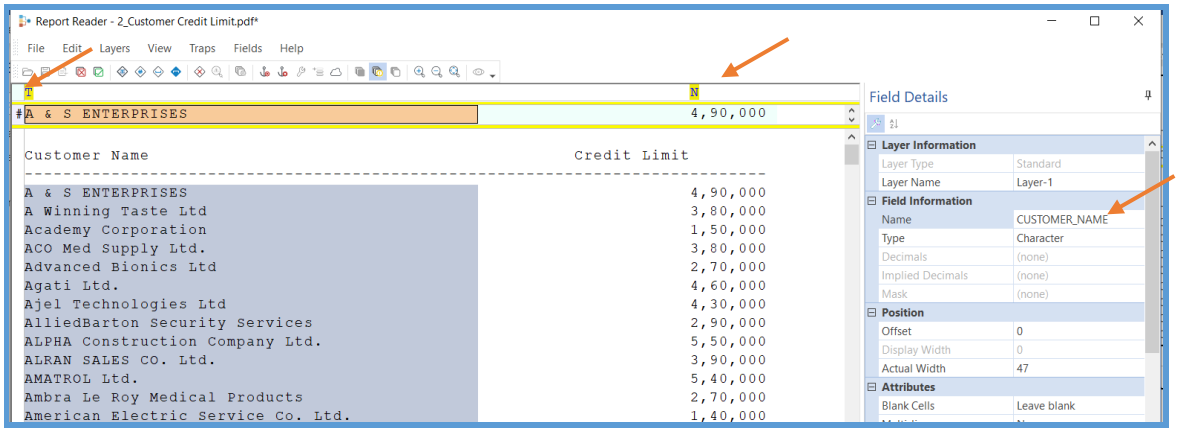
It will take the first row at top of the table as a sample line for defining columns.



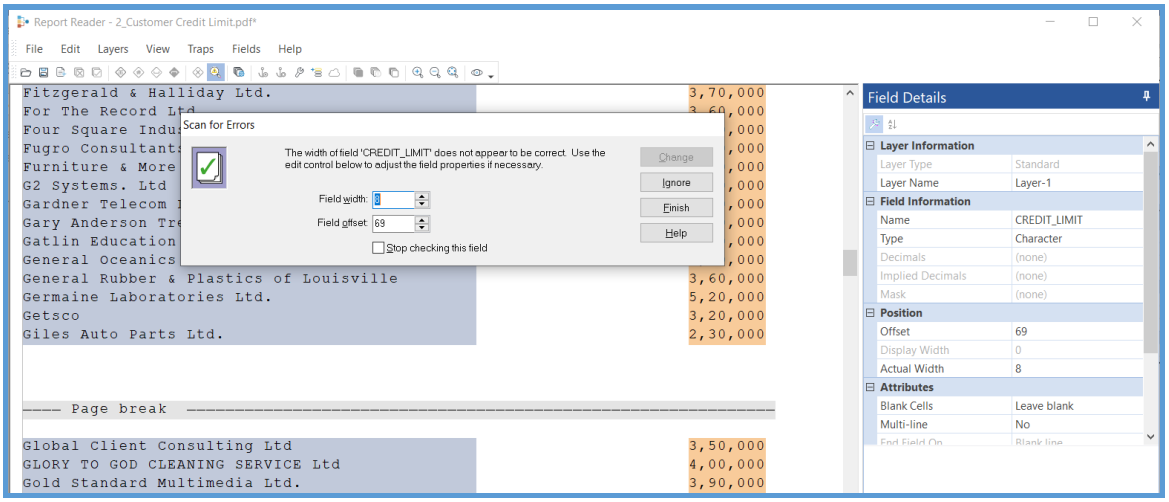
Now we need to put the Traps (Text Traps for text column and Numeric Trap for numeric column). Also we need to define the field nature on right side of the screen.



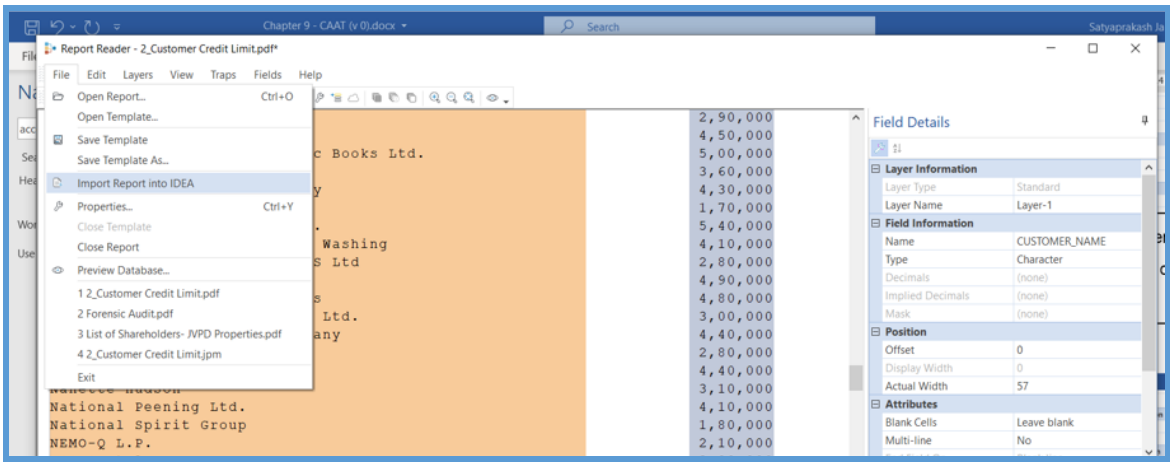
The we need to select the approximate width of eachfield as shown below:



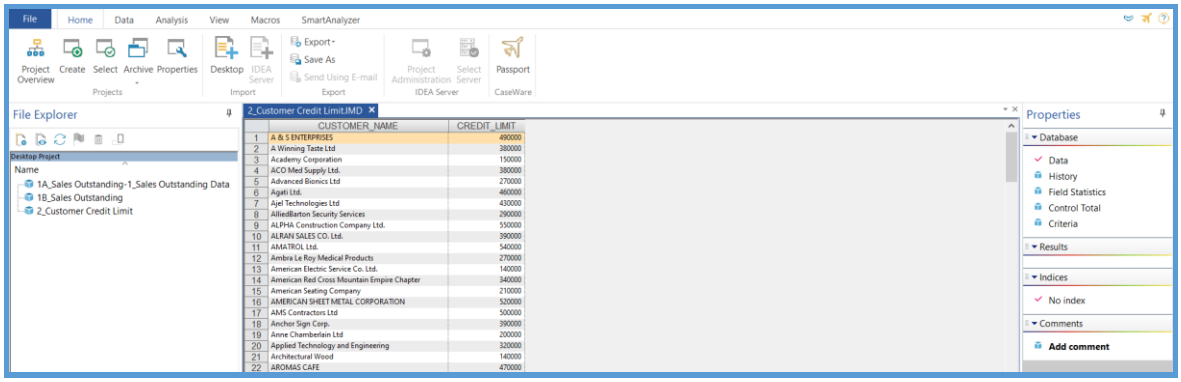
Once all field width, Field Name and Type defined (in Field details), we need to Save Layer. Then go for Scan for Errors. It will give option to increase or decrease the field width. Post which click on Finish and Save the Template.



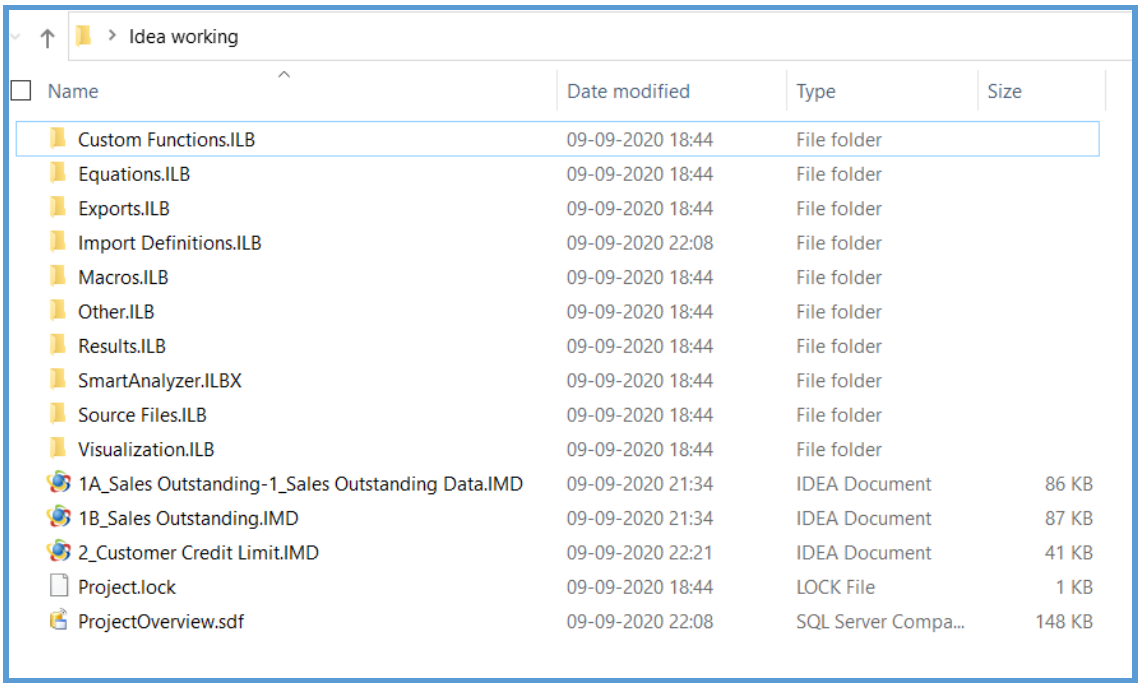
Then select the option “Import report into IDEA”.



It will load the file in IDEA as shown below:



After import of all these files, IDEA will automatically create imd files in the “Idea working” project folder. Like this, as we load more files in IDEA, it will create imd files in the project.

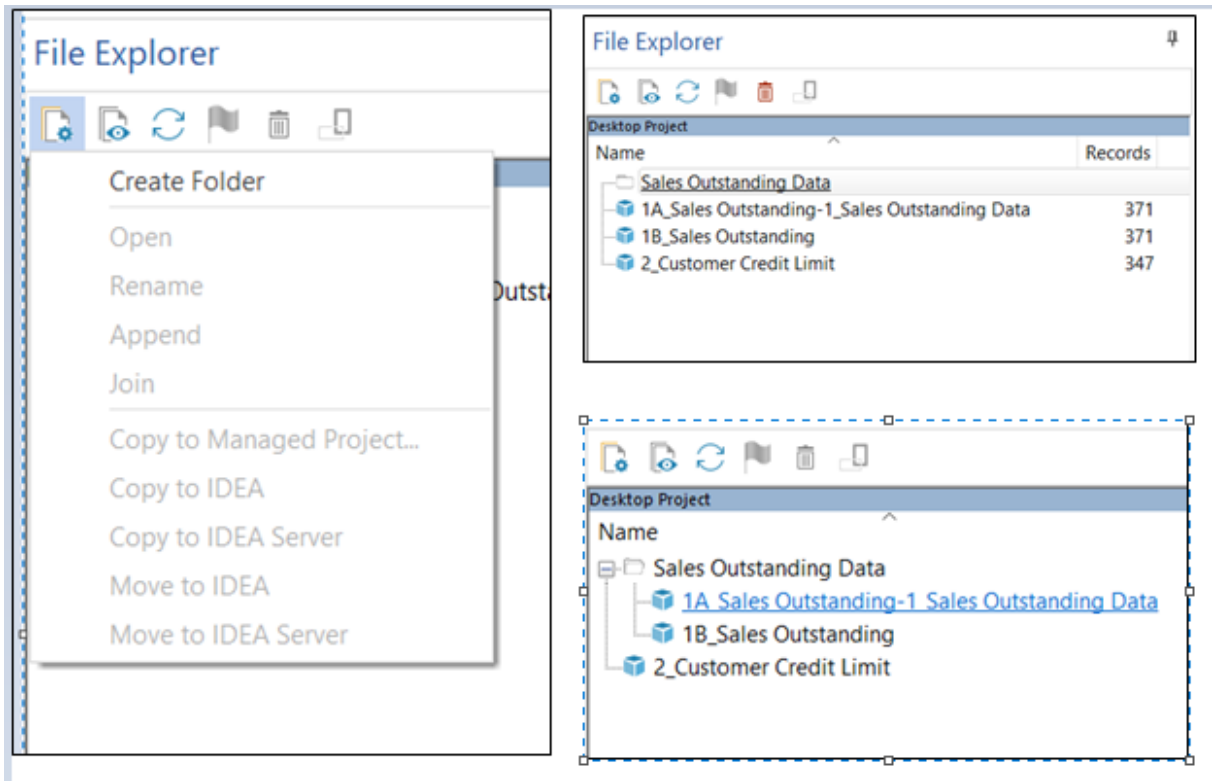


Now we have 3 files in File explorer on left side and the active file is highlighted in dark blue colour with file name in centre.

It is to be noted that whatever data has been imported in IDEA, original data can not be modified. No cell value can be changed. We can add only new fields.

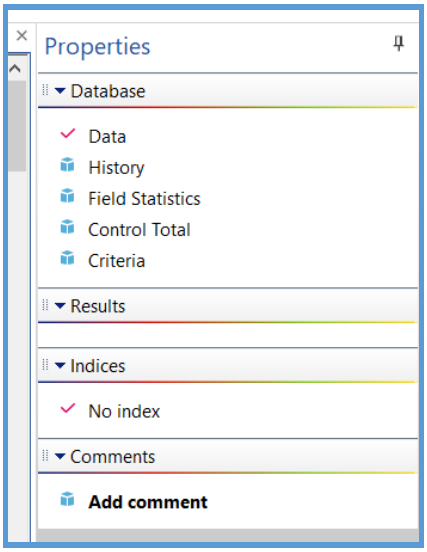
9.4 FILE EXPLORER

Using File Explorer, we can create folders and move the files to respective folders. In the below image, we created a folder named “Sale Outstanding Data”. Then we moved the files “1A_Sales Outstanding” and “1B_Sales Outstanding” to the folder. We have Refresh, Delete options as well. In case, we want to delete any file, we can use this option.

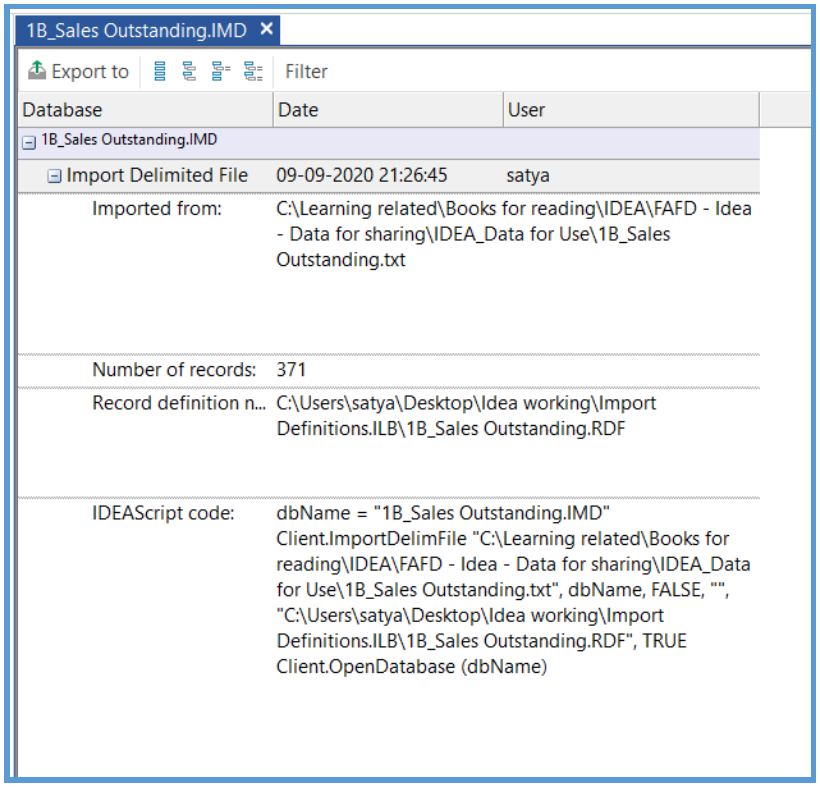


9.5 FILE PROPERTIES

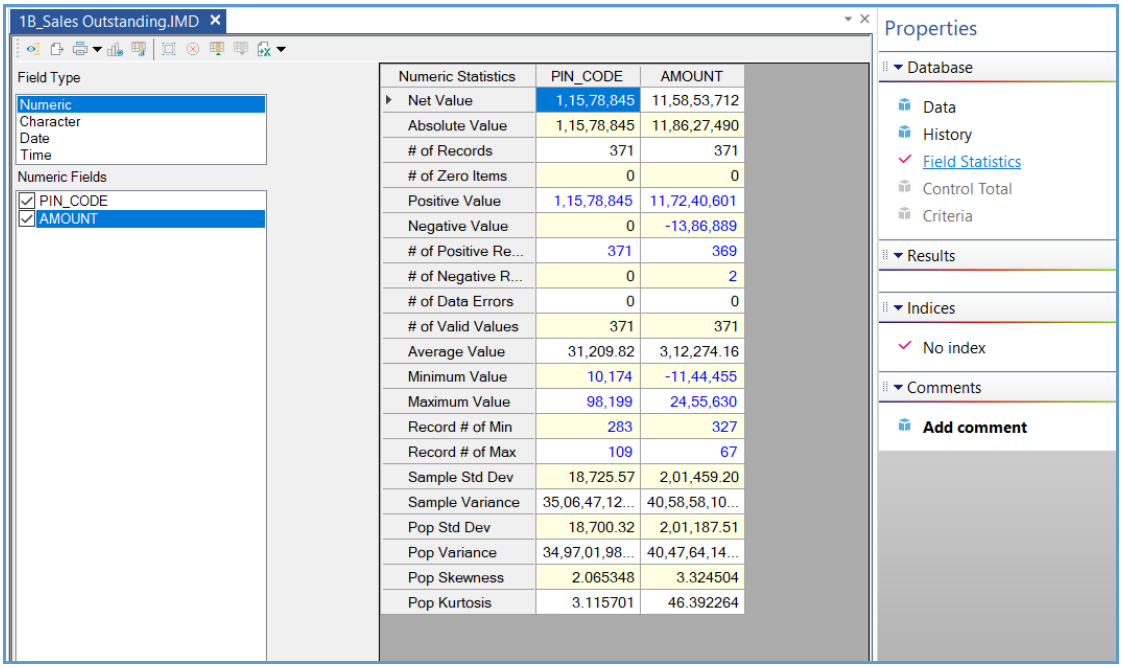
Once we open any file, the file properties option gets activated. In this, we get following options: Data, History, Field Statistics, Control Total and Criteria. By clicking each option, we can review data.



- **Data:** It is used to view the data in the file in a spreadsheet-like format with field names as field headings and record numbers as row numbers.
- **History option** records all the steps performed in that particular file since inception. All these steps can not be undone/modified/deleted. This History works as an audit evidence and accepted in Court as well. Till now, in this file, we have only imported this file and this step has been recorded in the History as shown below:

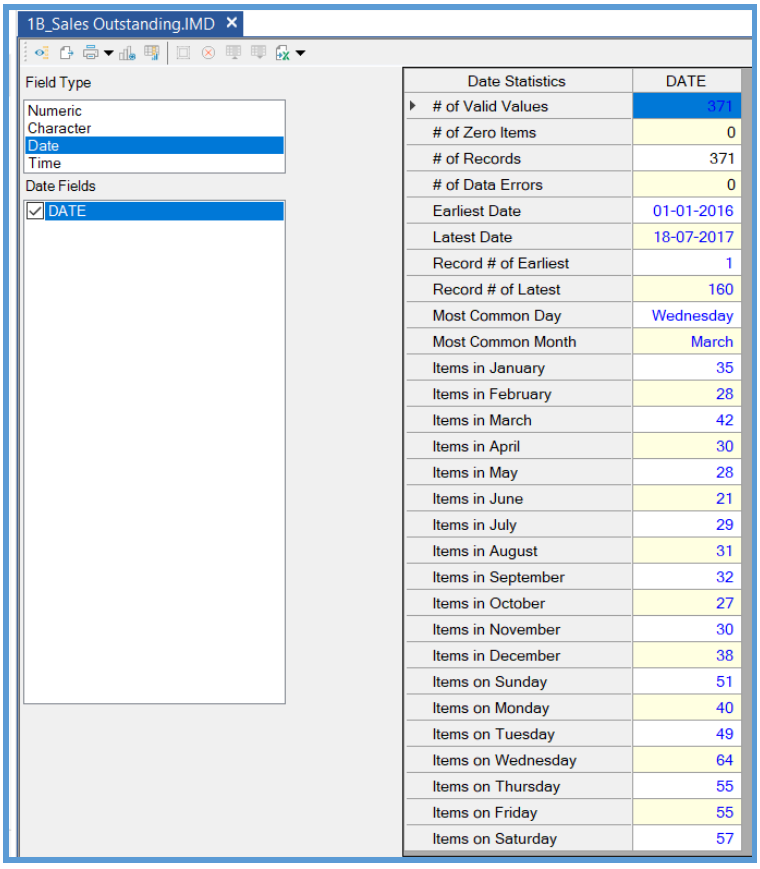


- **Field Statistics** option shows statistics for each field in the particular file. If we observe the statistics properly, we can make some initial observations from these statistics like :
  - Minimum and Maximum Value in Amount column,



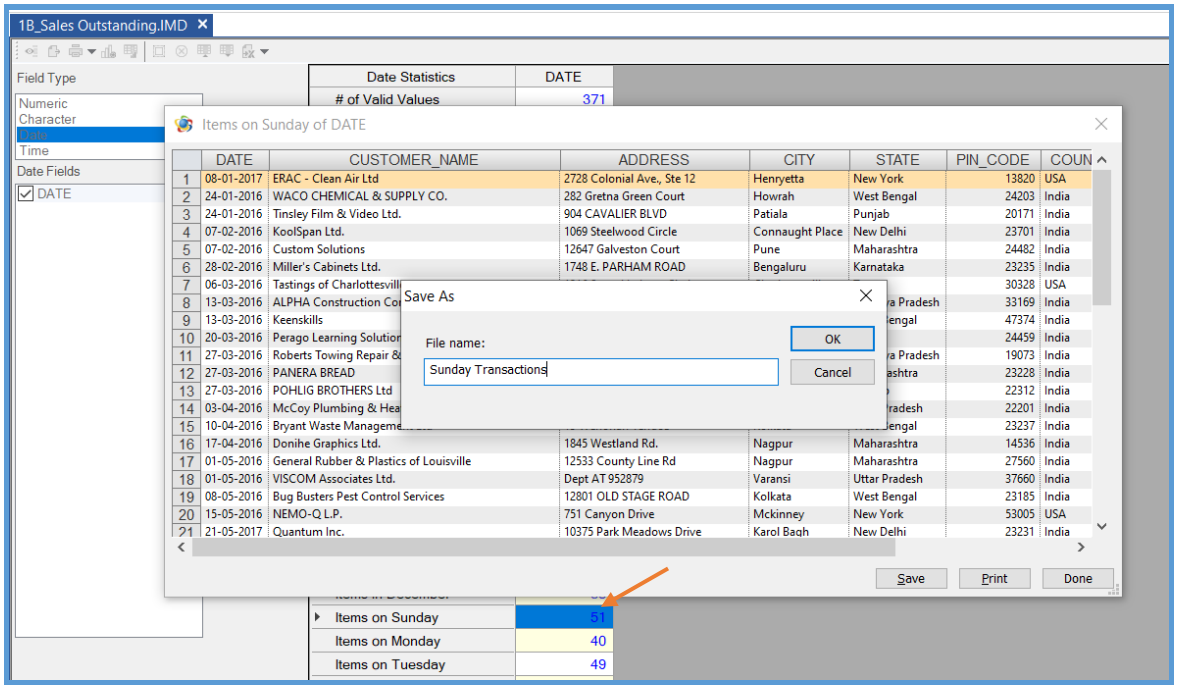
Numeric Statistics	PIN_CODE	AMOUNT
Net Value	1,15,78,845	11,58,53,712
Absolute Value	1,15,78,845	11,86,27,490
# of Records	371	371
# of Zero Items	0	0
Positive Value	1,15,78,845	11,72,40,601
Negative Value	0	-13,86,889
# of Positive Re...	371	369
# of Negative R...	0	2
# of Data Errors	0	0
# of Valid Values	371	371
Average Value	31,209.82	3,12,274.16
Minimum Value	10,174	-11,44,455
Maximum Value	98,199	24,55,630
Record # of Min	283	327
Record # of Max	109	67
Sample Std Dev	18,725.57	2,01,459.20
Sample Variance	35,06,47,12...	40,58,58,10...
Pop Std Dev	18,700.32	2,01,187.51
Pop Variance	34,97,01,98...	40,47,64,14...
Pop Skewness	2.065348	3.324504
Pop Kurtosis	3.115701	46.392264

- Earliest date, Latest Date, Day wise and month wise concentration in Date column,

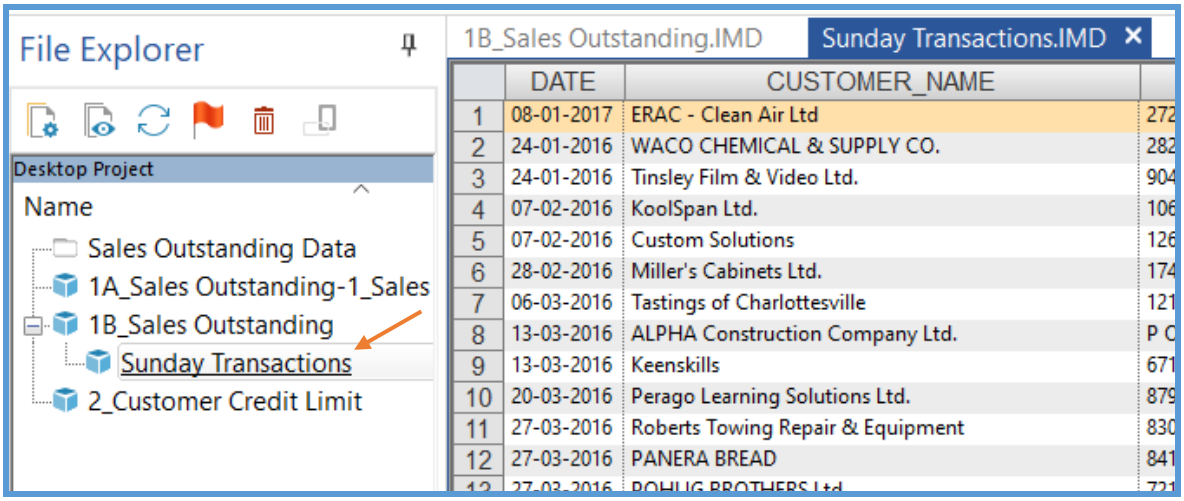


Date Statistics	DATE
# of Valid Values	371
# of Zero Items	0
# of Records	371
# of Data Errors	0
Earliest Date	01-01-2016
Latest Date	18-07-2017
Record # of Earliest	1
Record # of Latest	160
Most Common Day	Wednesday
Most Common Month	March
Items in January	35
Items in February	28
Items in March	42
Items in April	30
Items in May	28
Items in June	21
Items in July	29
Items in August	31
Items in September	32
Items in October	27
Items in November	30
Items in December	38
Items on Sunday	51
Items on Monday	40
Items on Tuesday	49
Items on Wednesday	64
Items on Thursday	55
Items on Friday	55
Items on Saturday	57

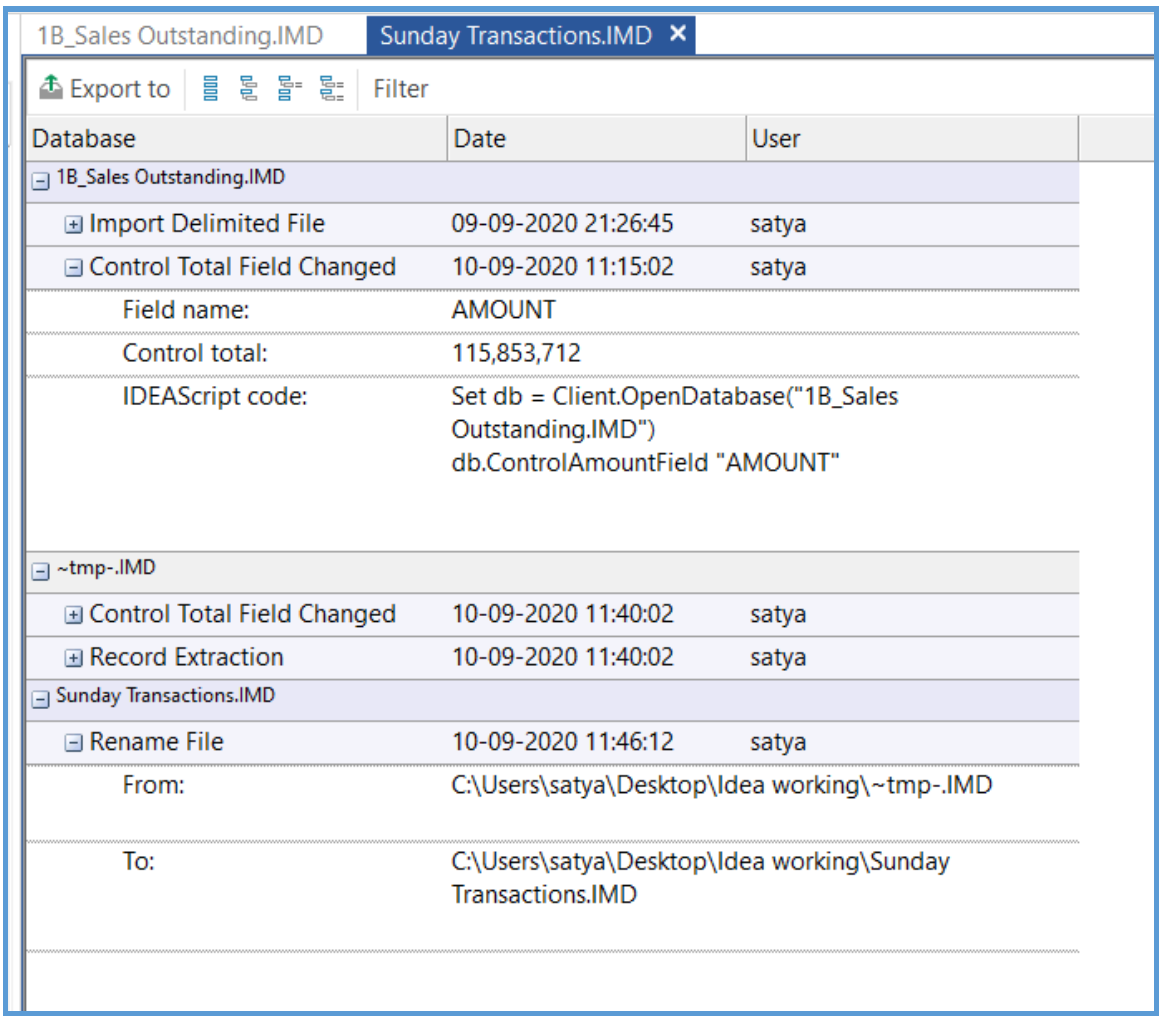
In case, we want to extract any data, say we want list of transaction done on Sunday, we need to click on hyperlink (highlighted in blue) and it will open a new window and we can SAVE / PRINT / DONE (to close). We are going to save the list as Sunday Transactions for sharing with client for further information.



Once we save the file as Sunday Transactions, it creates a thread below parent file and Child Parent relationship created in files and History has been updated accordingly.

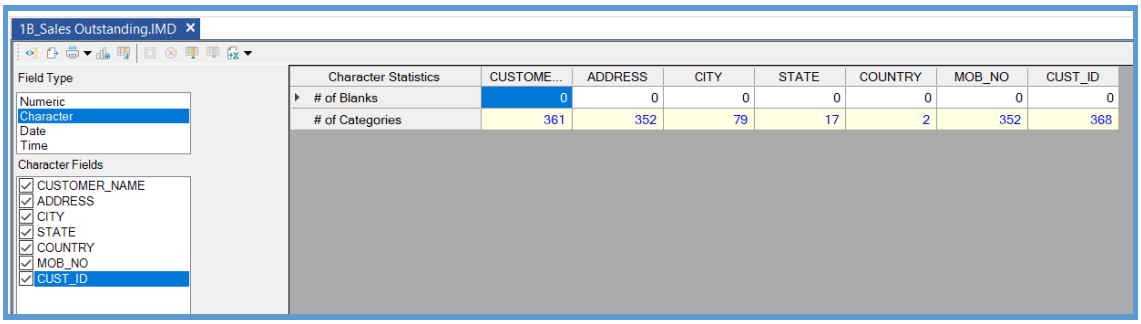


History update for Sunday Transactions:



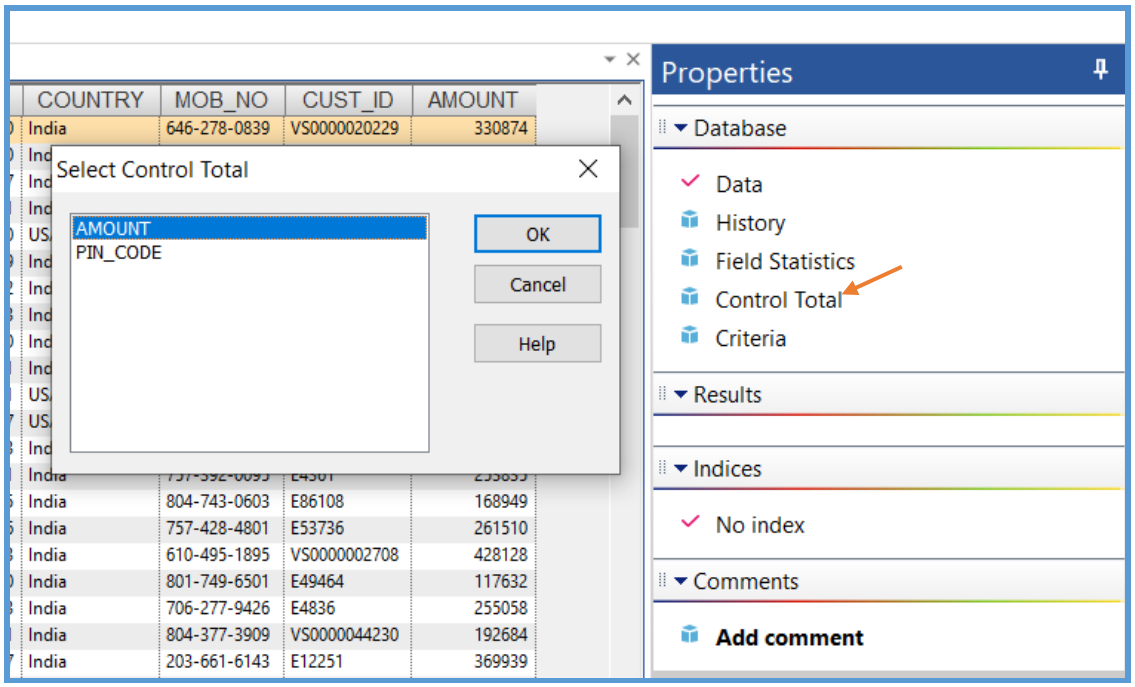
- Number of items in each field. Just by looking at all Character Field, we can try to identify the anomaly in data. In this example, we can see there are 361 line items in Customer Name column, whereas Mobile No. is only 352 and Cust ID is only 368. It shows that there are duplicate Cust ID and Mobile No., which we try to identify using Duplicate Option while doing Analysis.

  In case, we don't want any particular field, we can untick that field.

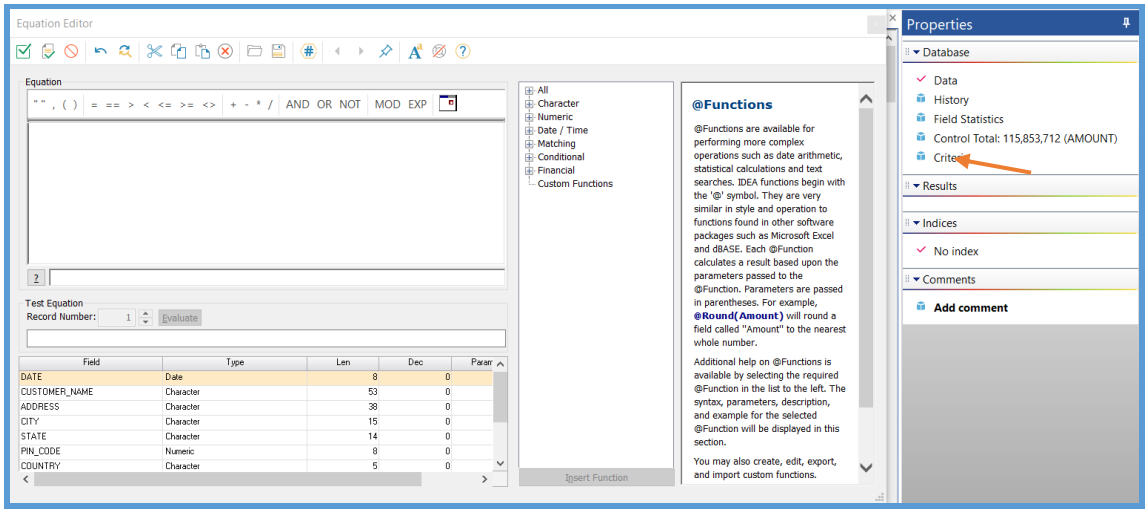


If we click on Data option, it will show the entire data. So whatever we do, original data cannot be manipulated.

- When we click on **Control Total** option, it will show option to select Numerical field for reconciliation purposes.

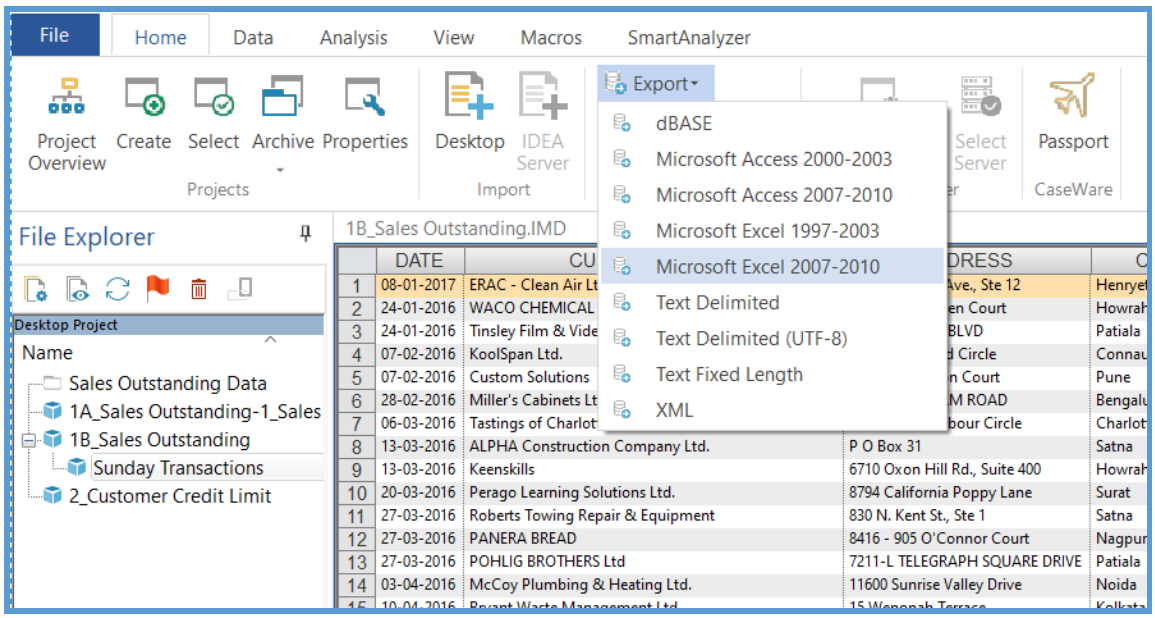


- When we click on **Criteria** option, it will open the Equation Editor option, which we can use for the purpose various calculation. We will see options in later part of chapter.

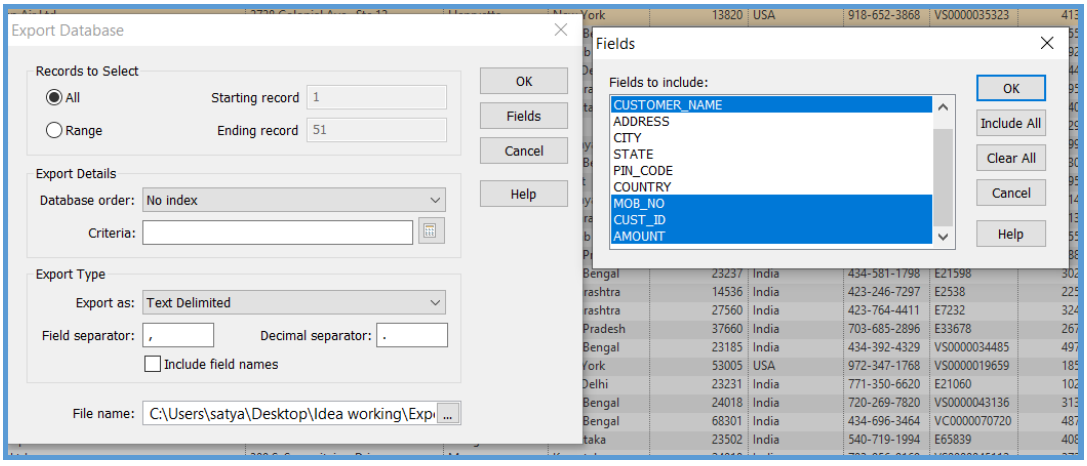


9.6 EXPORT OF FILES FROM IDEA

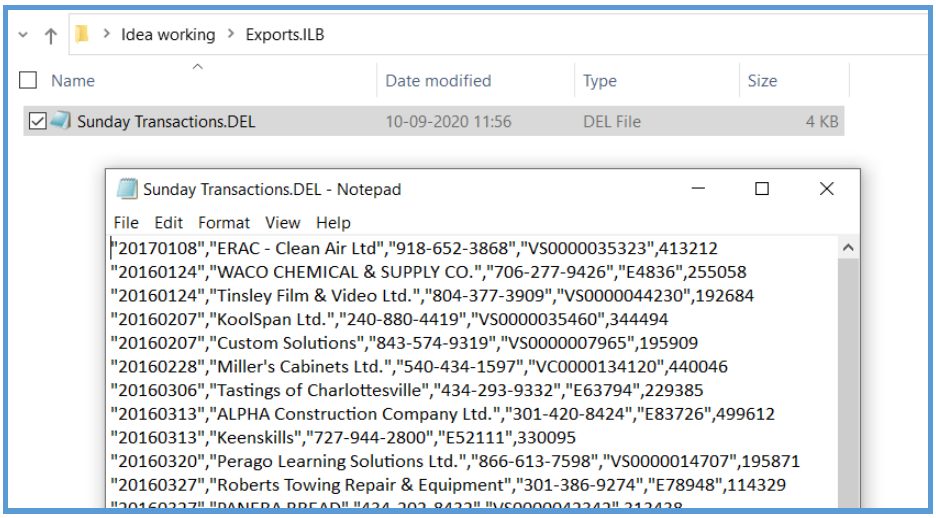
- **Export file:** We can Export the file in various formats as shown in the image. We are going to Export the Sunday Transactions file in Text Delimited format.



Once we click on Text Delimited option, it gives the option to store the data based on specified criteria, separator, location, etc. In Fields option, it gives the option to select/deselect the fields which we want in the output file as shown in the image.

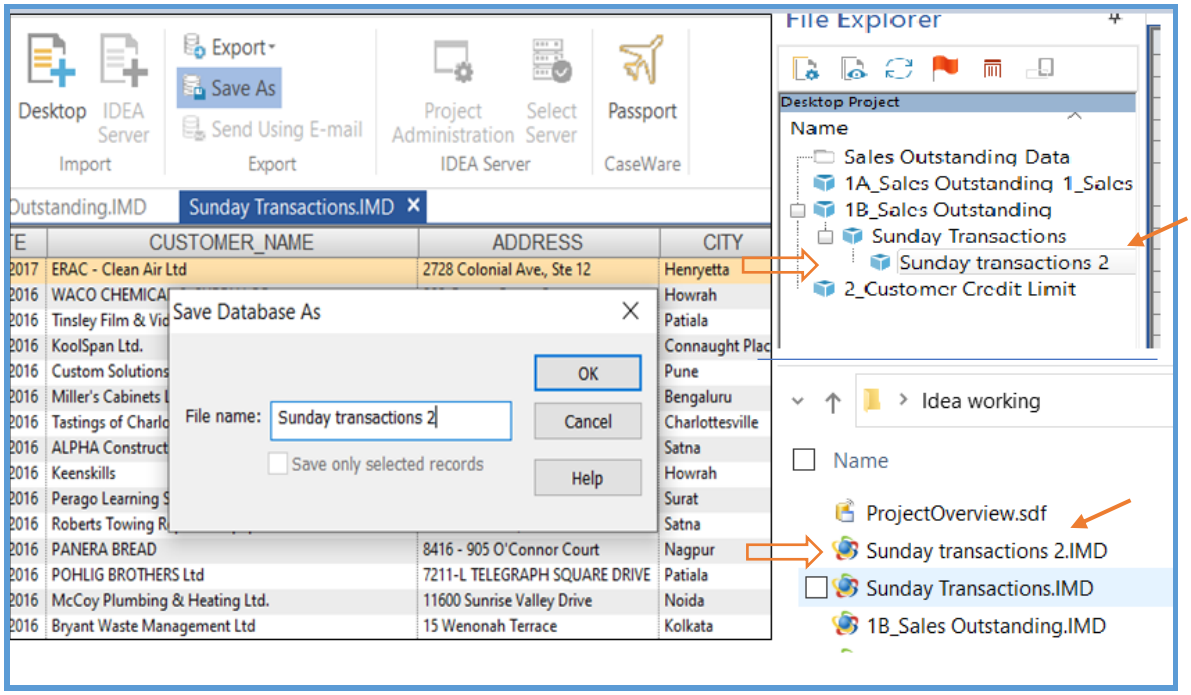


Once we click on OK button, the file is saved at the specified location. In this case, the default location is Exports. ILB in Idea working project.



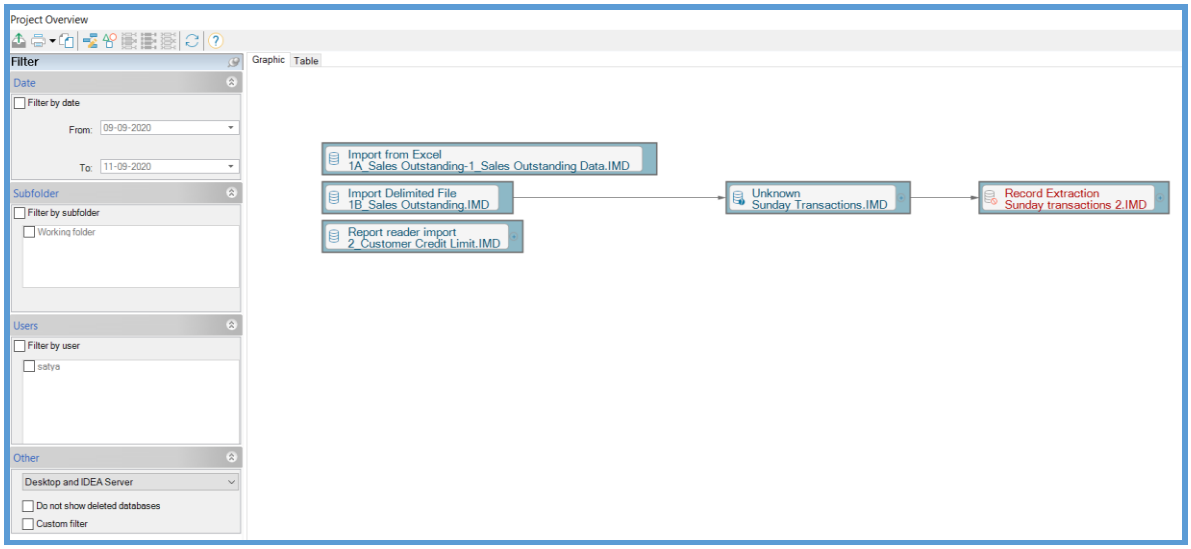
- **Save As:** When we click on Save As option, it gives the option to save the same file with a different name. We are saving Sunday Transactions file as "Sunday Transactions 2". By saving the file, another thread is created and it is also saved in the project folder.

  In case, we want to delete any file say (Sunday Transactions 2), then we can select the file click on delete button.



- **Project Overview:** This option is available in **Home Tab > Projects > Project Overview**. This feature records all the files created or deleted form the folder. Till now we have imported 3 files in IDEA. Created a child (Sunday Transactions) and another child (Sunday Transactions 2), which we deleted later on. In

the Project Overview, we can see the Names of file imported. Link of Child and Grand Child. It also highlights the deleted file as well in red colour.

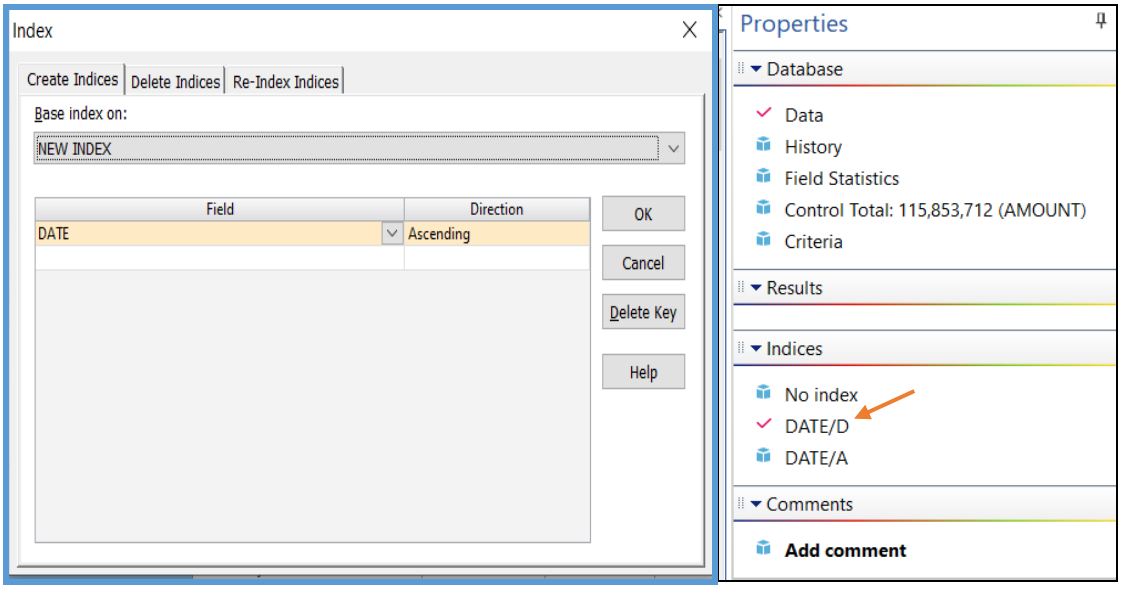


9.7 CLIPBOARD - COPY

This feature is available in Data Tab > Clipboard > Copy. Using this feature, we can copy the data and paste in another file (say word, excel).

9.8 ORDER

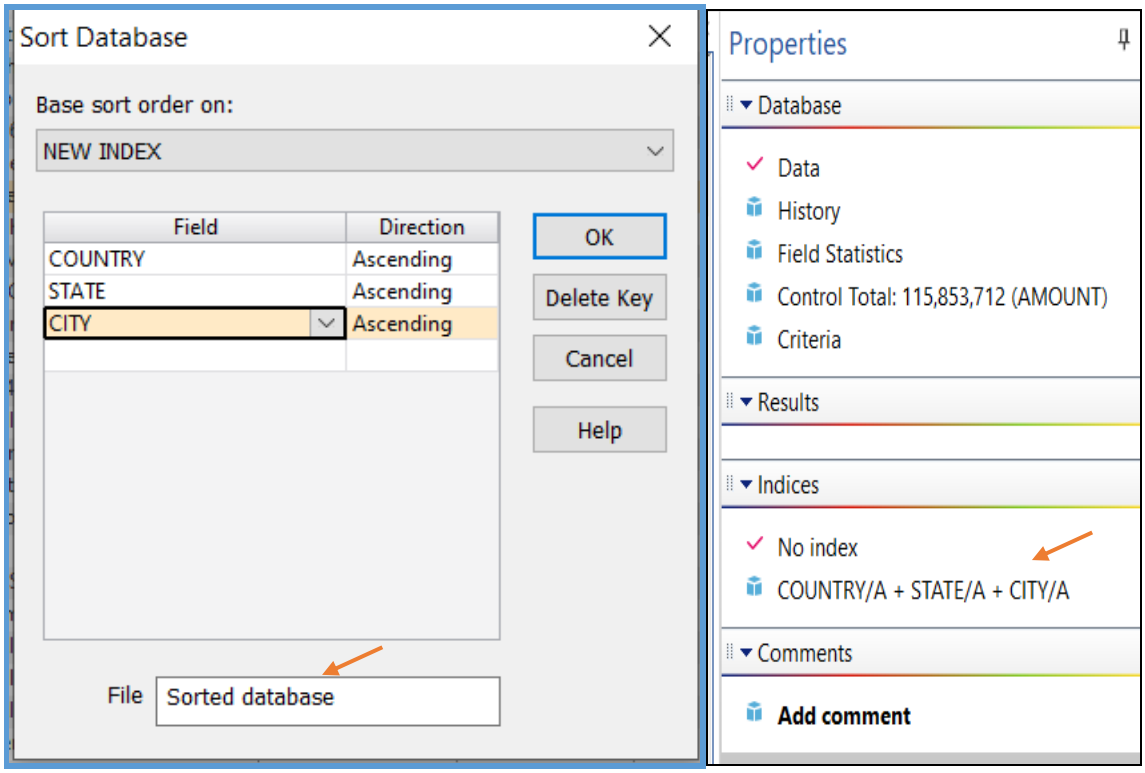
- **INDEX:** This feature is available in **Data Tab > Order > Index**. This feature helps to create, maintain or delete indices for the current database. Once we click on Index option, It gives 3 option i.e. Create Indices, Delete Indices and Re-Index Indices.



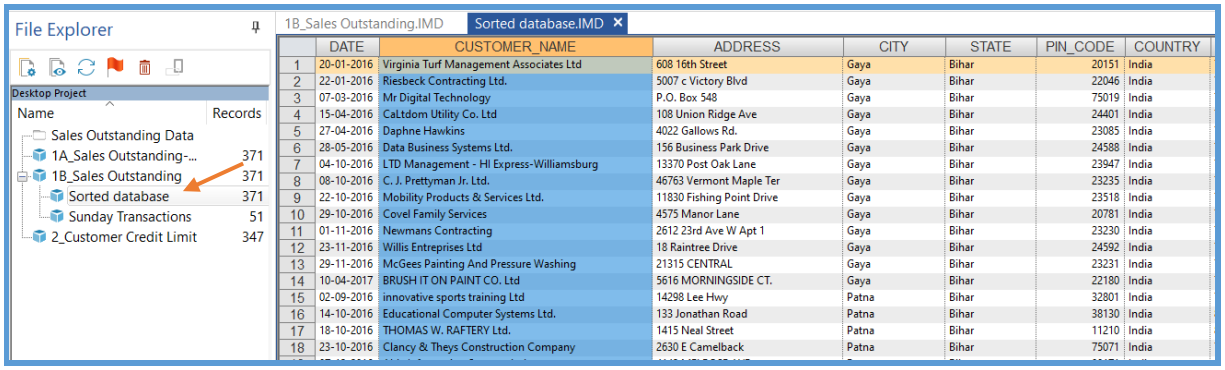
In the above image, we can see the field based on which we can create an Index. The same Index can be sorted in Ascending and Descending order. Also we can add up to 8 levels of field on below the other. In case, we want to delete any indices, we can click on option "Delete Indices" and select the Indices for deletion.

The list of indices created will be listed in the Indices under Properties head on the right side of the screen. Here we have created 2 indices – "DATE/A" and "DATE/D". Here Date is the column name and D stands for Descending and A stands for Ascending order. We can change view to "No Index", "DATE/D" and "DATE/A".

- **SORT:** This feature is available in **Data Tab > Order > Sort**. Using this option, we can sort the data. Like Index, we can sort upto 8 level of sorting. In below example, we have sorted 3 level of sorting.



It is to be noted that while sorting a new file "Sorted database" (a default name) is created in the File Explorer.

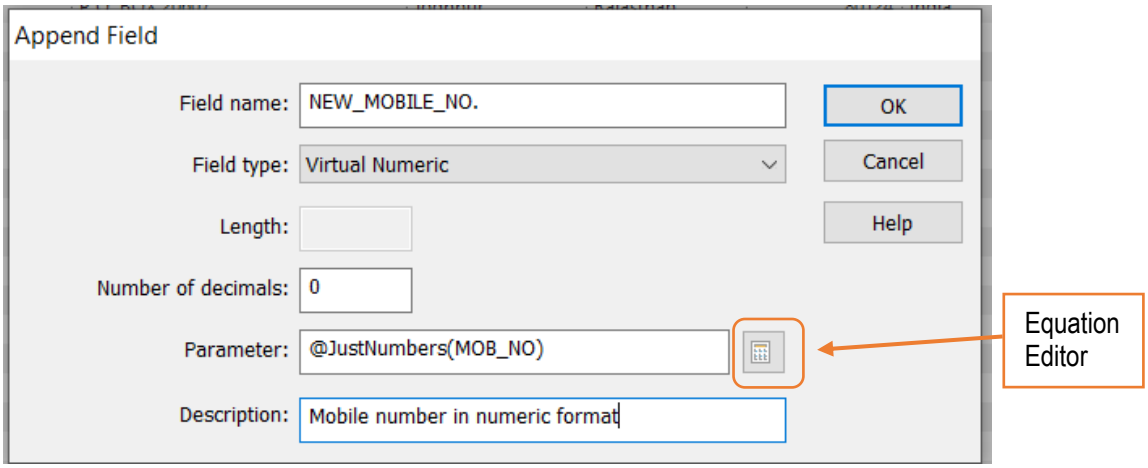


We can do the sorting just by double clicking on column name also.

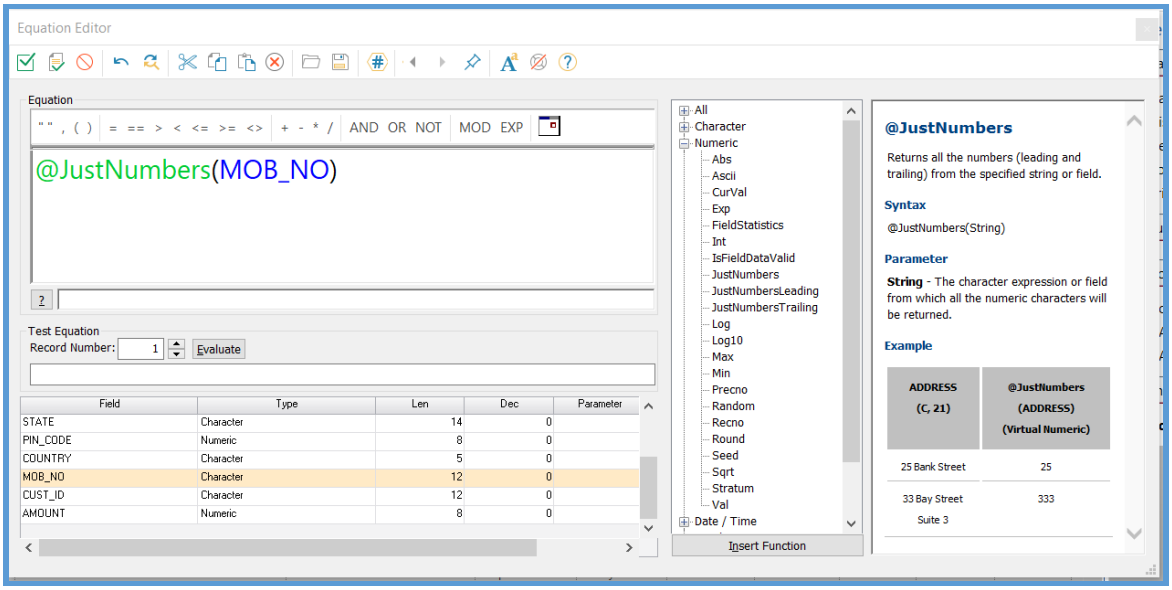
- **DESCRIBE SORTED ORDER:** This feature is available in Data Tab > Order > Describe Sorted Order. This option does not sort or index the database. This option informs IDEA of the inherent order of the database which prevents a sort or index if this order is required.

9.9 FIELDS

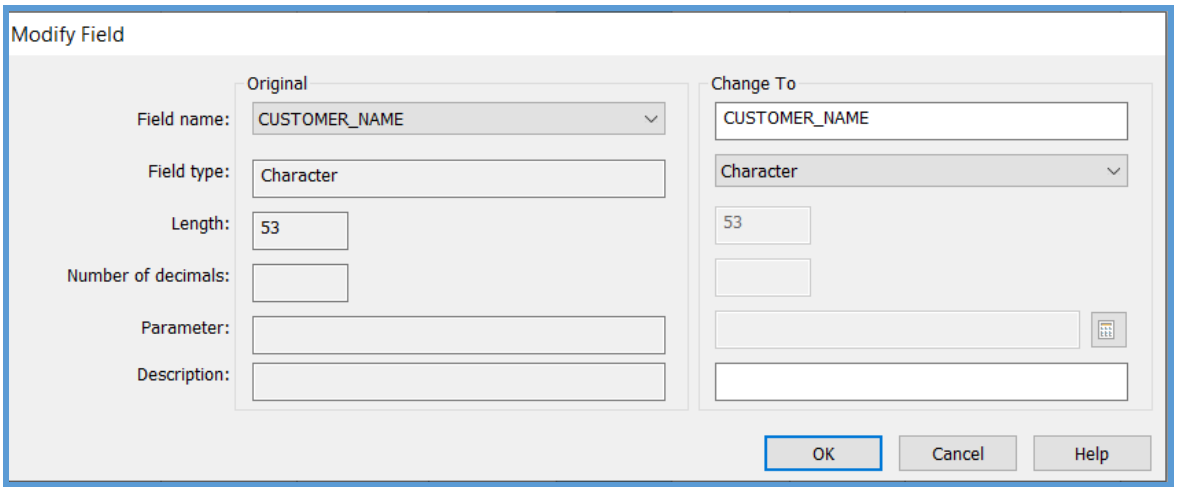
- **Append:** This feature helps to create a new **field (column)** in the current file. For example, we want to append a new column "New Mobile No." without "-" sign in it. In this, we can give the parameter and also write the description, if as required.



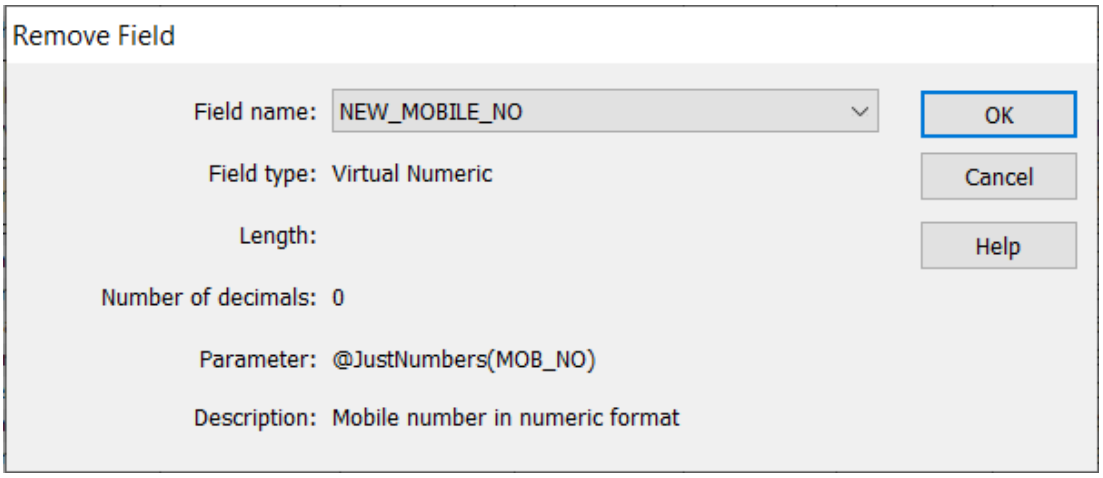
In the parameter, we need to put the equation. This can be done using Equation Editor as highlighted above. Once we click on that it opens Equation Editor window. While writing function, we need to remember that the function begins with @ sign (in excel, we use "=" sign). For function selection, in excel, we use Tab key, whereas in IDEA, we use Enter key. The function description is also shown in right side of the equation editor.



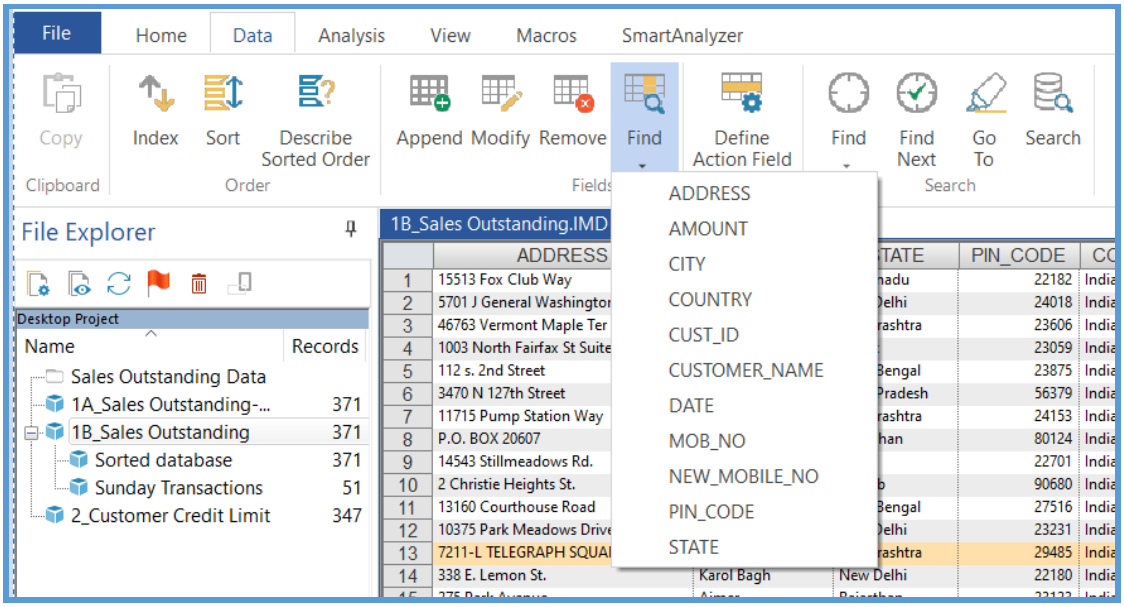
- **Modify:** Using this option, an existing field can be modified (Name of field, Field type, etc.). It is to be noted that if we modify any column, which has been used in creation of new field, then modifying the base column will also affect the new column.



- **Remove:** Using this option, we can remove the field from the database. The availability of field name for removal depends on version of IDEA software used. In version 10.3, it shows only the new field i.e. “New_Mobile_No” added in the database for removal.



- **Find:** This feature helps to highlight a field in the database.

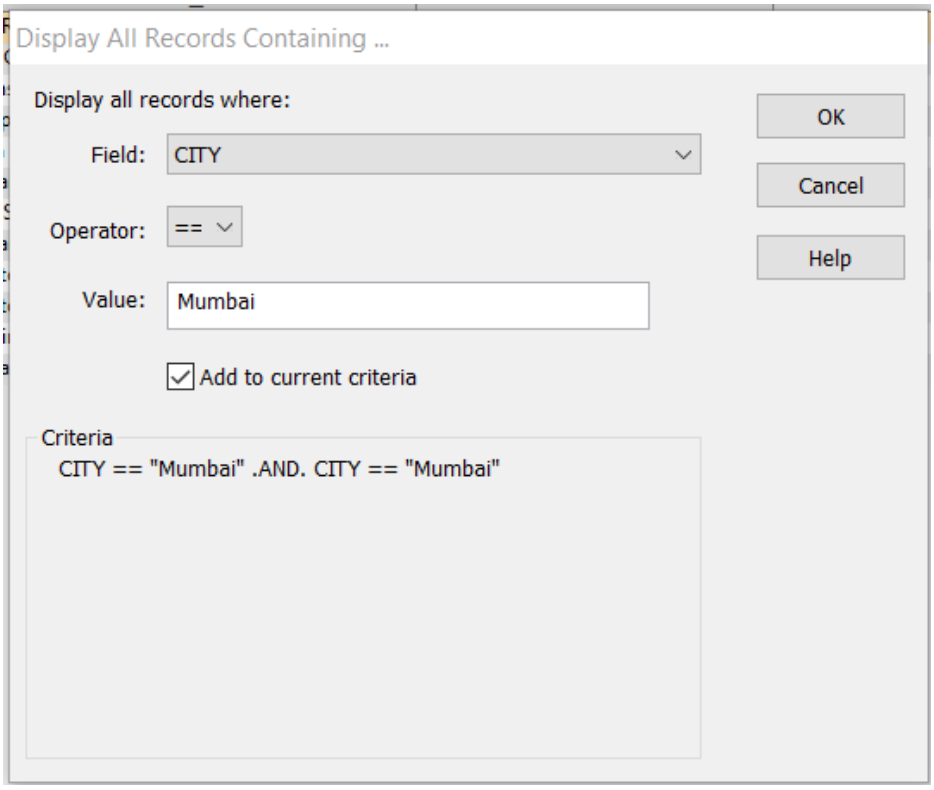


When we click on Find option, it shows the list of Columns in the current database and when we select a particular column, it will highlight that column.

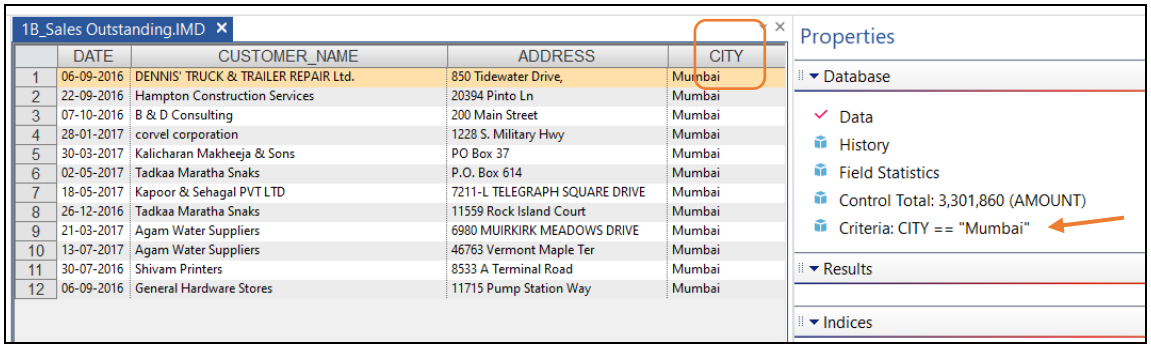
9.10 SEARCH

Under this menu option, there are 4 options available – Find, Find Next, Go To and Search.

- **Find:** This option helps to find a particular value or text. In the Find, 2 options are available i.e. "Find a record" and "Display all records containing the criteria". For example, we want to display all records which contain Mumbai, then we need to select the CITY Field as shown below:

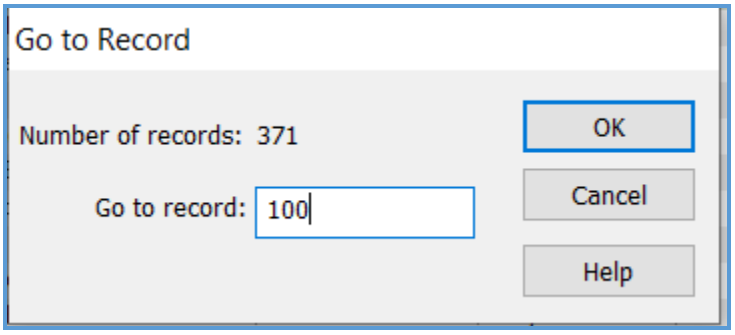


Remember that the value field is case sensitive. The output is displayed as follows:

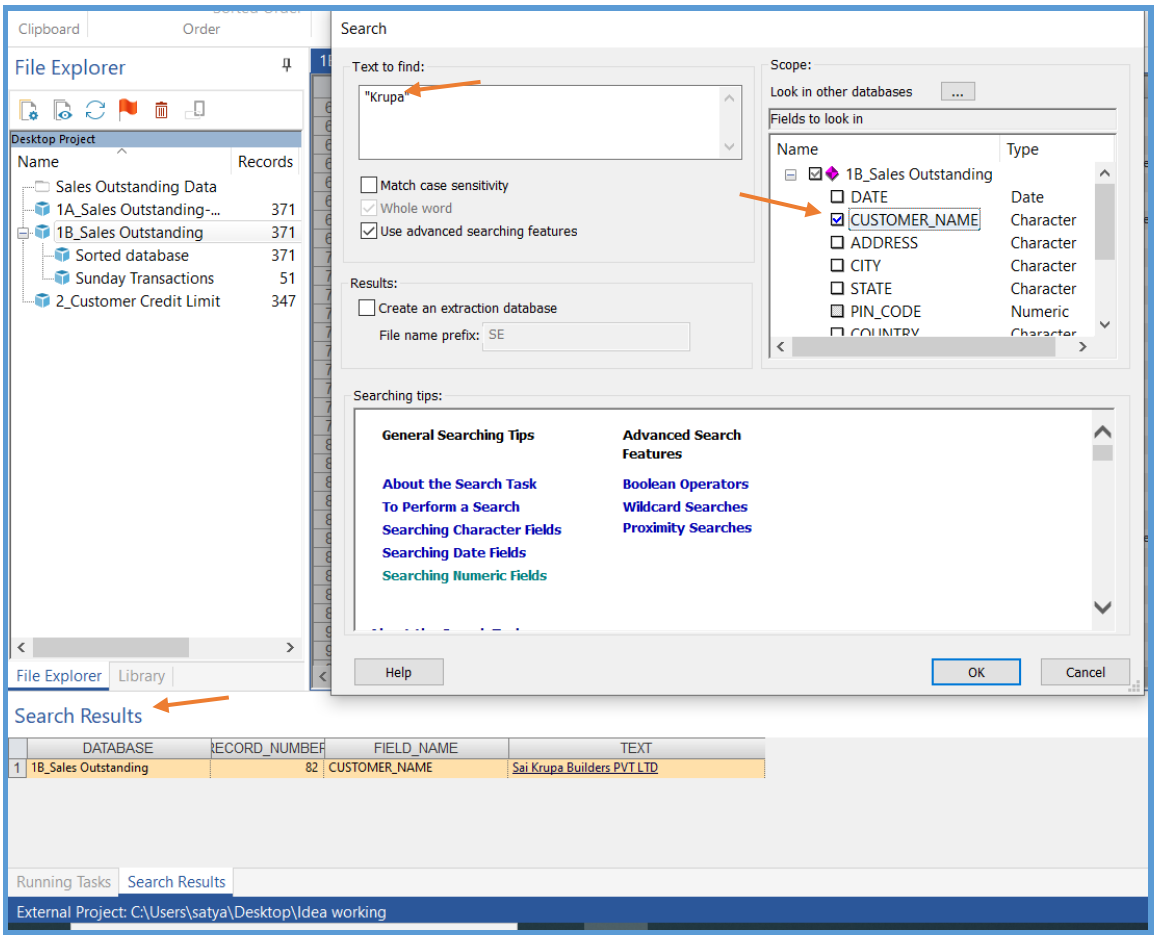


We can right click on criteria in Properties and clear the display to see the original (full) data.

- **Find Next:** This option will help to locate the next occurrence of the specified criteria.
- **Go To:** This option will highlight the specified record. Say we want to go to 100th record, we can type the record number and it will take to that row.



- **Search:** This feature helps to search a value based on criteria. It is not case sensitive. We need to write the criteria and select the columns in which we want to search the value. If found, the list of items is shown below in the screen.



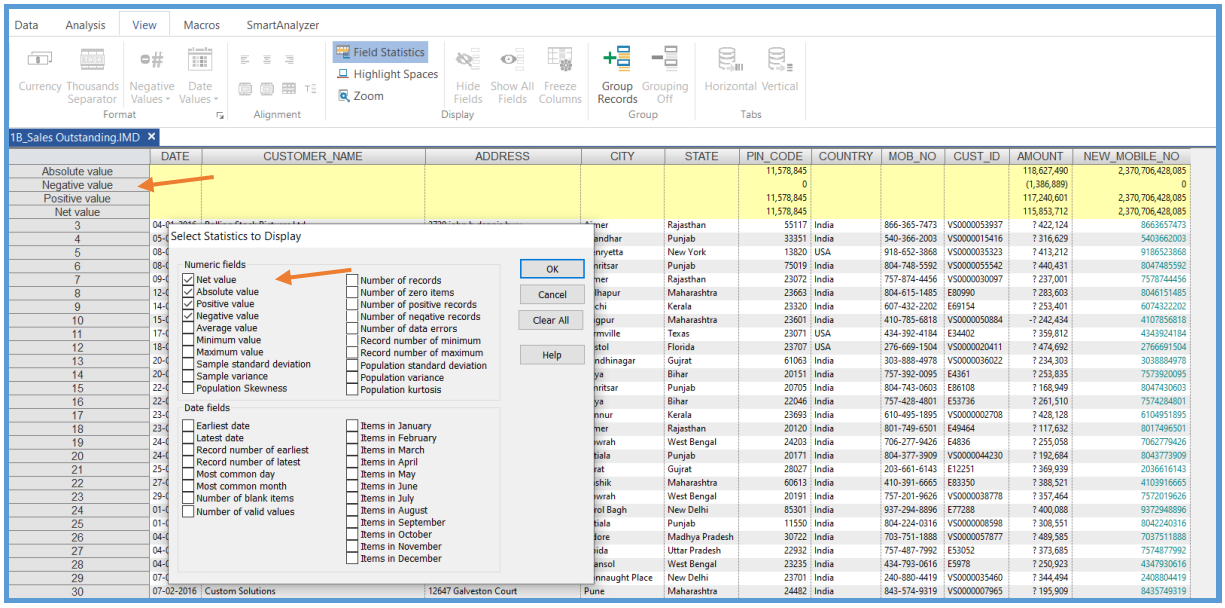
IDEA displays the results of a search in the Search Results fly-out window. Within this window, new searches overwrite old ones. From the right-click menu in the Search Results window, we can re-run the search (as is or make refinements) or print the results. We may sort the results by double-clicking any of the columns within the Search Results window. A directional arrow appears in the field header to indicate

the sort order (ascending or descending). IDEA Server allows multiple users to be searching the same database simultaneously without the risk of overwriting another user's search results.

9.11 VIEW TAB

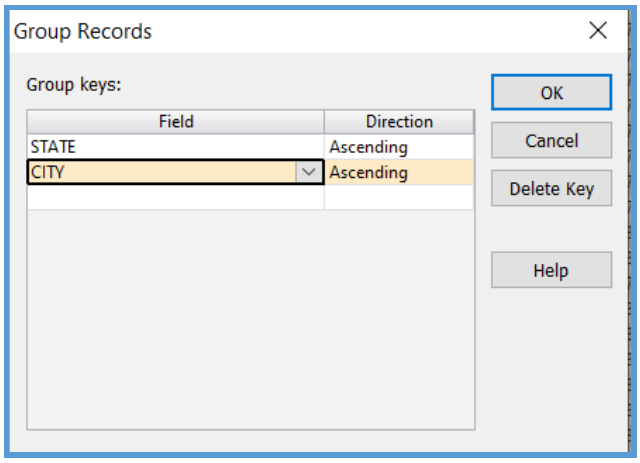
In the View Tab, we have formatting options. The options will be active based on the field selection. For example, date format will be active when we select date field. The list of format options is as follows:

- **Currency Format:** Displays the currency symbol in the selected field.
- **Thousand Separator:** It displays thousand separator in the selected field.
- **Negative Values:** It defines the format for negative values.
- **Date Values:** It defines the date format for date values.
- **Alignment:** Using the alignment option, we can align the column left, right and centre. There is also option to Stack left, Stack Right and Unstack all.
- **Field Statistics:** When we select options in Field Statistics and click OK, then it displays the values over the database:



- **Highlight Spaces:** It display spaces within the data as blocks.
- **Zoom:** It adjust the Zoom factor.
- **Hide Fields:** This option hides the selected fields from the view.
- **Show all Fields:** It displays all fields in the database.
- **Freeze Columns:** It prevents selected column from scrolling off the screen.

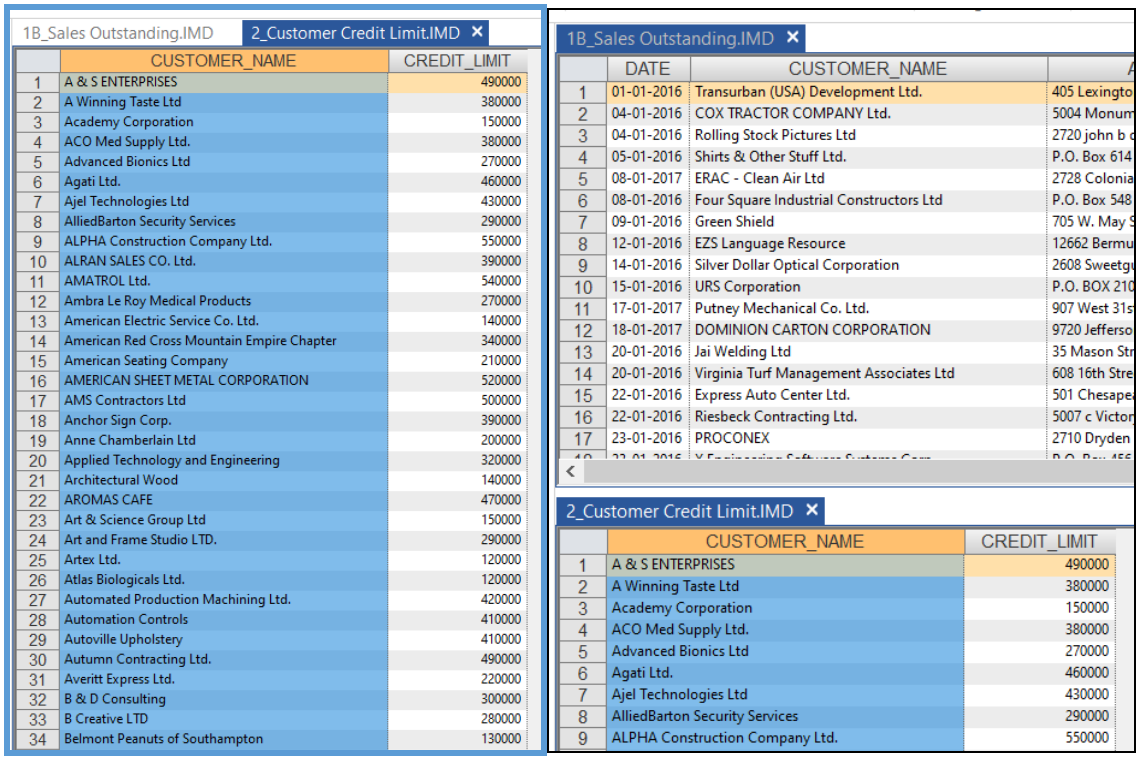
- **Group Records:** It groups the records in the current database by selected keys. In the below example, we have grouped records based on State, then by City. In the output, we can see the first column being the State, followed by City column, The details will be visible, when we click on expand button ("+" button before state / city name).



1B_Sales Outstanding.IMD

	STATE	CITY	DATE	CUSTOI
1	Bihar			
2		Gaya (14 Records)		
3			20-01-2016	Virginia Turf Managemen
4			22-01-2016	Riesbeck Contracting Ltd
5			07-03-2016	Mr Digital Technology
6			15-04-2016	CaLtdom Utility Co. Ltd
7			27-04-2016	Daphne Hawkins
8			28-05-2016	Data Business Systems Lt
9			04-10-2016	LTD Management - HI Ex
10			08-10-2016	C. J. Prettyman Jr. Ltd.
11			22-10-2016	Mobility Products & Serv
12			29-10-2016	Covel Family Services
13			01-11-2016	Newmans Contracting
14			23-11-2016	Willis Entreprises Ltd
15			29-11-2016	McGees Painting And Pre
16			10-04-2017	BRUSH IT ON PAINT CO.
17		Patna (10 Records)		
18	California			
19		BELTSVILLE (1 Records)		
20		Grand Rapids (1 Records)		
21		Hampton (1 Records)		
22		Oak Ridge (1 Records)		
23		ROANOKE (1 Records)		
24		Roanoke (1 Records)		
25		STANTON (1 Records)		
26	Florida			
27	Gujrat			

- **Grouping Off:** It turns grouping mode off.
- **Horizontal Tabs:** This helps to view the databases in horizontal format. When we have more than 2 databases open, we can view them simultaneously using Horizontal option in Tabs. 1st image where, we have 2 databases open. In second image we have both databases in horizontal format.



1B_Sales Outstanding.IMD | 2_Customer Credit Limit.IMD

	CUSTOMER_NAME	CREDIT_LIMIT
1	A & S ENTERPRISES	490000
2	A Winning Taste Ltd	380000
3	Academy Corporation	150000
4	ACO Med Supply Ltd.	380000
5	Advanced Bionics Ltd	270000
6	Agati Ltd.	460000
7	Ajel Technologies Ltd	430000
8	AlliedBarton Security Services	290000
9	ALPHA Construction Company Ltd.	550000
10	ALRAN SALES CO. Ltd.	390000
11	AMATROL Ltd.	540000
12	Ambra Le Roy Medical Products	270000
13	American Electric Service Co. Ltd.	140000
14	American Red Cross Mountain Empire Chapter	340000
15	American Seating Company	210000
16	AMERICAN SHEET METAL CORPORATION	520000
17	AMS Contractors Ltd	500000
18	Anchor Sign Corp.	390000
19	Anne Chamberlain Ltd	200000
20	Applied Technology and Engineering	320000
21	Architectural Wood	140000
22	AROMAS CAFE	470000
23	Art & Science Group Ltd	150000
24	Art and Frame Studio LTD.	290000
25	Artex Ltd.	120000
26	Atlas Biologicals Ltd.	120000
27	Automated Production Machining Ltd.	420000
28	Automation Controls	410000
29	Autoville Upholstery	410000
30	Autumn Contracting Ltd.	490000
31	Averitt Express Ltd.	220000
32	B & D Consulting	300000
33	B Creative LTD	280000
34	Belmont Peanuts of Southampton	130000

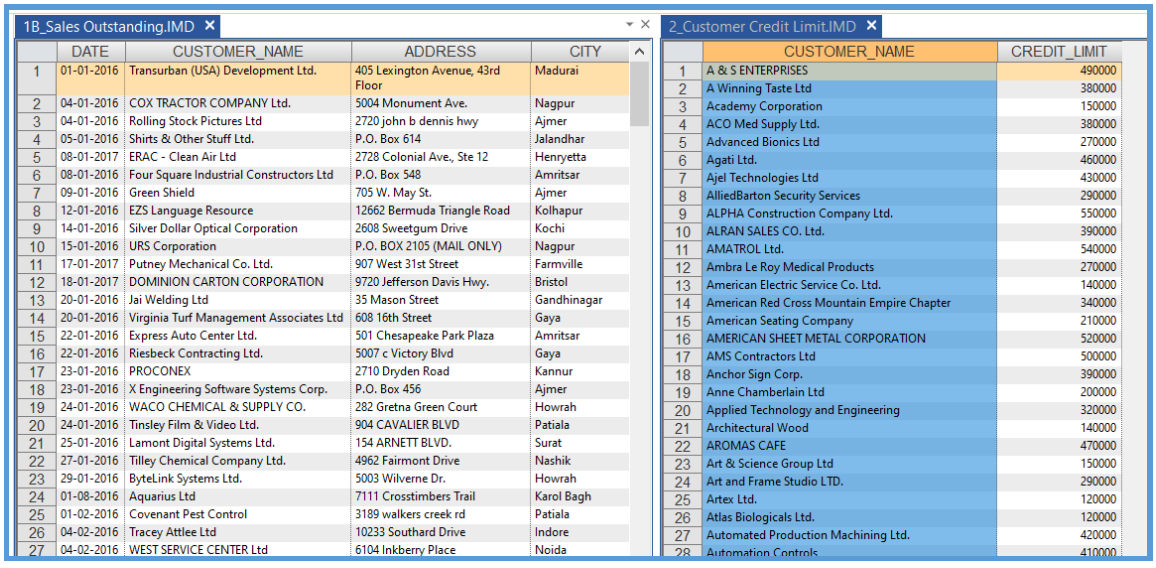
1B_Sales Outstanding.IMD

	DATE	CUSTOMER_NAME	A
1	01-01-2016	Transurban (USA) Development Ltd.	405 Lexingto
2	04-01-2016	COX TRACTOR COMPANY Ltd.	5004 Monum
3	04-01-2016	Rolling Stock Pictures Ltd	2720 john b
4	05-01-2016	Shirts & Other Stuff Ltd.	P.O. Box 614
5	08-01-2017	ERAC - Clean Air Ltd	2728 Colonia
6	08-01-2016	Four Square Industrial Constructors Ltd	P.O. Box 548
7	09-01-2016	Green Shield	705 W. May S
8	12-01-2016	EZS Language Resource	12662 Bermu
9	14-01-2016	Silver Dollar Optical Corporation	2608 Sweetg
10	15-01-2016	URS Corporation	P.O. BOX 210
11	17-01-2017	Putney Mechanical Co. Ltd.	907 West 31s
12	18-01-2017	DOMINION CARTON CORPORATION	9720 Jefferso
13	20-01-2016	Jai Welding Ltd	35 Mason Str
14	20-01-2016	Virginia Turf Management Associates Ltd	608 16th Stre
15	22-01-2016	Express Auto Center Ltd.	501 Chesape
16	22-01-2016	Riesbeck Contracting Ltd.	5007 c Victor
17	23-01-2016	PROCONEX	2710 Dryden

2_Customer Credit Limit.IMD

	CUSTOMER_NAME	CREDIT_LIMIT
1	A & S ENTERPRISES	490000
2	A Winning Taste Ltd	380000
3	Academy Corporation	150000
4	ACO Med Supply Ltd.	380000
5	Advanced Bionics Ltd	270000
6	Agati Ltd.	460000
7	Ajel Technologies Ltd	430000
8	AlliedBarton Security Services	290000
9	ALPHA Construction Company Ltd.	550000

- **Vertical Tabs:** It organises the open databases by moving the active database into its own vertical tab.



1B_Sales Outstanding.IMD

	DATE	CUSTOMER_NAME	ADDRESS	CITY
1	01-01-2016	Transurban (USA) Development Ltd.	405 Lexington Avenue, 43rd Floor	Madurai
2	04-01-2016	COX TRACTOR COMPANY Ltd.	5004 Monument Ave.	Nagpur
3	04-01-2016	Rolling Stock Pictures Ltd	2720 john b dennis hwy	Ajmer
4	05-01-2016	Shirts & Other Stuff Ltd.	P.O. Box 614	Jalandhar
5	08-01-2017	ERAC - Clean Air Ltd	2728 Colonial Ave., Ste 12	Henryetta
6	08-01-2016	Four Square Industrial Constructors Ltd	P.O. Box 548	Amritsar
7	09-01-2016	Green Shield	705 W. May St.	Ajmer
8	12-01-2016	EZS Language Resource	12662 Bermuda Triangle Road	Kolhapur
9	14-01-2016	Silver Dollar Optical Corporation	2608 Sweetgum Drive	Kochi
10	15-01-2016	URS Corporation	P.O. BOX 2105 (MAIL ONLY)	Nagpur
11	17-01-2017	Putney Mechanical Co. Ltd.	907 West 31st Street	Farmville
12	18-01-2017	DOMINION CARTON CORPORATION	9720 Jefferson Davis Hwy.	Bristol
13	20-01-2016	Jai Welding Ltd	35 Mason Street	Gandhinagar
14	20-01-2016	Virginia Turf Management Associates Ltd	608 16th Street	Gaya
15	22-01-2016	Express Auto Center Ltd.	501 Chesapeake Park Plaza	Amritsar
16	22-01-2016	Riesbeck Contracting Ltd.	5007 c Victory Blvd	Gaya
17	23-01-2016	PROCONEX	2710 Dryden Road	Kannur
18	23-01-2016	X Engineering Software Systems Corp.	P.O. Box 456	Ajmer
19	24-01-2016	WACO CHEMICAL & SUPPLY CO.	282 Gretna Green Court	Howrah
20	24-01-2016	Tinsley Film & Video Ltd.	904 CAVALIER BLVD	Patiala
21	25-01-2016	Lamont Digital Systems Ltd.	154 ARNETT BLVD.	Surat
22	27-01-2016	Tilley Chemical Company Ltd.	4962 Fairmont Drive	Nashik
23	29-01-2016	ByteLink Systems Ltd.	5003 Wilverne Dr.	Howrah
24	01-08-2016	Aquarius Ltd	7111 Crosstimbers Trail	Karol Bagh
25	01-02-2016	Covenant Pest Control	3189 walkers creek rd	Patiala
26	04-02-2016	Tracey Attlee Ltd	10233 Southard Drive	Indore
27	04-02-2016	WEST SERVICE CENTER Ltd	6104 Inkberry Place	Noida

2_Customer Credit Limit.IMD

	CUSTOMER_NAME	CREDIT_LIMIT
1	A & S ENTERPRISES	490000
2	A Winning Taste Ltd	380000
3	Academy Corporation	150000
4	ACO Med Supply Ltd.	380000
5	Advanced Bionics Ltd	270000
6	Agati Ltd.	460000
7	Ajel Technologies Ltd	430000
8	AlliedBarton Security Services	290000
9	ALPHA Construction Company Ltd.	550000
10	ALRAN SALES CO. Ltd.	390000
11	AMATROL Ltd.	540000
12	Ambra Le Roy Medical Products	270000
13	American Electric Service Co. Ltd.	140000
14	American Red Cross Mountain Empire Chapter	340000
15	American Seating Company	210000
16	AMERICAN SHEET METAL CORPORATION	520000
17	AMS Contractors Ltd	500000
18	Anchor Sign Corp.	390000
19	Anne Chamberlain Ltd	200000
20	Applied Technology and Engineering	320000
21	Architectural Wood	140000
22	AROMAS CAFE	470000
23	Art & Science Group Ltd	150000
24	Art and Frame Studio LTD.	290000
25	Artex Ltd.	120000
26	Atlas Biologicals Ltd.	120000
27	Automated Production Machining Ltd.	420000
28	Automation Controls	410000

9.12 SUMMARY

Computer Assisted Audit Techniques (CAATs) refers to using technology for increasing the effectiveness and efficiency of auditing. CAATs enable auditors to do more with less and add value through the assurance process which is more robust and comprehensive. CAATs are tools for drawing inferences and gathering relevant and reliable evidence as per requirements of the assignment. CAATs provide direct access to electronic information and empower auditors not only to perform their existing audits more efficiently and effectively but also facilitate them in knowing how to create and execute new type of IT related audit assignments. CAATs may produce a large proportion of the audit evidence developed on audits and, as a result, the auditor should carefully plan for and exhibit due professional care in the use of CAATs.

In IDEA, we create project like a folder and whatever we perform in IDEA, all those works will be stored in that project. We can import various type of files in IDEA using Import Assistant, mainly Microsoft Excel, Microsoft Access, Print report and Adobe PDF, Text, XML, dBASE, etc. Using File Explorer, we can create folders and move the files to respective folders. Once we open any file, the file properties option gets activated. In this, we get following options: Data, History, Field Statistics, Control Total and Criteria. Data is used to view the data in the file in a spreadsheet-like format with field names as field headings and record numbers as row numbers. History option records all the steps performed in that particular file since inception. All these steps can not be undone/modified/deleted. This History works as an audit evidence and accepted in Court as well. Field Statistics option shows statistics for each field in the particular file.

We can export files in various format including Excel, Access and Text format. Project Overview option records all the files created or deleted form the folder. In IDEA, we can do formatting and sorting of the data as we do in Excel. New fields can be added or deleted in/from the database using the option Append, Modify and Remove. We can also arrange various databases in horizontal and vertical form.

REFERENCE

[1] https://idea.caseware.com

CHAPTER

10 DATA ANALYSIS IN CAATs

LEARNING OBJECTIVES

- Understanding extraction of data
- Understanding Gap and Duplicates in data
- Understanding Categorisation of data
- Understanding setting relationship between databases
- Understanding data visualization
- Understanding various function in IDEA

In last Chapter, we imported 3 files in IDEA. We added 2 more files in IDEA. In this Chapter, out main objective is to analyse all these files using various options in IDEA.

- File 1 – 1B_Sales Outstanding – Party-wise amount outstanding
- File 2 – Customer Credit Limit – Credit Limit assigned to each party
- File 3 – Sales database for 2 years sales
- File 4 – Sales data for Benford's law

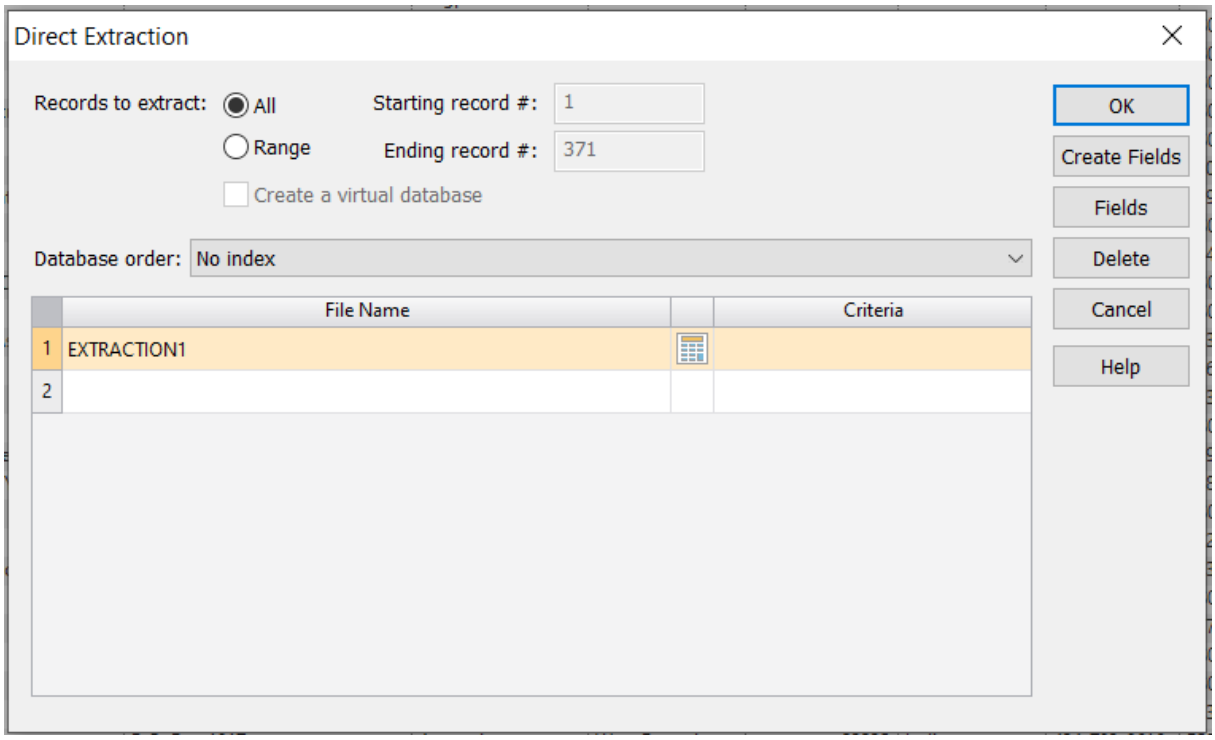
10.1 EXTRACTION OF DATA

The Extraction task is used to copy records that meet the specified criteria from an existing database and place them in another database. This task makes it possible to work on a subset of the data and speed the analysis process.

The extraction can be done in 4 different ways:

i) **Direct Extraction:** This option is available in Analysis Tab > Extract > Direct. It helps to create a new database containing the records that meet the criteria from the current database.

Once we click on Direct Extraction option, a dialoguebox opens as below:



Records to extract: We need to define the range of records within which we want to do extraction. It can be done in full data range or in a specified range by defining Starting record and Ending record.

File Name: We can define the File Name as well. The new extracted data will be given that file name as defined here.

Equation Editor / Criteria option: Also using Equation Editor / Criteria option, we can define the rule for data extraction.

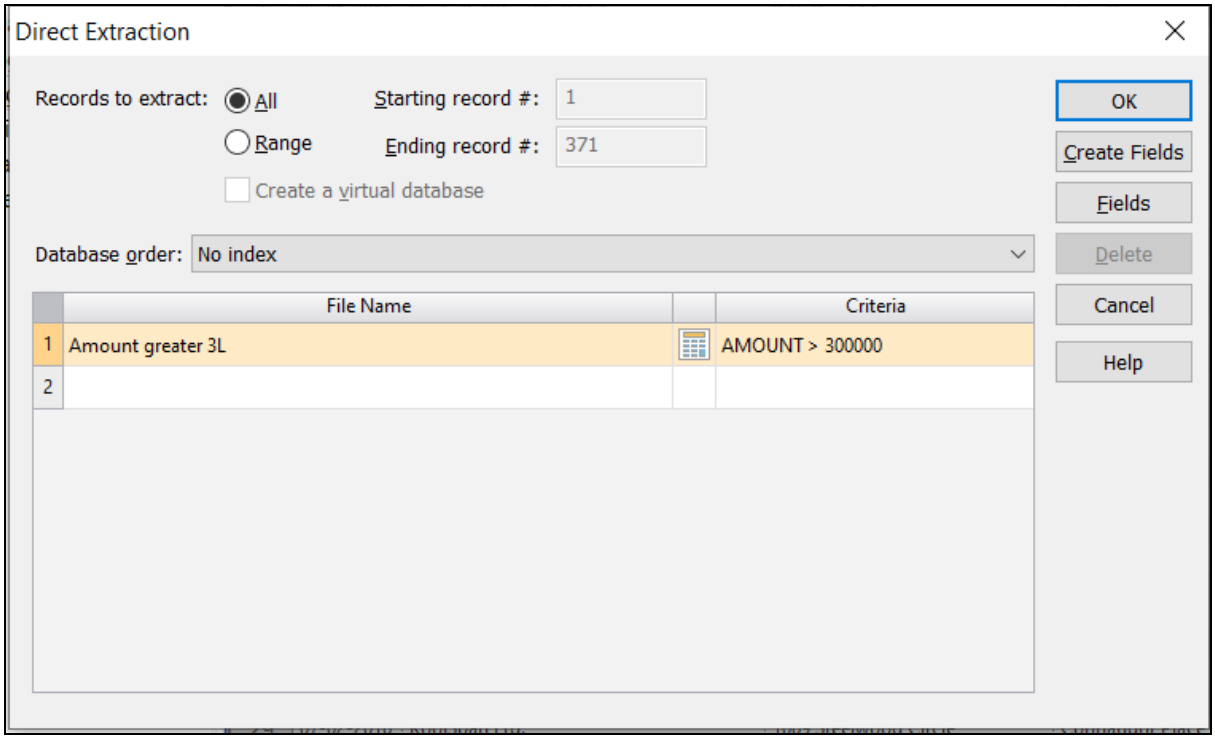
Create Fields: This help to add a new field in the output file.

Fields: This helps to define the list of fields/columns in the output file. By default, it includes all fields. In case, we don't want specific file, we can deselect that column.

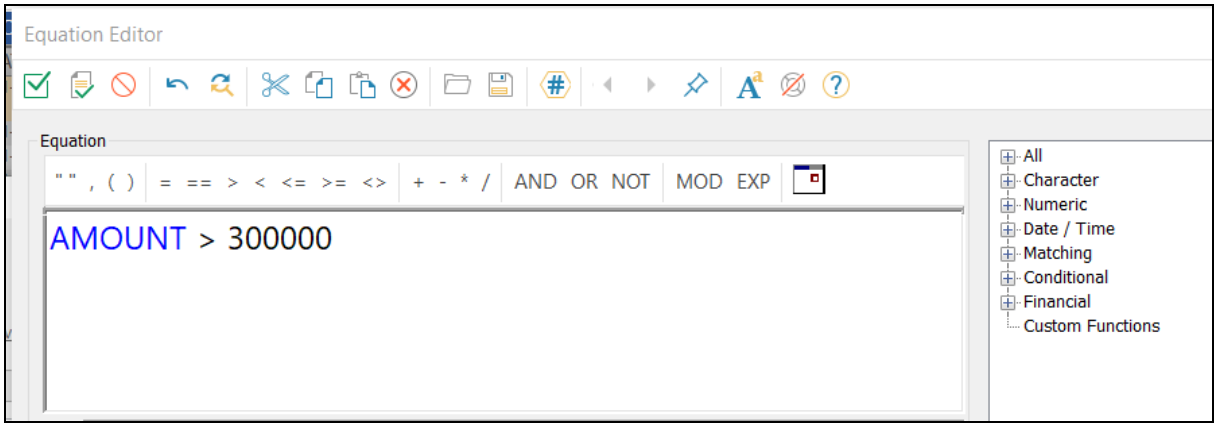
Delete: In case, we want to delete any Extraction rule, then we can use this option.

In this example, in entire outstanding database, we want to extract all the parties, whose outstanding amount is greater than 300000. The file name will be "Amount greater 3L" as shown below.

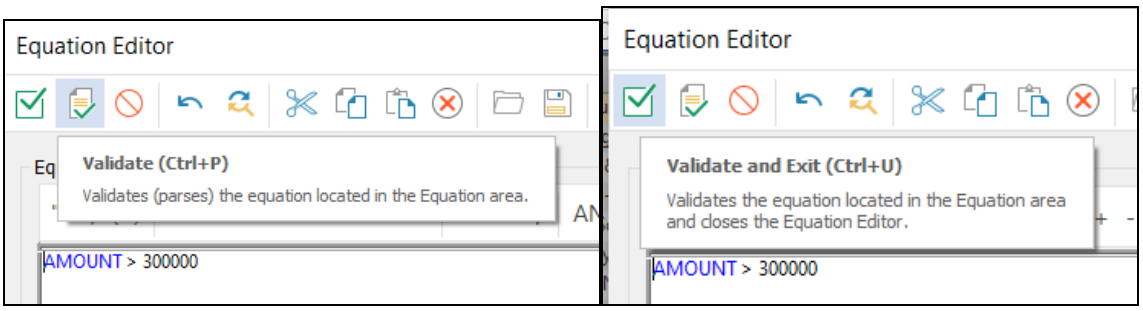
Note: It is to be noted that it does not allow use of symbols in File name.



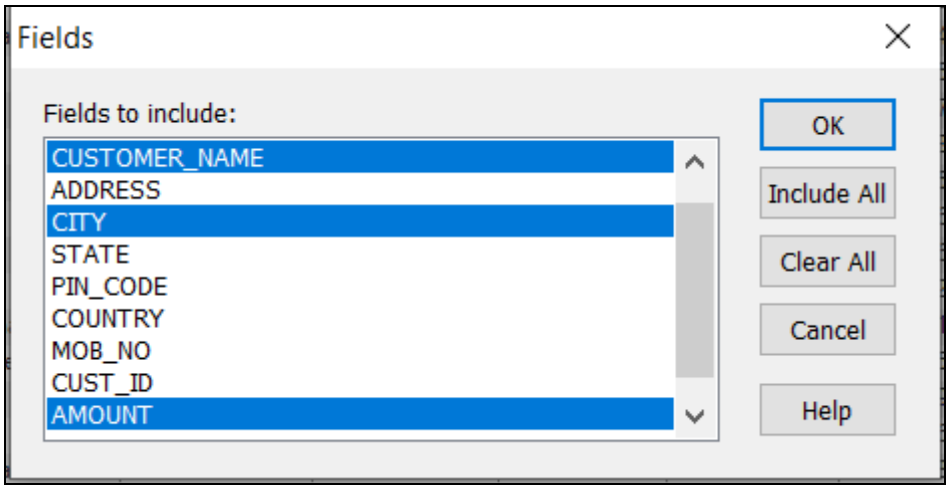
Using the Equation Editor, we have given the equation as "Amount > 300000" as shown below:



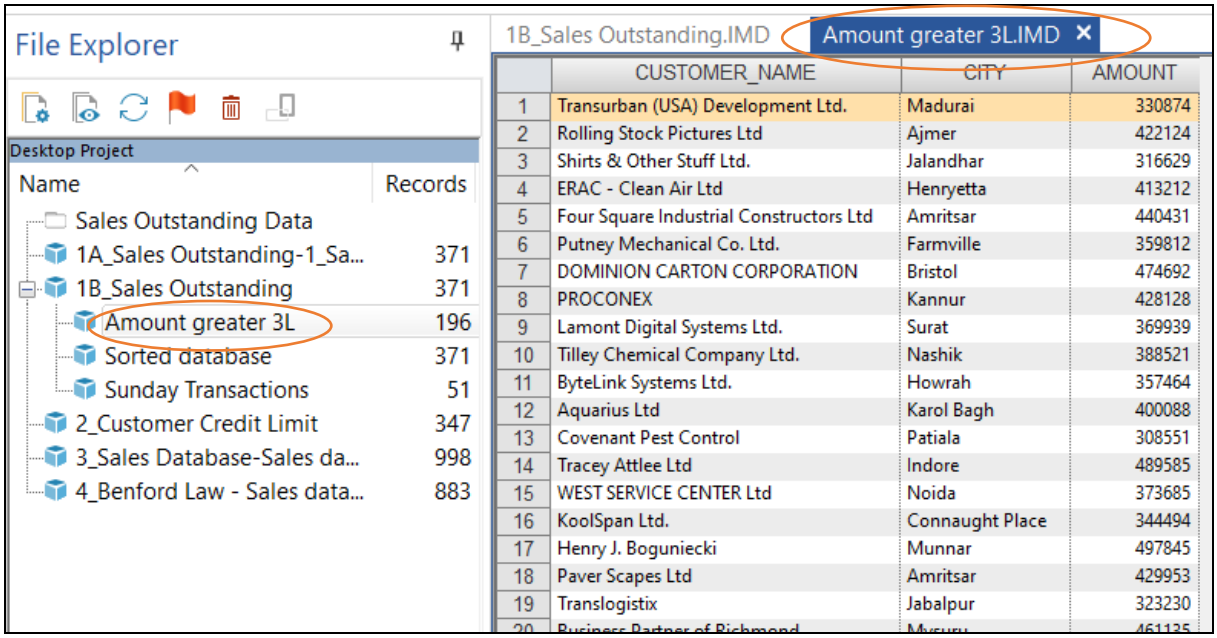
After entering the equation, we should validate that equation in Equation Editor.



We need to define the fields, which we need in Output file. Here, we have selected only 3 columns as highlighted below:

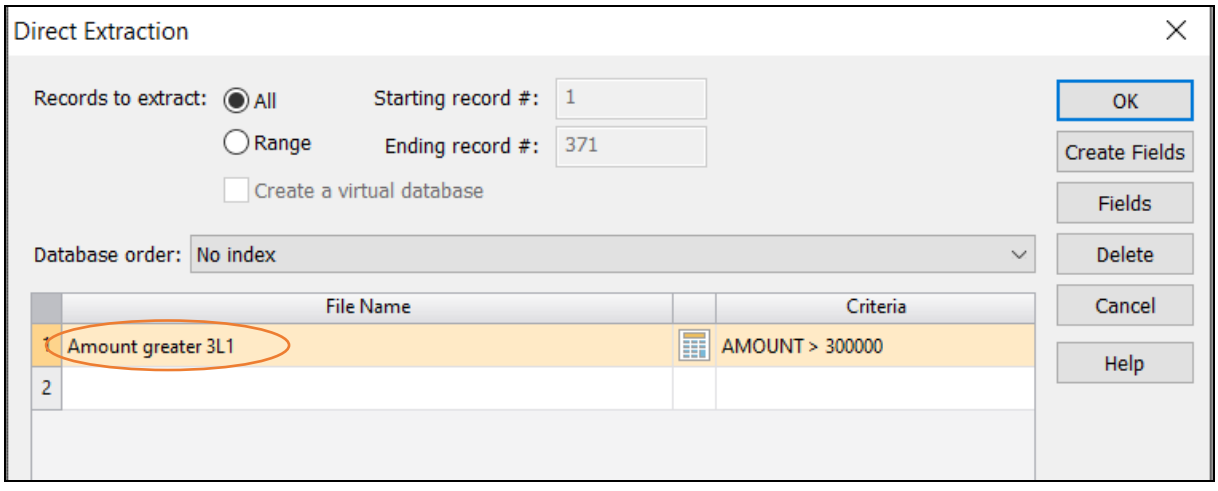


When we click ok for Direct Extraction, a new tab is created with the file name given by us and also a new child is created below 1B_Sales Outstanding file. It is to be noted that the output file has only 3 columns as defined by us.

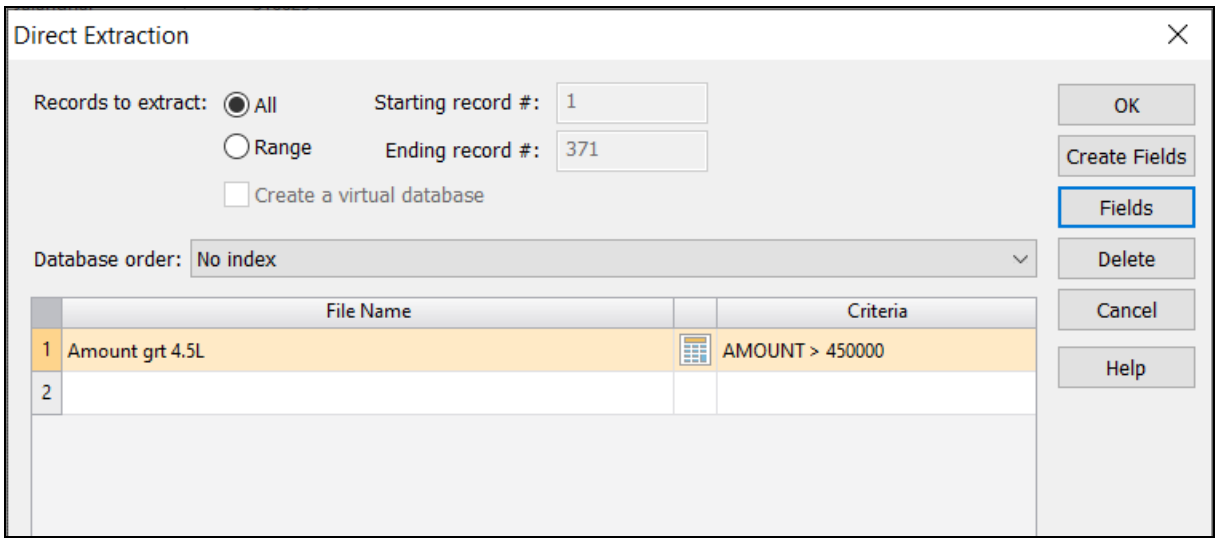


Re-run option

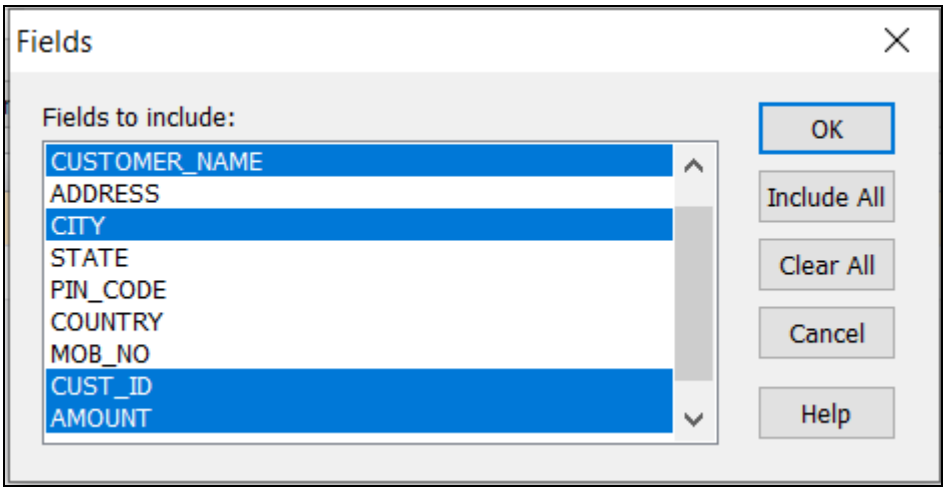
In case, we want to Re-run the above query for all amount greater than 450000, we need to click on Re-run option in Analysis Tab. When we click on Re-run option, it opens the same query with number 1 added at the end of File Name.



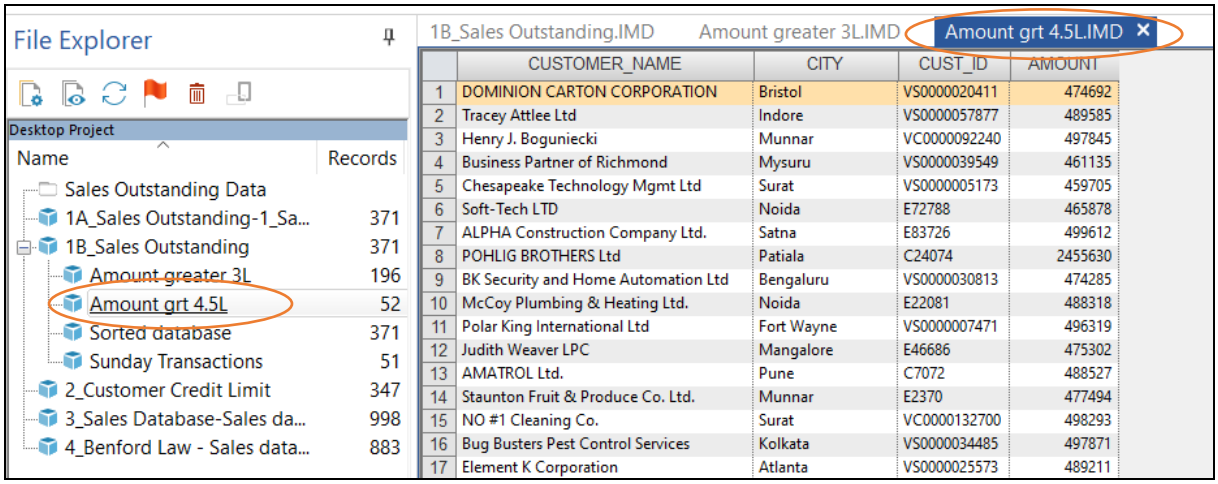
We can rename the file as “Amount grt 4.5L”. We can either change the equation in Equation Edition or directly in Criteria as shown below.



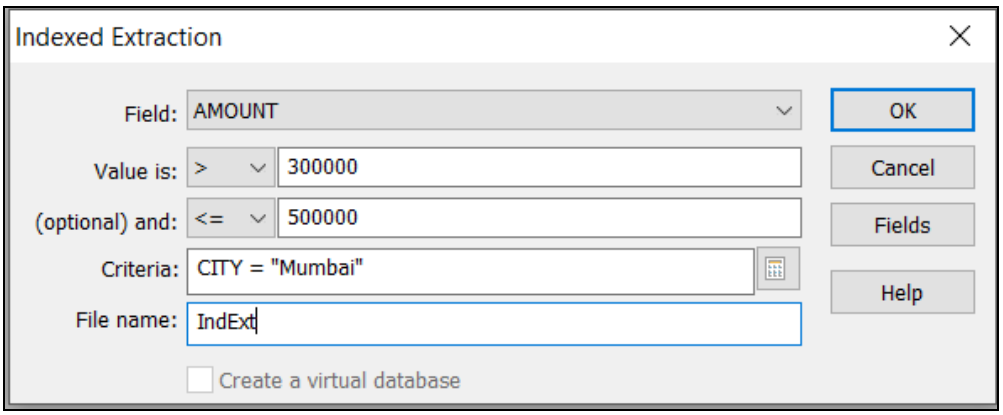
While doing Re-run, we can also redefine the number of columns we want in output file. Here, we have selected "CUST_ID" column in addition to 3 columns selected previously.



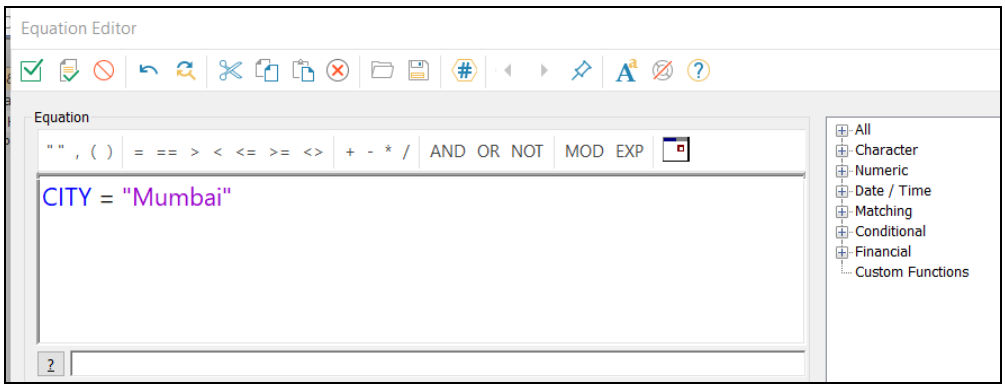
The output file is created and a child is created in "1B_Sales Outstanding" as shown below:



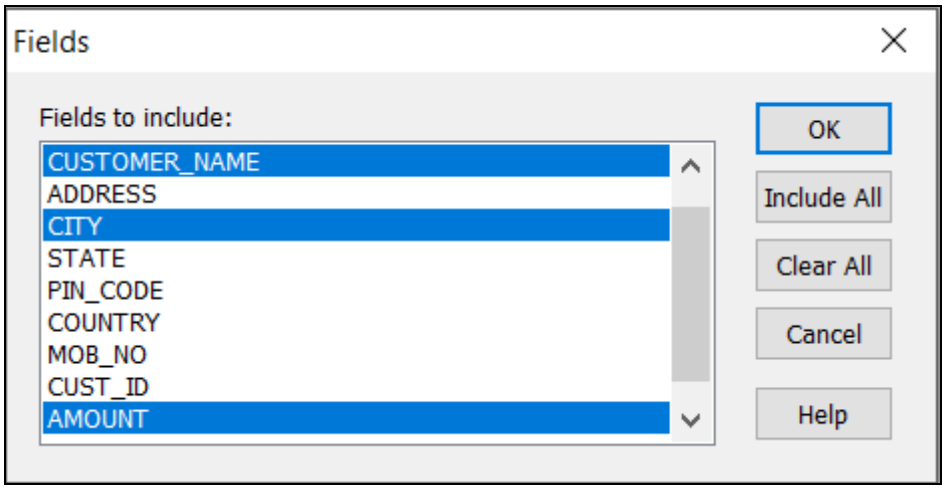
ii) **Indexed Extraction:** The Indexed Extraction task is used to create a new database that contains records that meet the criteria we specify. In the below example, in Amount field, we want all value > 300000 and <= 500000 as shown below. Here, the default file name is "IndExt". We can change it. Currently we are going ahead with default name.



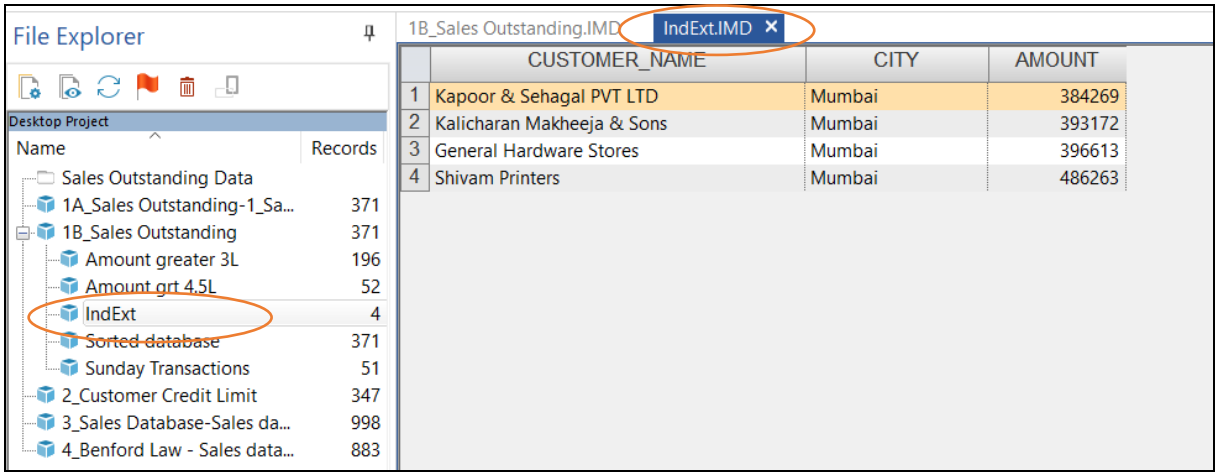
We also want the same for Mumbai city. Hence, in criteria option, we can directly enter the equation as "CITY = "Mumbai". We can also use Equation Editor also for equation writing. It is to be remembered that IDEA is case sensitive.



For the output file, we can select the number of columns we want in output. Here, we want only 3 columns i.e. Customer Name, City and Amount.



Once we click OK, the new file and a child is created as shown below.

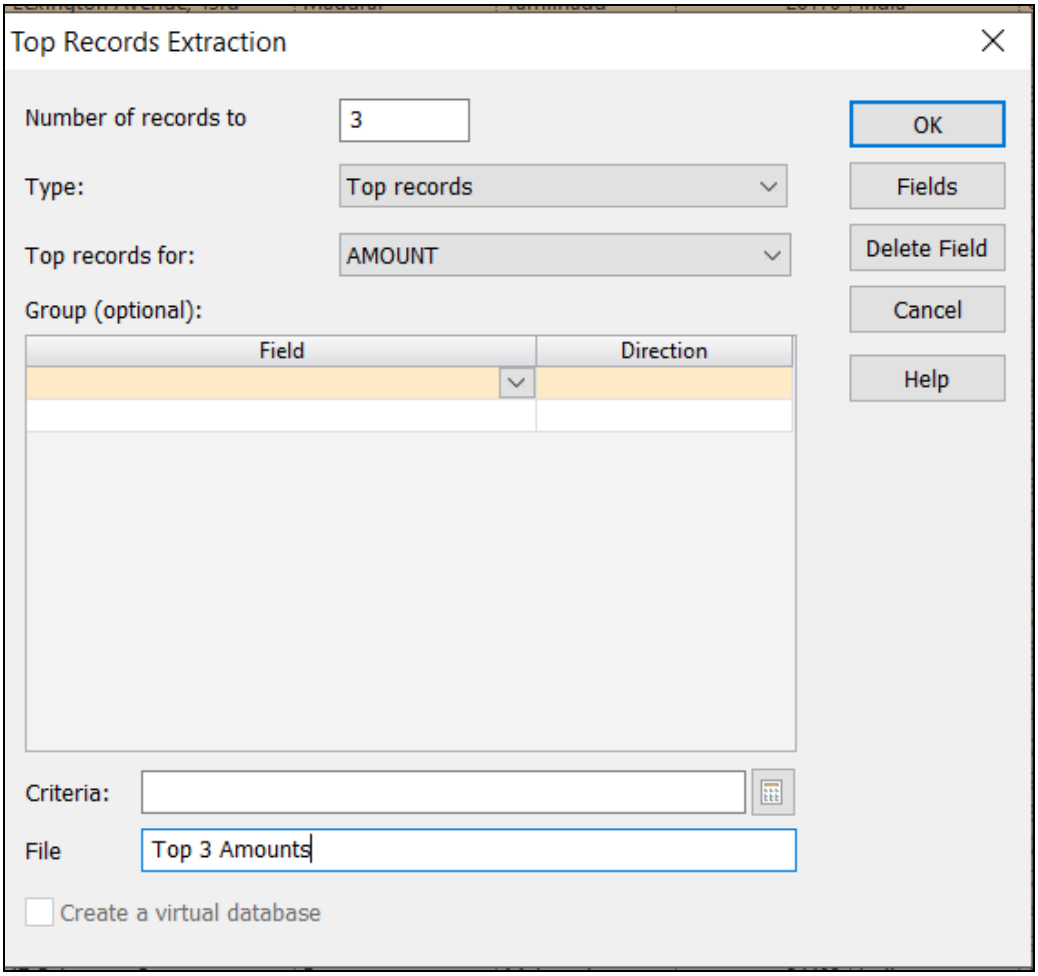


In case, we want to modify the output or condition, we can simply click on Re-run option and redefine the things and another file will be created.

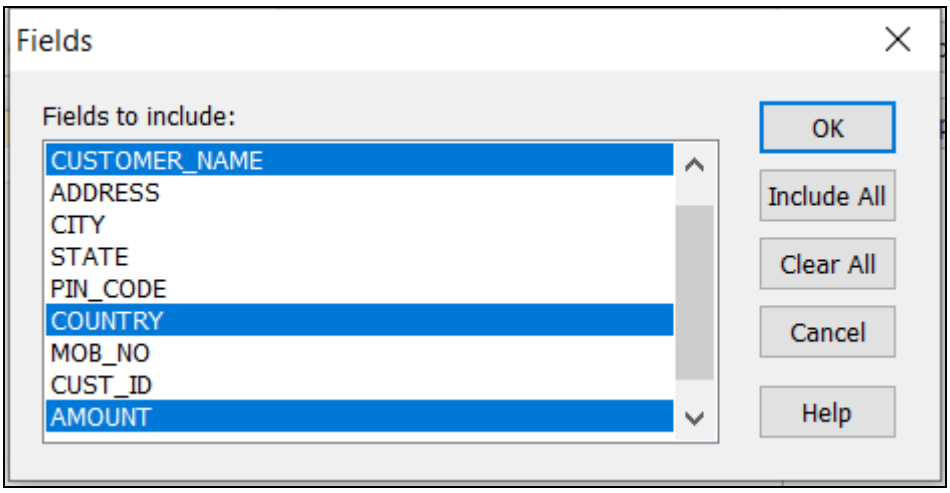
iii) **Top Records Extraction:** We use this option to extract the specified number of records within a unique key (or key combination) to another database. In many cases, we use this kind of extraction to view a sample of the database for testing or other purposes.

Example 1: In this example, we want "Top 3 Amounts" from entire database.

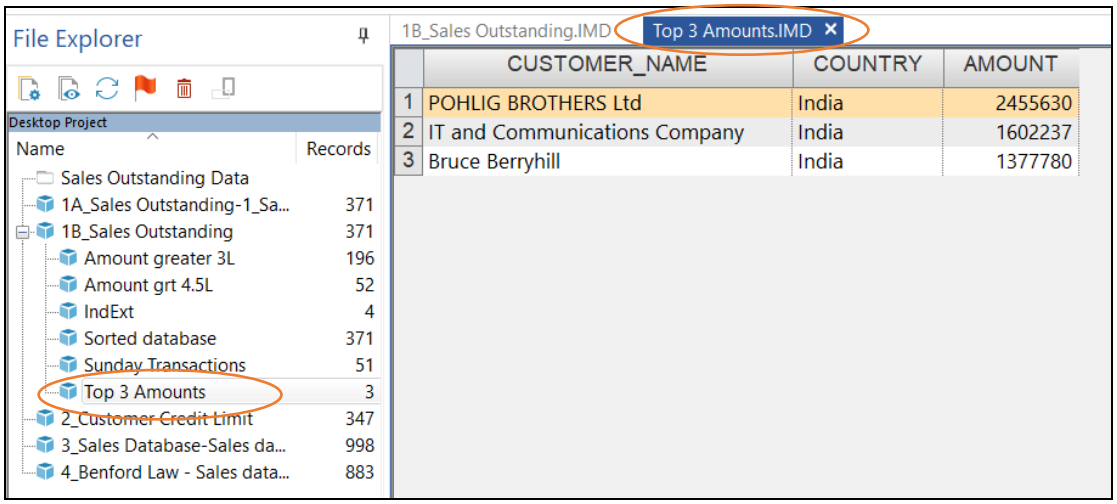
- **Type :** Top records (we also have choice for Bottom records)
- **Top records for :** Amount (we can select other numerical field as well depending on requirement)
- **File :** Top 3 Amounts



In Fields dialoguebox, we have selected only 3 columns for output file i.e. Customer Name, Country and Amount.

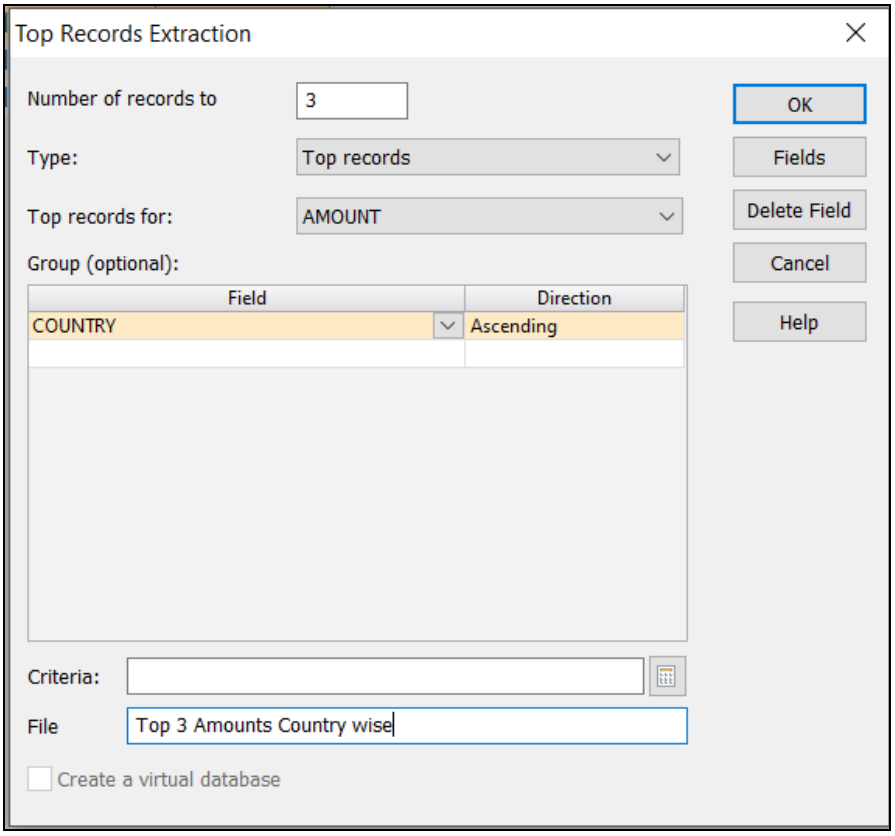


The output file is created as shown below:

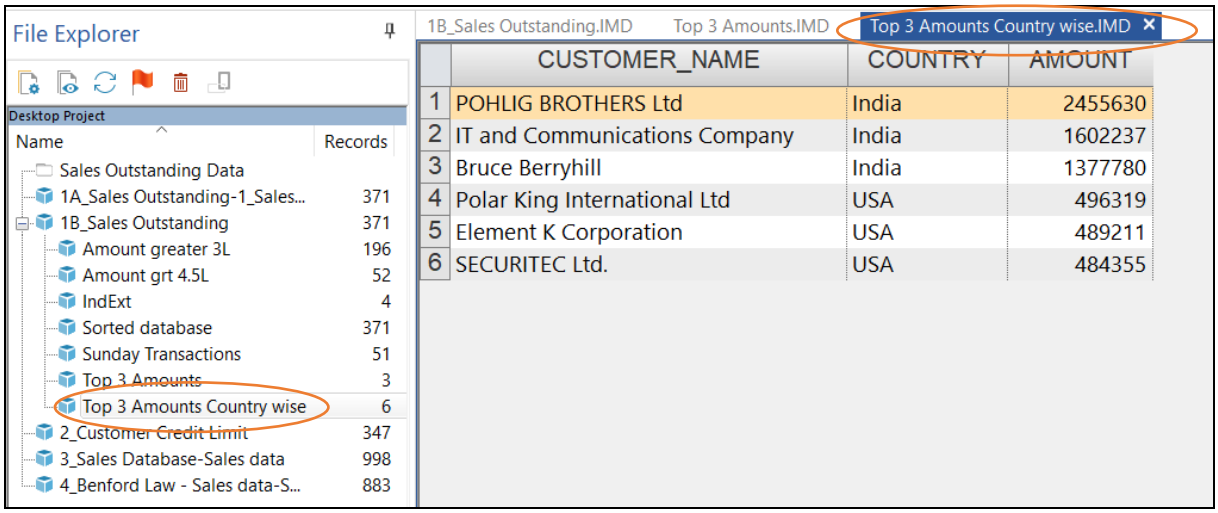


It can be seen that the output file contains only 3 line items based on amount.

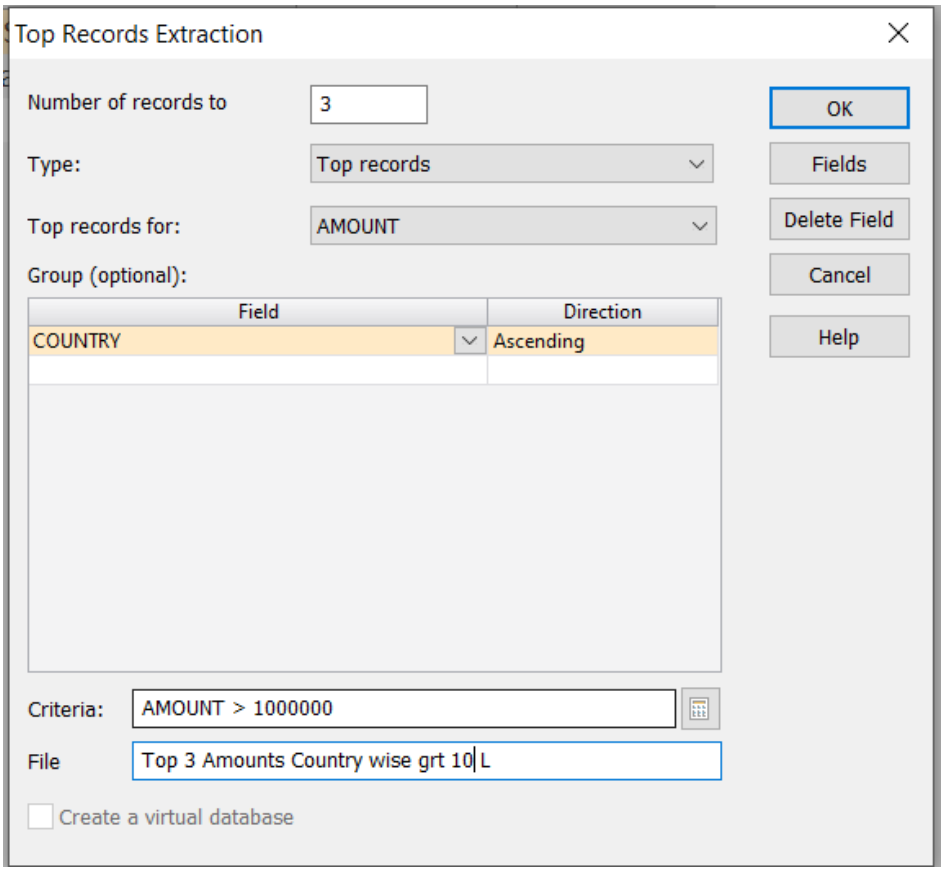
Example 2: Now we want Top 3 Amounts Country wise. Hence, we click on Re-run and change the file name as "Top 3 Amounts Country wise". In the Group (optional), we select field "Country" as shown below. Rest all options remained unchanged.



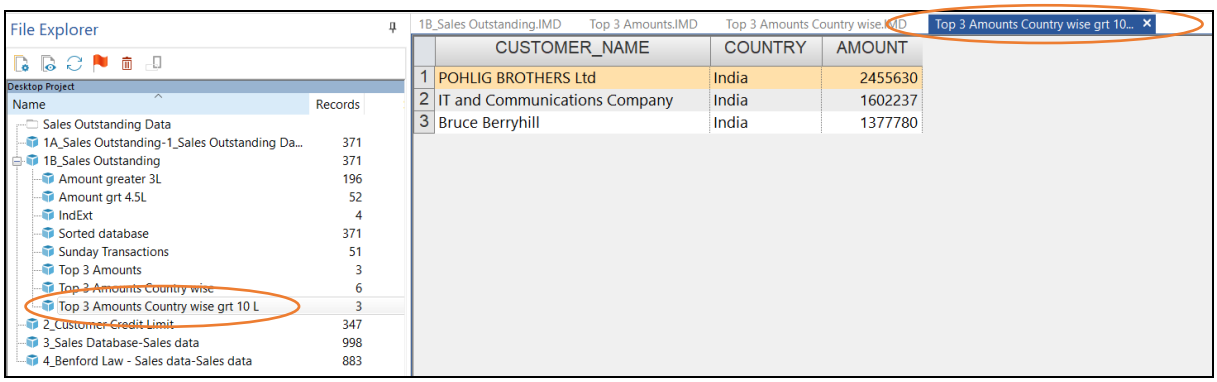
When we click OK, a new Output file is created as shown below. Now we have Top 3 Customers in each Country.



Example 3: Now we want Output based on criteria i.e. Amount > 1000000.We click on Re-run and add the criteria as shown below and modified the file name.

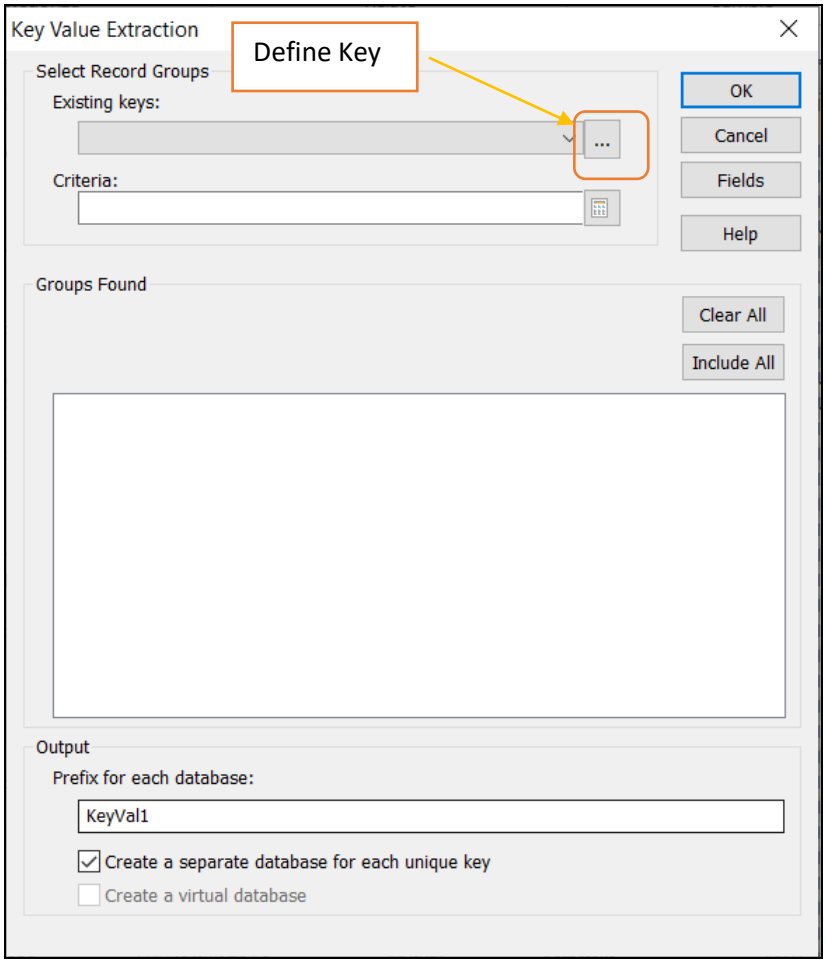


The new output file is created as shown below.

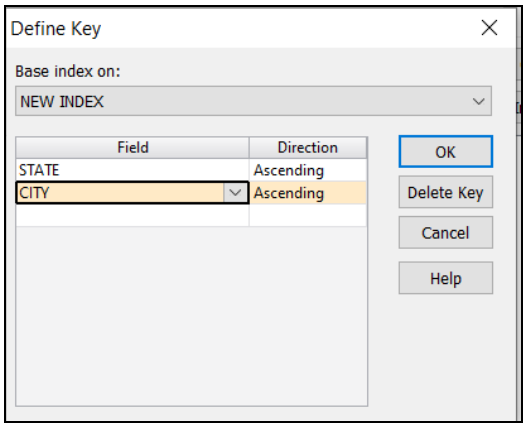


iv) **Key Value Extraction:** We use this option to create multiple output databases that contain unique values based on the field you define as a key. The main advantage ofusing a Key Value Extraction is that it efficiently splits the records in a database into many subsets. This feature is like multiple filter in excel.

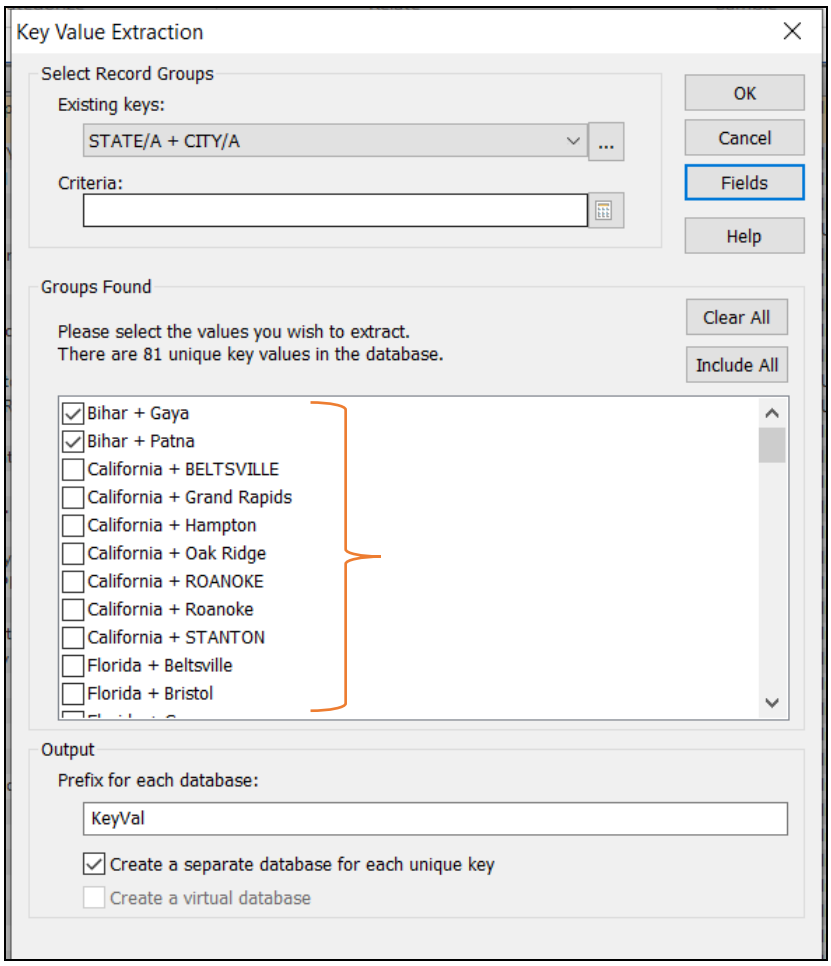
In the below example, we clicked on Key Value Extraction option.



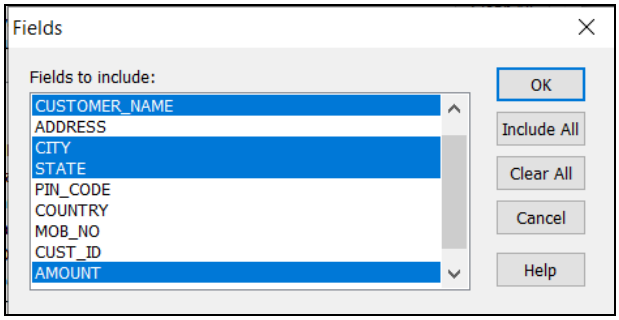
Here, we have defined State and City as the Key value shown below:



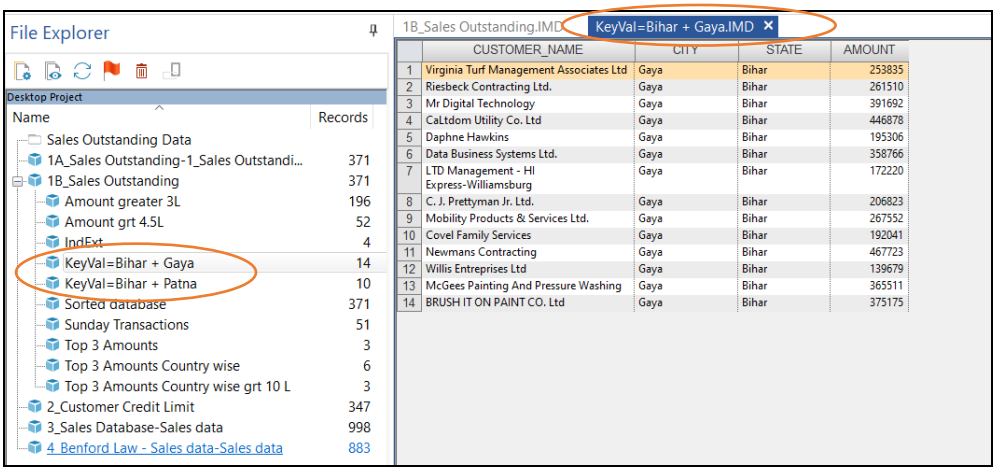
Once we click OK in Define Key option, we get a list of groups (as shown below). It shows that there are 81 combinations. If we click OK now, it will create 81 separate files. But for the sake of convenience, we have selected only 2 combinations as shown in image. Each file/database will contain the prefix “KeyVal”.



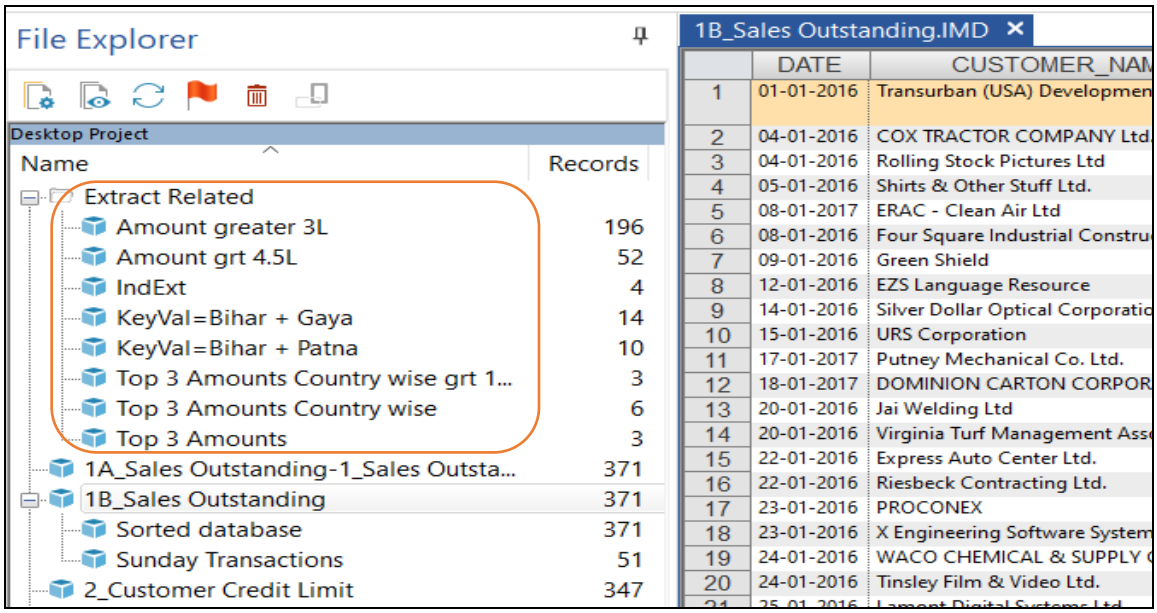
In the fields dialogue box, we have selected only 4 columns i.e. Customer Name, City, State and Amount.



Two new output files/databases are created as shown below:



We created a folder “Extract related” in File Explorer and moved all the extraction related files in that folder.



10.2 SUMMARISATION

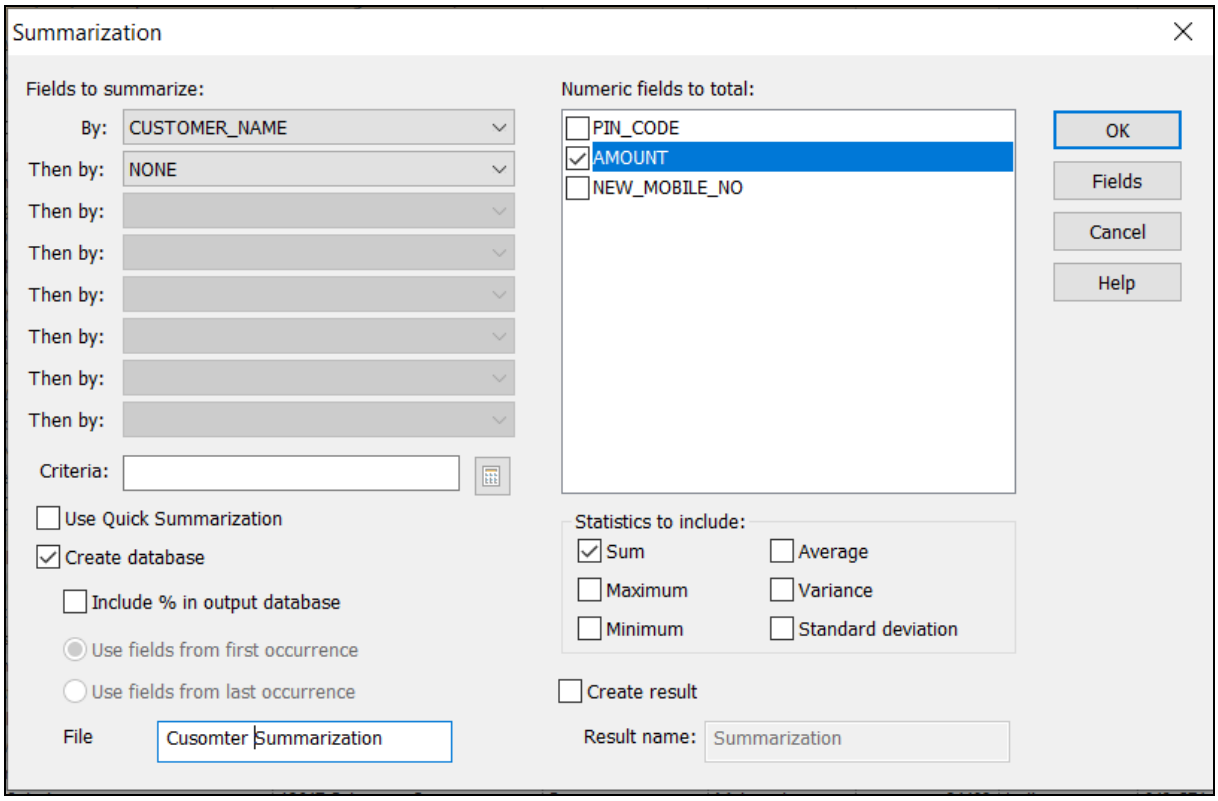
This feature is available in **Analysis Tab > Categorize > Summarization**. The Summarization group the records based on the fields specified. It’s then possible to create a result with certain statistics, including sum, minimum, average, or maximum value for a specified field. We can do summarisation up to 8 levels.

Example 1: We have seen in field statistics that there are multiple names, mobile number and Cust ID. In this example, we want to summarise the data based on Customer Name.

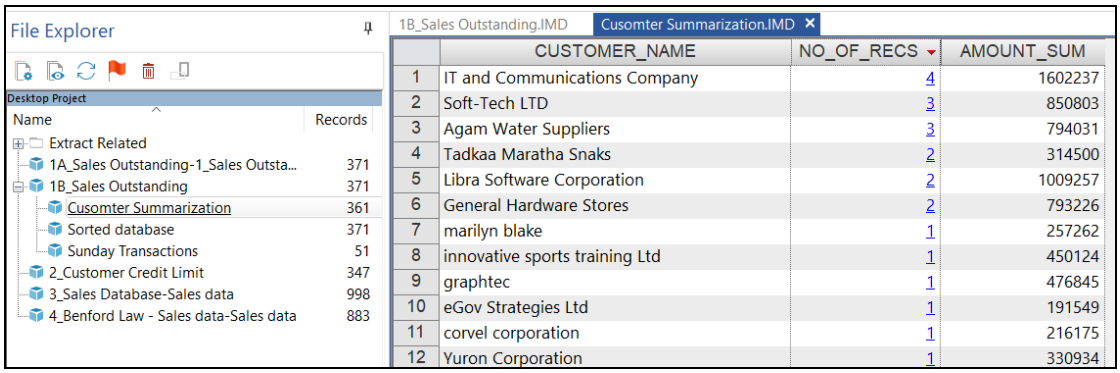
When we click on Summarization option, we get the Summarisation dialoguebox.

— **Filed to summarise by :** As we want to summarise by Customer, hence, we select “CUSTOMER_NAME”.

- **Numeric fields to total :** It shows all the numeric fields i.e. AMOUNT, PIN_CODE and NEW MOBILE NUMBER. We are selecting AMOUNT column.
- **Statistics to include :** Sum is ticked
- By default there is tick on “Create Database” box. It will create a separate file for this summarisation.
- **File name :** Customer Summarization



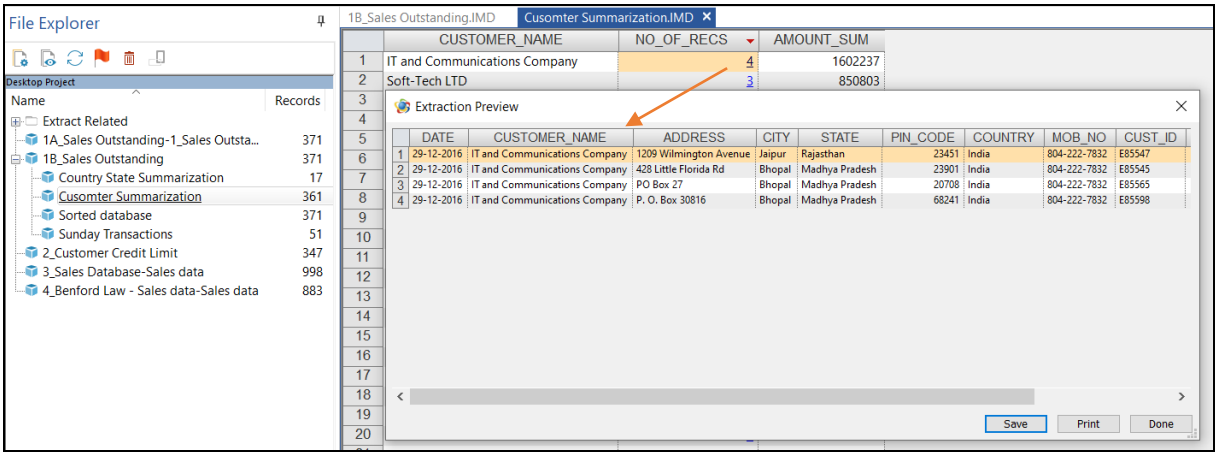
After giving all input, when we click on OK button, a separate new database is created in File Explorer named as “Customer Summarization” with 3 columns i.e. Customer Name, Number of Records and Amount sum.



	CUSTOMER_NAME	NO_OF_RECS	AMOUNT_SUM
1	IT and Communications Company	4	1602237
2	Soft-Tech LTD	3	850803
3	Agam Water Suppliers	3	794031
4	Tadkaa Maratha Snaks	2	314500
5	Libra Software Corporation	2	1009257
6	General Hardware Stores	2	793226
7	marilyn blake	1	257262
8	innovative sports training Ltd	1	450124
9	graphtec	1	476845
10	eGov Strategies Ltd	1	191549
11	corvel corporation	1	216175
12	Yuron Corporation	1	330934

When we sort the column "NO_OF_RECS" in descending order, we can see that there are many parties with common name. It can be possible that they might be having separate Customer ID or they are duplicate records.

If we click on any record number (say 4) highlighted in blue colour, then it shows the list of those 4 records as follows:

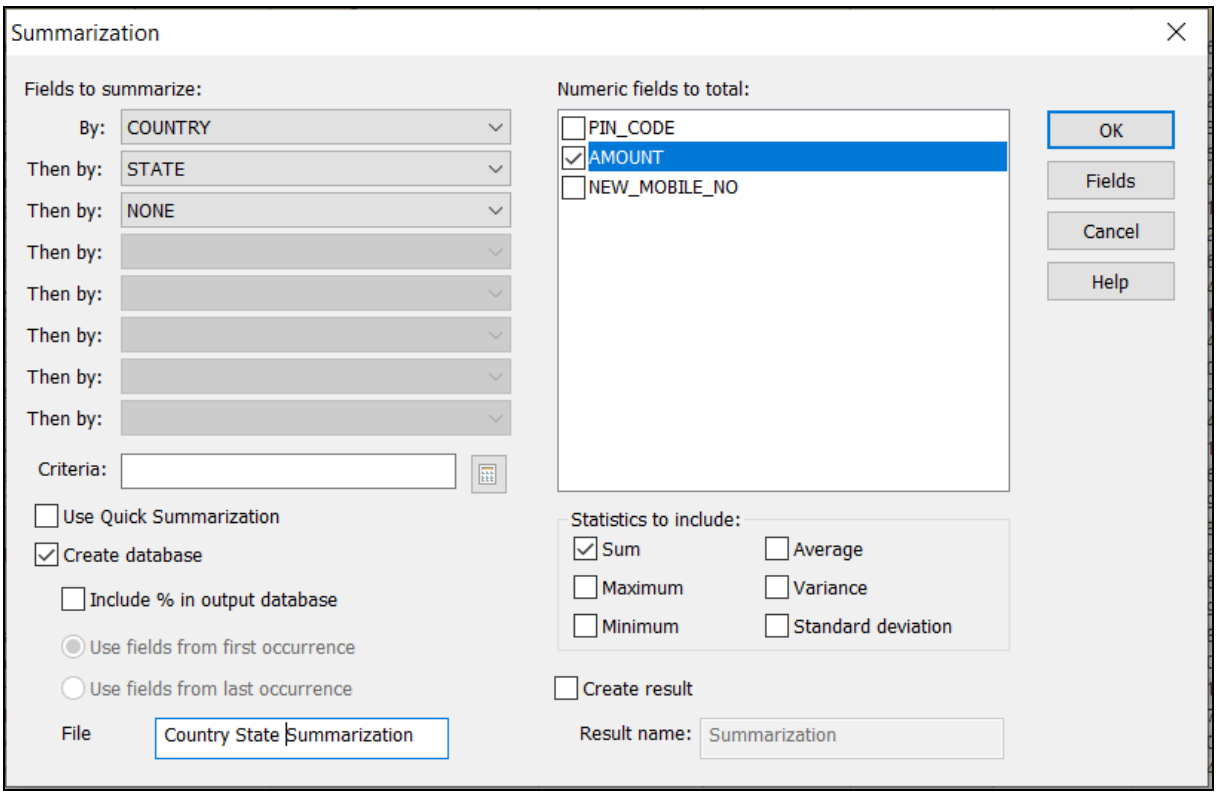


Once review done, we can Save / Print / Done (close) these records as per our requirement.

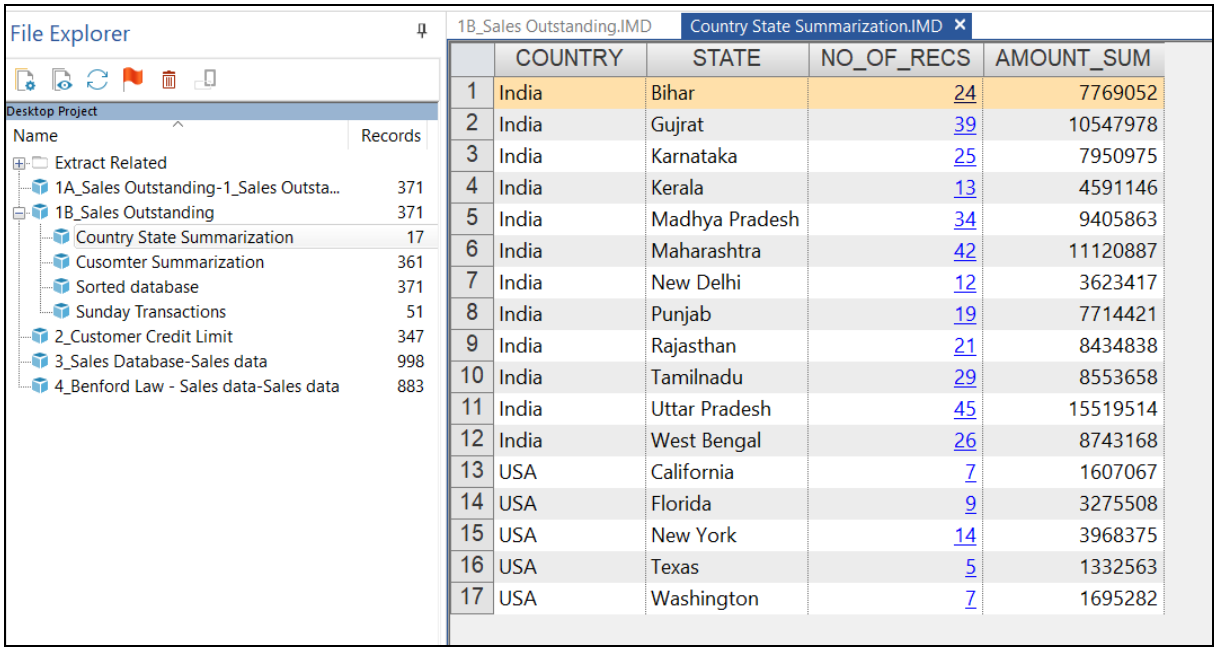
Example 2: Now we want to summarise by Country and State. Here instead of Re-run, we are going to create query again. That it why, we need to go back to original base file i.e. "1B_Sales Outstanding" and make query in this file.

The inputs are as under:

- **Filed to summarise by :** First by Country and then by State
- **Numeric fields to total :** AMOUNT column.
- **Statistics to include :** Sum is ticked
- By default there is tick on "Create Database" box. It will create a separate file for this summarisation.
- **File name :** Country State Summarization



Once we click on OK button, a new file is created in File Explorer with summary by Country and State showing number of records and sum of Amount :



	COUNTRY	STATE	NO_OF_RECS	AMOUNT_SUM
1	India	Bihar	24	7769052
2	India	Gujrat	39	10547978
3	India	Karnataka	25	7950975
4	India	Kerala	13	4591146
5	India	Madhya Pradesh	34	9405863
6	India	Maharashtra	42	11120887
7	India	New Delhi	12	3623417
8	India	Punjab	19	7714421
9	India	Rajasthan	21	8434838
10	India	Tamilnadu	29	8553658
11	India	Uttar Pradesh	45	15519514
12	India	West Bengal	26	8743168
13	USA	California	7	1607067
14	USA	Florida	9	3275508
15	USA	New York	14	3968375
16	USA	Texas	5	1332563
17	USA	Washington	7	1695282

10.3 STRATIFICATION

Stratification helps to create a view of data. Creating a view lets us build graphs, drill down into the data, and perform other analysis tasks. A database option allows to do more after the stratification. A stratified database contains a special STRATUM field that shows which value band holds a particular record.

We use the Stratification task to place the data in layers or strata by a date, numeric, or character field. It's then possible to create a result with certain statistics, including sum, minimum, average, or maximum value for a specified field. It helps to look at the distribution of related data with greater ease. What we see in the output is the number of records that fall between the limits we set and the total of a numeric field for each layer. In addition, we can click a link to see specific records within a layer.

We need to provide following inputs:

- **Group By:** Defines a key field used to group the stratification result. This will create a stratification result for each unique value in the selected field.
- **Field to stratify:** The field which we want to stratify
- **Field to total on:** The numerical field which to be totalled
- **Criteria:** Defines the criteria used to select records for processing.
- **Cut-Offs:** Specifies cut-offs to filter group records to be stratified.
- **Lower Limit:** Defines the lowest value that you use to start defining your numeric intervals. The default is 0.
- **Upper Limit:** Defines the highest value that we use to start defining the numeric intervals.
- **Increment:** We need to define the interval for stratification. The increment can be changed as per need even increased or decreased or keep constant based on increment.
- **Output File Name:** Specifies the name of the optional database created when runningthe Stratification task. This will be a copy of the original database but withan additional field identifying which layer each record falls into.

Stratification relies on the creation of value bands and placing the records from a database within these bands. Each value band represents a range of data and provides a means of categorizing the data in certain ways. We can stratify databases based on character, numeric, or date fields. IDEA supports up to 1,000 stratification bands. After we stratify the database, we can use the results for a number of tasks such as:

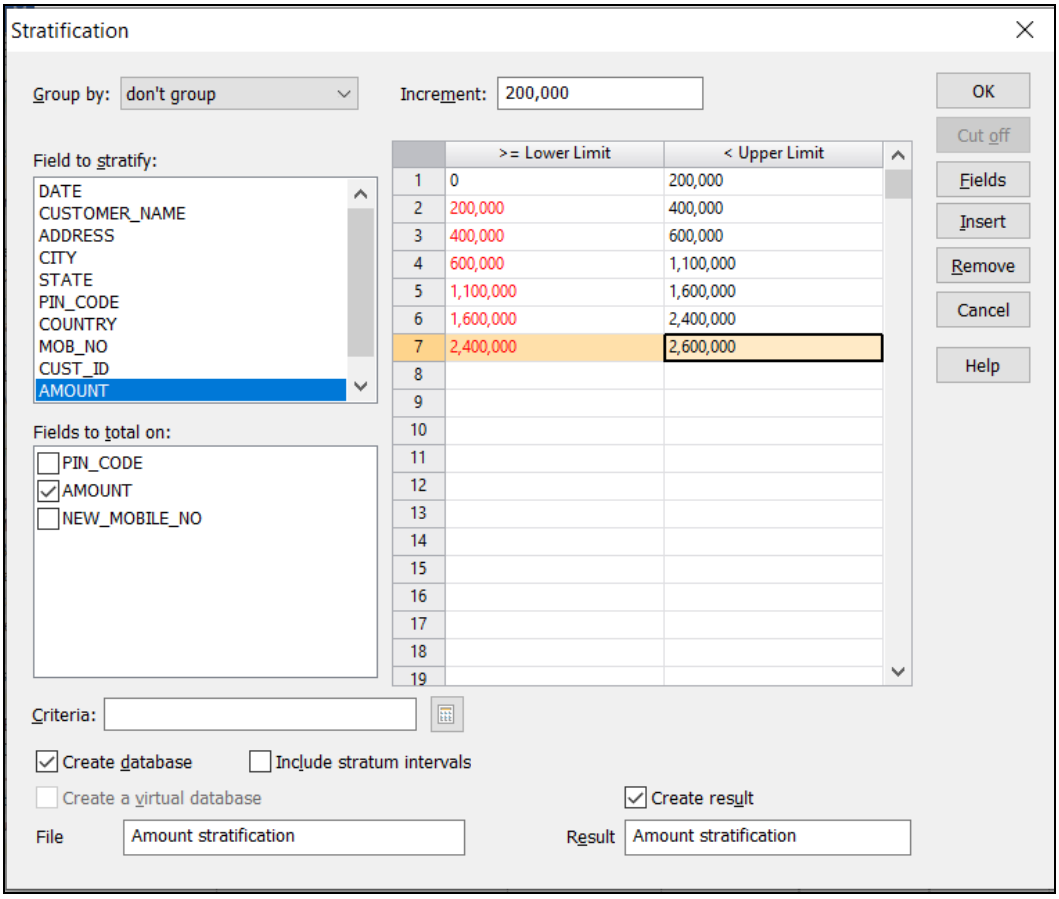
- Totalling the number of records in each band, then using the results to look for expected trends.
- Setting high and low cut-off values to test for exceptional items.
- Creating both a database and a result that contains the details of the band range each record falls within. We can then use this database for further analysis.

Once you have the database stratified, you can obtain random samples from each band to use for further analysis.

Example 1: Stratify the data based on Amount column

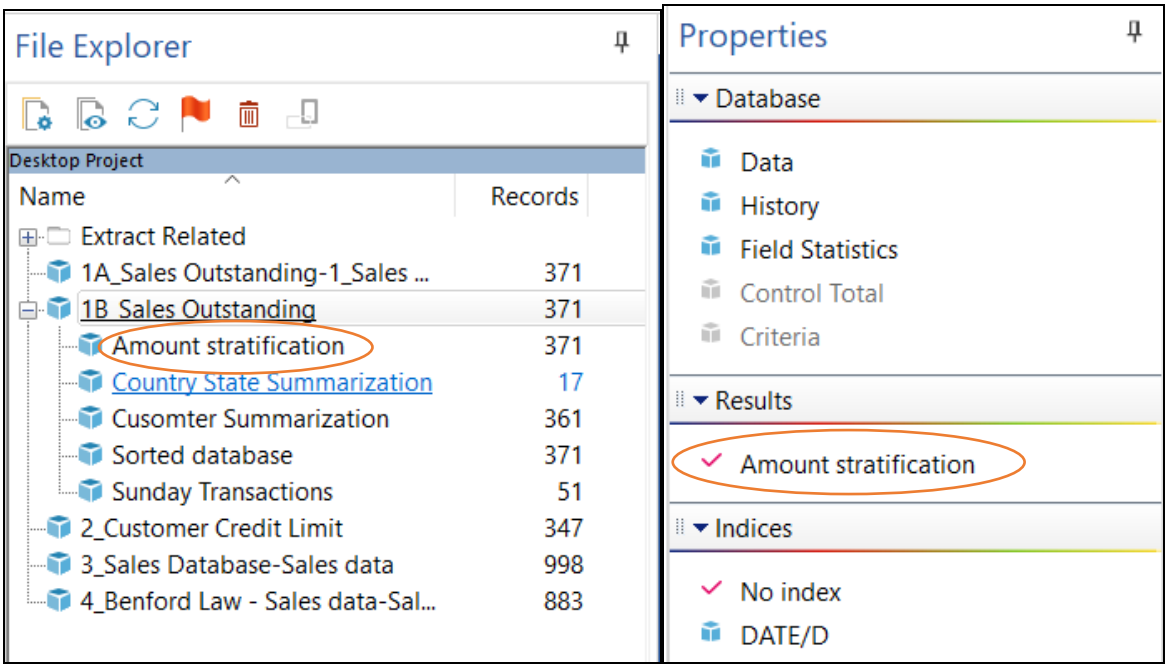
We click on **Analysis Tab > Categorise > Stratification**. We give following inputs:

- **Group By:** don't group
- **Field to stratify:** Amount
- **Field to total on:** Amount
- **Criteria:** No criteria (optional)
- **Cut-Offs:** No cut-off (optional) – not active
- **Lower Limit:** System automatically picks lowest amount as Lower limit. We can manually change that to any value. In this example, we are making Lower Limit as 0.
- **Upper Limit:** The Upper limit will be defined as per increment
- **Increment:** We want first 3 intervals of ₹ 200000, then 2 intervals of ₹ 500000, then 1 interval of ₹ 800000, then last interval of ₹ 200000.
- **Output File Name:** Amount stratification



Note: For increasing the increment, we need to click in the blank space next to Lower Limit. The increment will be automatically added in the Lower limit value and in the next row, the Upper Limit will become Lower limit.

As we clicked on option "Create database", it will create a new database in File Explorer. As there is tick on "Create result", the result will also be created and can be accessed from Properties window.



The Output is as follows:

1B_Sales Outstanding.IMD

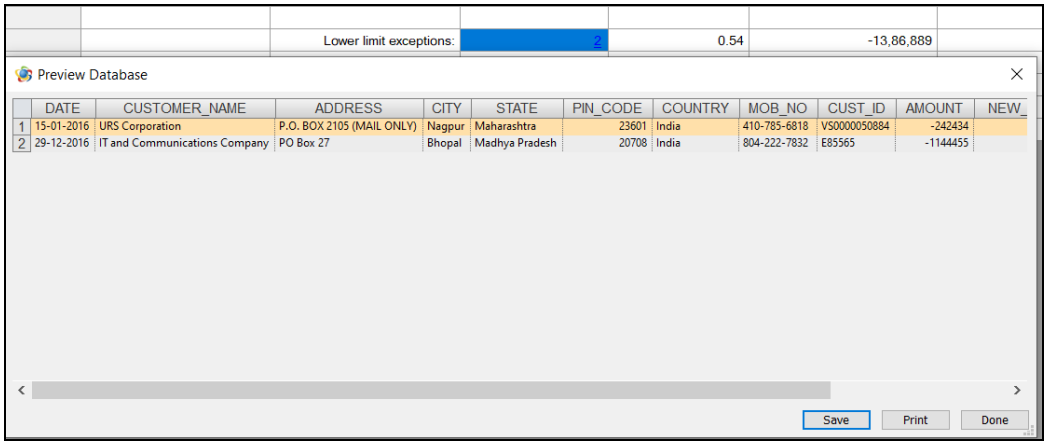
Totalled on: AMOUNT

Stratum #	>= L Limit	< U Limit	# Records	(%) # Records	AMOUNT	(%) AMOUNT
1	0	2,00,000	86	23.18	1,29,98,822	11.22
2	2,00,000	4,00,000	194	52.29	5,91,66,895	51.07
3	4,00,000	6,00,000	84	22.64	3,80,87,732	32.88
4	6,00,000	11,00,000	2	0.54	15,51,505	1.34
5	11,00,000	16,00,000	1	0.27	13,77,780	1.19
6	16,00,000	24,00,000	1	0.27	16,02,237	1.38
7	24,00,000	26,00,000	1	0.27	24,55,630	2.12
		Lower limit exceptions:	2	0.54	-13,86,889	-1.20
		Upper limit exceptions:	0	0.00	0	0.00
		Totals:	371	100.00	11,58,53,712	100.00

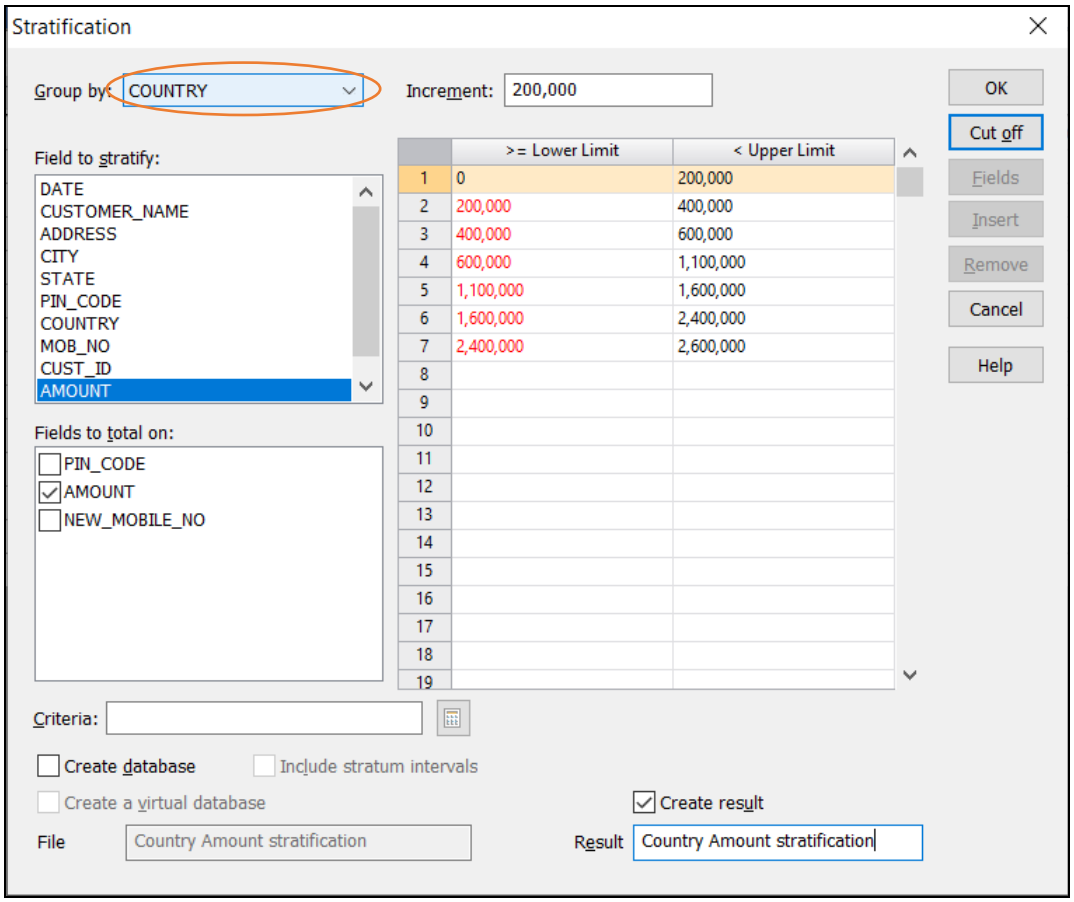
It can be seen that the range started from 0 and first 3 intervals are of ₹ 200000, then next 2 intervals are of ₹ 500000, then one interval of ₹ 800000 and last interval of ₹ 200000.

Here, we can see major concentration is in 2nd Stratum i.e. 200000 to 400000 – 194 instances accounting for 52.29 % total number of records.

There are 2 transactions which are shown as “Lower limit exceptions”, which means there are 2 transactions below lowest limit i.e. 0. When we click on the hyperlink, the output is as follows:



In the same example, if we want to group this data by Country, then we click on Re-run and in “Group by” option, we select field “Country”. This time we are not clicking on option “Create database”. Hence, no new database will be created.



The output will look like this for each country:

1B_Sales Outstanding.IMD

Totalled on: AMOUNT COUNTRY= India

Stratum #	>= L Limit	< U Limit	# Records	(%) # Records	AMOUNT	(%) AMOUNT
1	0	2,00,000	75	22.80	1,14,22,903	10.99
2	2,00,000	4,00,000	172	52.28	5,29,25,921	50.90
3	4,00,000	6,00,000	75	22.80	3,40,25,830	32.73
4	6,00,000	11,00,000	2	0.61	15,51,505	1.49
5	11,00,000	16,00,000	1	0.30	13,77,780	1.33
6	16,00,000	24,00,000	1	0.30	16,02,237	1.54
7	24,00,000	26,00,000	1	0.30	24,55,630	2.36
		Lower limit exceptions:	2	0.61	-13,86,889	-1.33
		Upper limit exceptions:	0	0.00	0	0.00
		Totals:	329	100.00	10,39,74,917	100.00

1B_Sales Outstanding.IMD

Totalled on: AMOUNT COUNTRY= USA

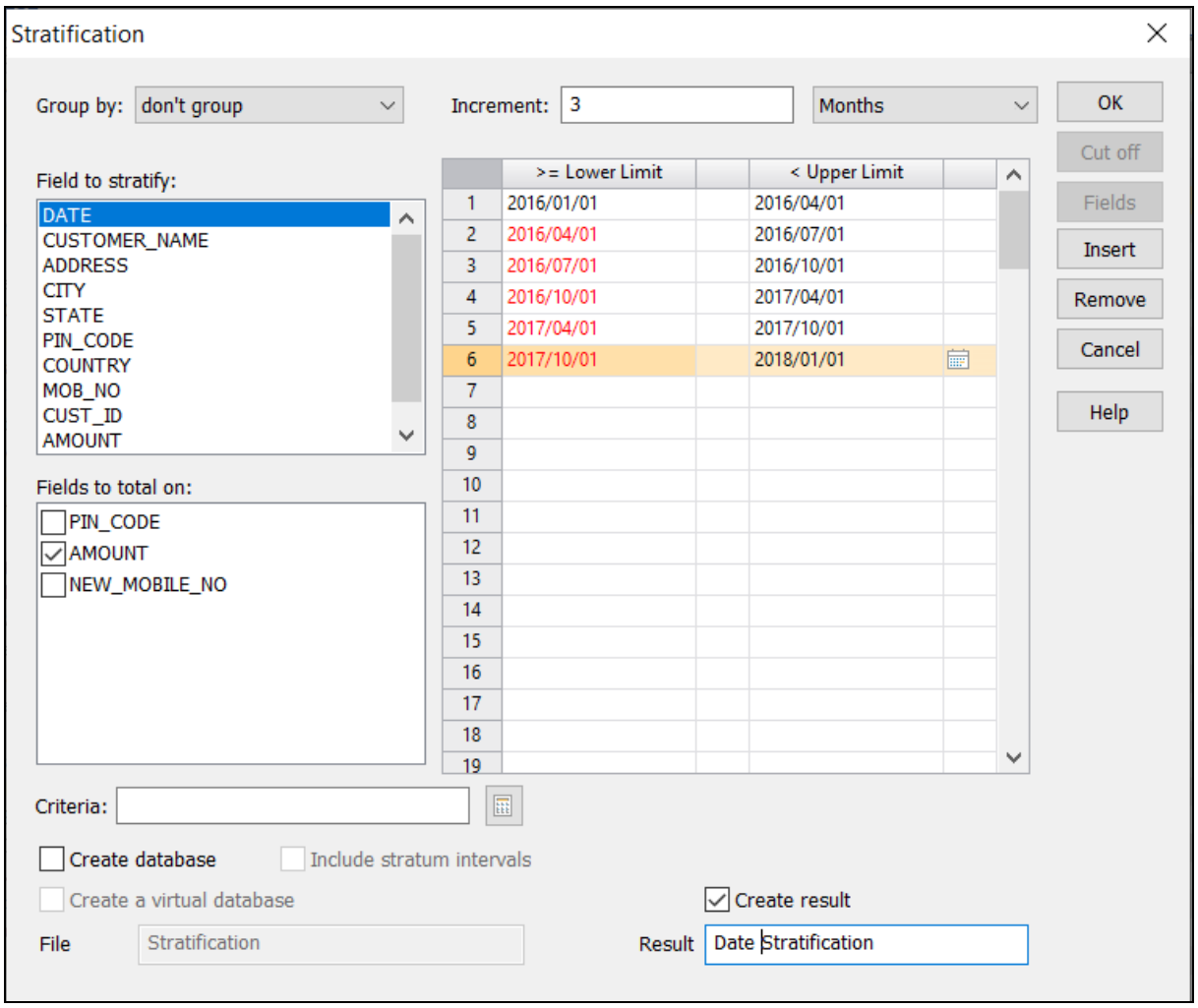
Stratum #	>= L Limit	< U Limit	# Records	(%) # Records	AMOUNT	(%) AMOUNT
1	0	2,00,000	11	26.19	15,75,919	13.27
2	2,00,000	4,00,000	22	52.38	62,40,974	52.54
3	4,00,000	6,00,000	9	21.43	40,61,902	34.19
4	6,00,000	11,00,000	0	0.00	0	0.00
5	11,00,000	16,00,000	0	0.00	0	0.00
6	16,00,000	24,00,000	0	0.00	0	0.00
7	24,00,000	26,00,000	0	0.00	0	0.00
		Lower limit exceptions:	0	0.00	0	0.00
		Upper limit exceptions:	0	0.00	0	0.00
		Totals:	42	100.00	1,18,78,795	100.00

Example 2: Stratify the data based on Date column

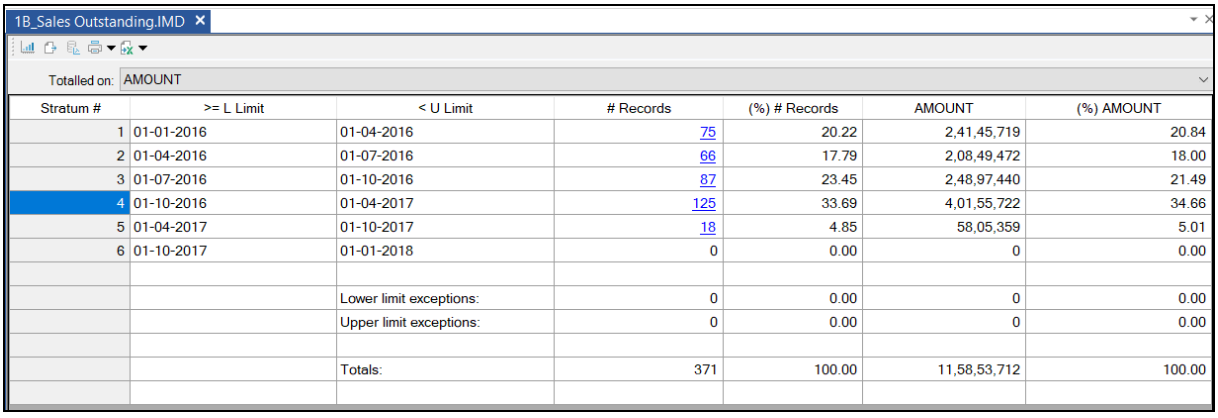
Following inputs given:

- **Field to stratify:** Date column
- **Field to total on:** Amount
- **Increment:** Month (first 3 interval by 3 month, the next 2 interval by 6 month, last interval by 3 month).

  Here we have choice to select interval either in Days, Weeks, Months or Years. For adding the incremental date in Lower Limit date, we need to click in the blank space in the Upper limit, next to Lower Limit in same row. The increment will be added in Lower Limit as well as it will be displayed in Lower Limit in next row automatically.



The output is as follows:



1B_Sales Outstanding.IMD

Totalled on: AMOUNT

Stratum #	>= L Limit	< U Limit	# Records	(%) # Records	AMOUNT	(%) AMOUNT
1	01-01-2016	01-04-2016	75	20.22	2,41,45,719	20.84
2	01-04-2016	01-07-2016	66	17.79	2,08,49,472	18.00
3	01-07-2016	01-10-2016	87	23.45	2,48,97,440	21.49
4	01-10-2016	01-04-2017	125	33.69	4,01,55,722	34.66
5	01-04-2017	01-10-2017	18	4.85	58,05,359	5.01
6	01-10-2017	01-01-2018	0	0.00	0	0.00
		Lower limit exceptions:	0	0.00	0	0.00
		Upper limit exceptions:	0	0.00	0	0.00
		Totals:	371	100.00	11,58,53,712	100.00

Here, we can observe that major concentration is 4th Stratum i.e. 01-10-2016 to 01-04-2017 – 125 records i.e. 33.69% of total number of records.

10.4 AGING

This feature available in **Analysis Tab > Categorize > Ageing**. This feature helps to create the aging of the database from a specified date for up to 6 specified intervals. We can reduce the intervals as per our requirement. In order to use this analysis, the database must have at least one date field and one numeric field.

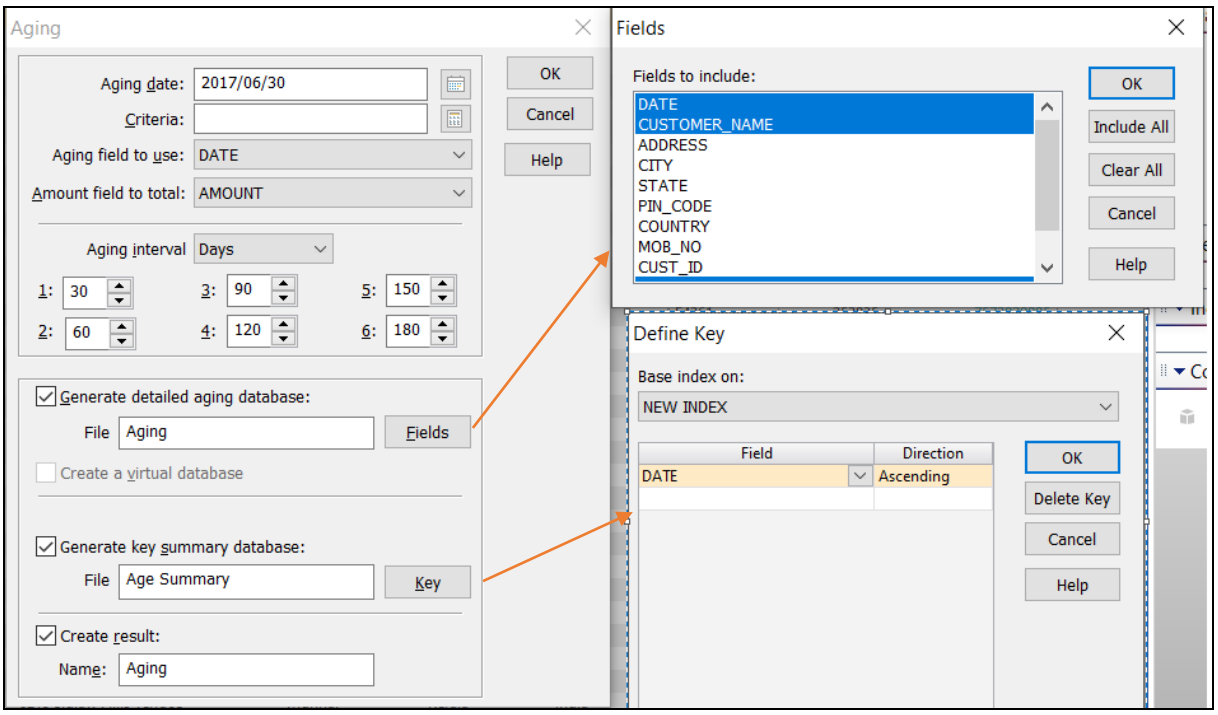
For example, we want to check an Accounts Receivable database for clients who are behind in their payments by so many months and then display the amount that they're behind. We decide on six continuous intervals to use for the analysis, such as 30, 60, 90, 120, 180, and 240 days (normally, the intervals are evenly sized, but we can use any intervals needed). We need to define the Aging Interval whether aging occurs in days, months, or years.

Before we start aging, it is advisable to the earliest and latest dates in Field Statistics. It helps in preparing the aging.

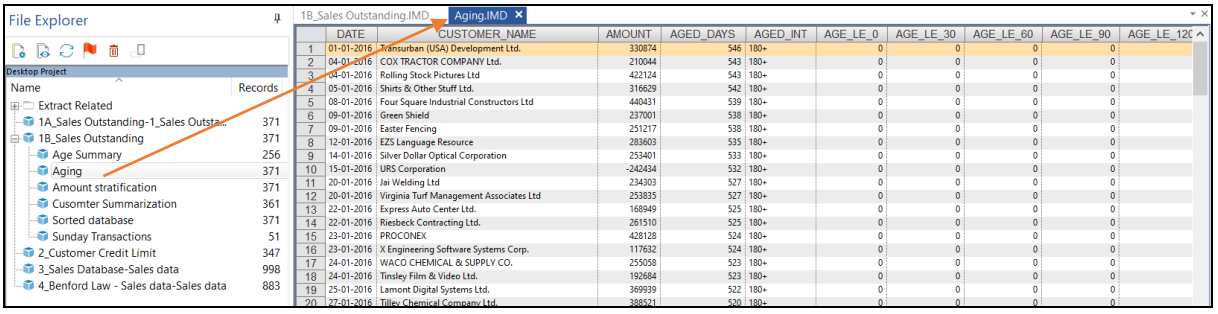
Example 1: In our example, the earliest date in 01-01-2016 and Latest date is 18-07-2017. We want to prepare aging as on 30-06-2017.

The inputs are:

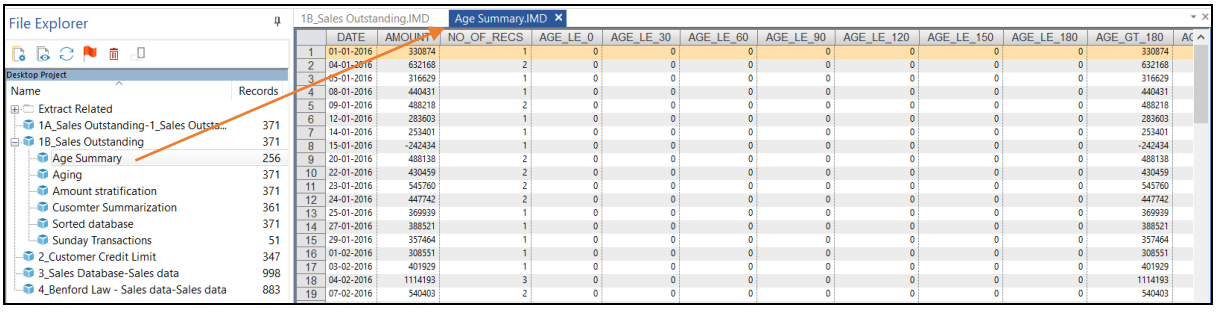
- **Aging date:** 2016/06/30 - we need to enter the date in the format specified by IDEA (i.e. YYYY/MM/DD). Alternatively, we can use calendar as well to select the date.
- **Aging field to use:** Date
- **Amount field to total:** Amount
- **Aging interval:** Days (Default) with the default intervals (30, 60, 90, 120, 150, 180). If we want, we can select Month and Year as well.
- **Generate detailed aging database:** Option ticked.
- **File name:** Aging (default); In the field option, we have selected only 3 columns i.e. Date, Customer Name and Amount
- **Generate kay summary database:** option ticked
- **File name:** Age summary (default); Key is Date
- **Create result:** Aging



The output for Aging database is as follows: It shows aging against each customer.

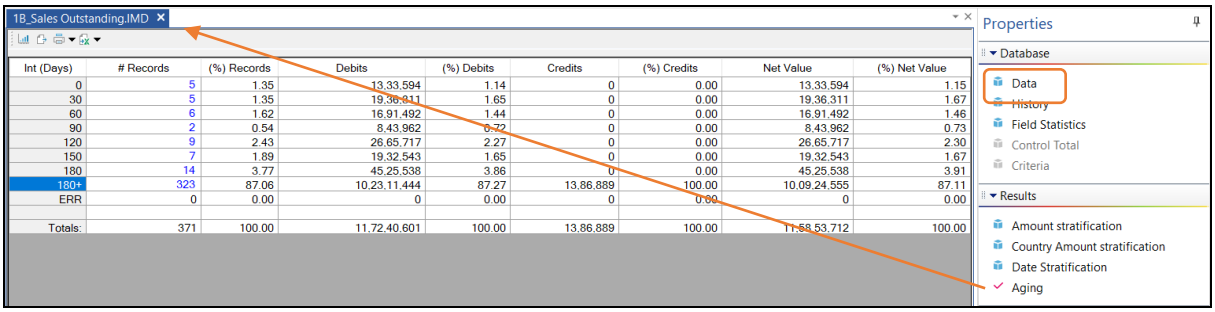


The output for Summary database is as follows: it summarises date-wise each transaction and its gaining.



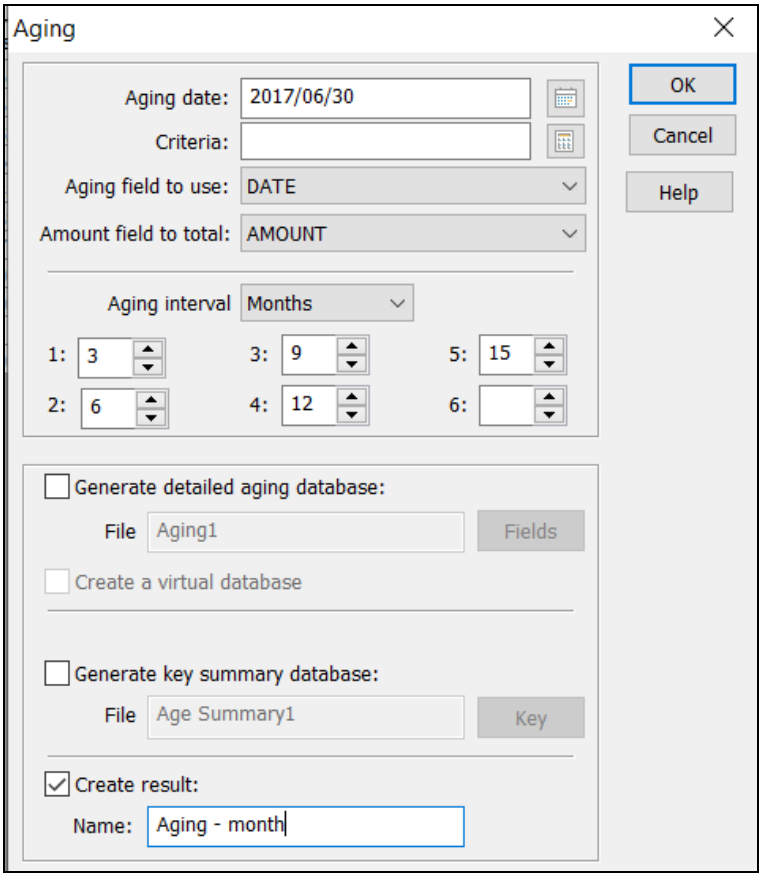
The output for aging result is as follows:

It summarises all the transactions and groups them as per intervals specified. In the result, we can see that 87.06% of records are more than 180 days old. It means there is ling outstanding, which needs to be focused. In case, we want detailed list of those parties, we can click on hyperlink.

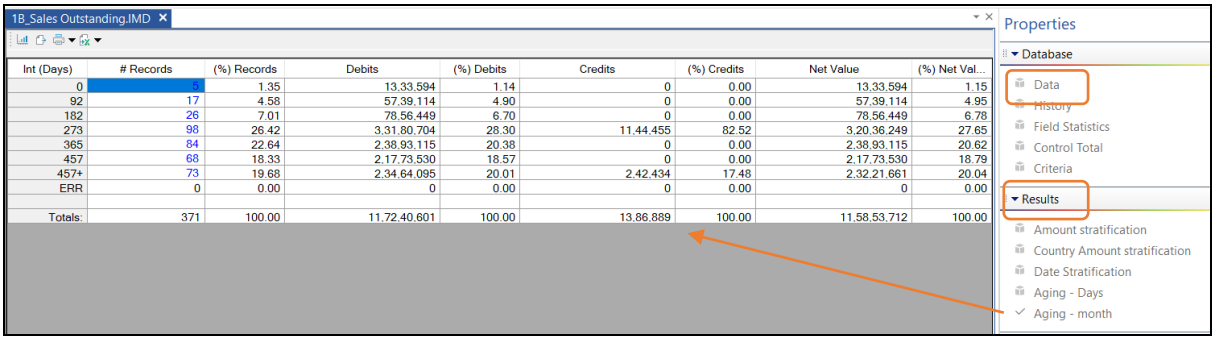


Note: The original data charged to Aging. In case, we want to see original data, we need to click on Data button in Database under Properties.

Example 2: In case, we want to change the interval to months (3, 6, 9, 12, 15). Not adding the 6th interval.



The month aging output is as follows:



Int (Days)	# Records	(%) Records	Debits	(%) Debits	Credits	(%) Credits	Net Value	(%) Net Val...
0	5	1.35	13,33,594	1.14	0	0.00	13,33,594	1.15
92	17	4.58	57,39,114	4.90	0	0.00	57,39,114	4.95
182	26	7.01	78,56,449	6.70	0	0.00	78,56,449	6.78
273	98	26.42	3,31,80,704	28.30	11,44,455	82.52	3,20,36,249	27.65
365	84	22.64	2,38,93,115	20.38	0	0.00	2,38,93,115	20.62
457	68	18.33	2,17,73,530	18.57	0	0.00	2,17,73,530	18.79
457+	73	19.68	2,34,64,095	20.01	2,42,434	17.48	2,32,21,661	20.04
ERR	0	0.00	0	0.00	0	0.00	0	0.00
Totals:	371	100.00	11,72,40,601	100.00	13,86,889	100.00	11,58,53,712	100.00

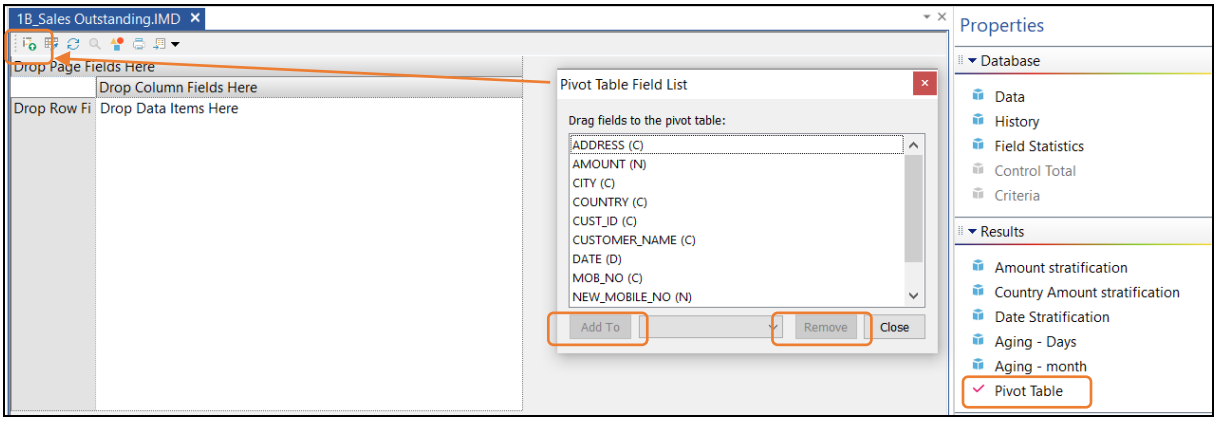
We can change the view of results by clicking on each heading i.e.

- Amount stratification
- Country Amount Stratification
- Date Stratification
- Aging – Days
- Aging – Months

10.5 PIVOT TABLE

The PivotTable is used to display specified data using a grid format. Pivot tables are helpful because it can help in summarizing the information or gain a new view of the data without becoming overwhelmed with detail. It creates a dynamic summary of the data in table format. This feature is available in **Analysis Tab > Categorize > Pivot Table**. Once we click on Pivot Table option, it asks for the name. the default name is “Pivot Table”.

The following structure appears:



On the left side, we have blank Pivot Table structure and on right side we have Pivot Table Field List. From the field list, we need to drag and drop the Fields in either Row or Column area as we want to present. Alternatively, we can use “Add To” and “Remove” buttons. These 2 buttons will be activated once we select a

field. Data values need to be dropped in Data area. In case, we want to filter the data, we need to drop the fields in Page area.

It is to be noted that against each field name, there is either C or N is written. Here, C stands for Character Field and N stands for Numeric Field.

By chance, if we close the Field list, then we need to click at the option above in the image.

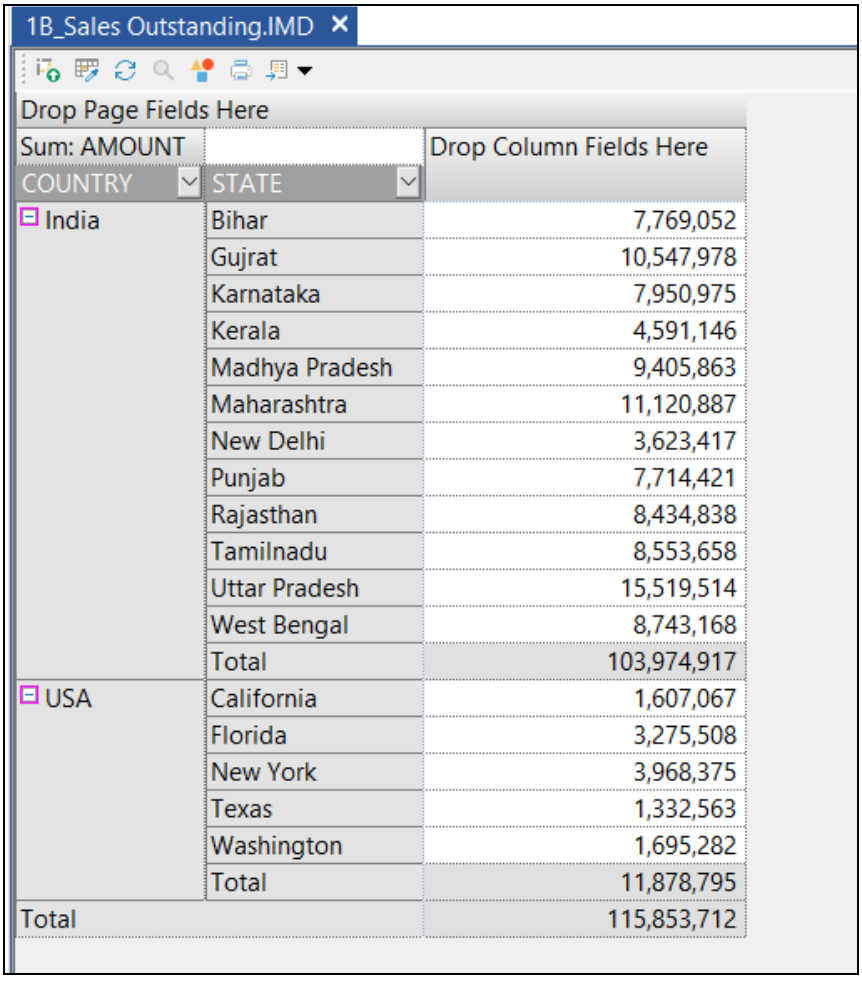
Example 1: Country-State-wise summary

— **Row Area** – First Country and then State field

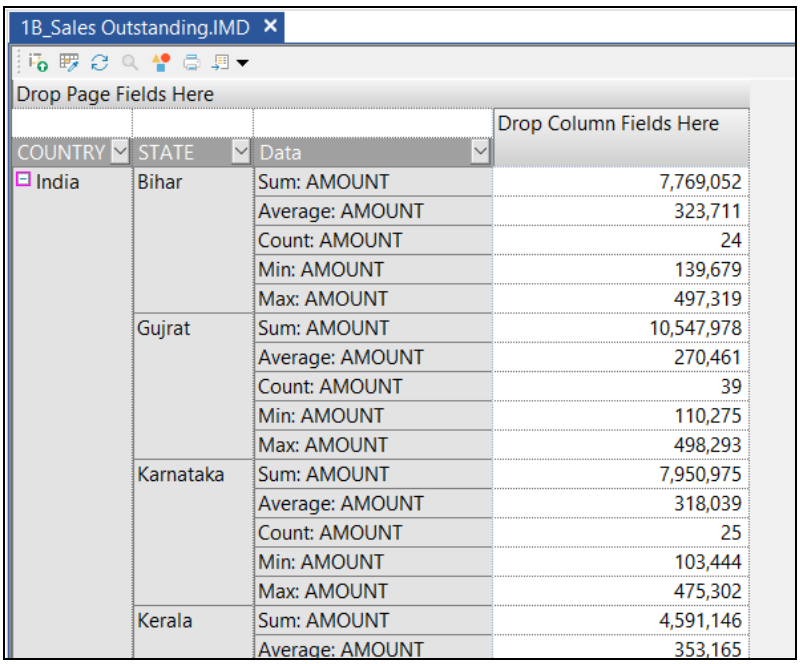
— **Column Area** – No field (Blank)

— **Data Area** – Amount field

We can add additional columns also.



We can change the display of Amount column to Sum, Average, Count, Minimum and Maximum.

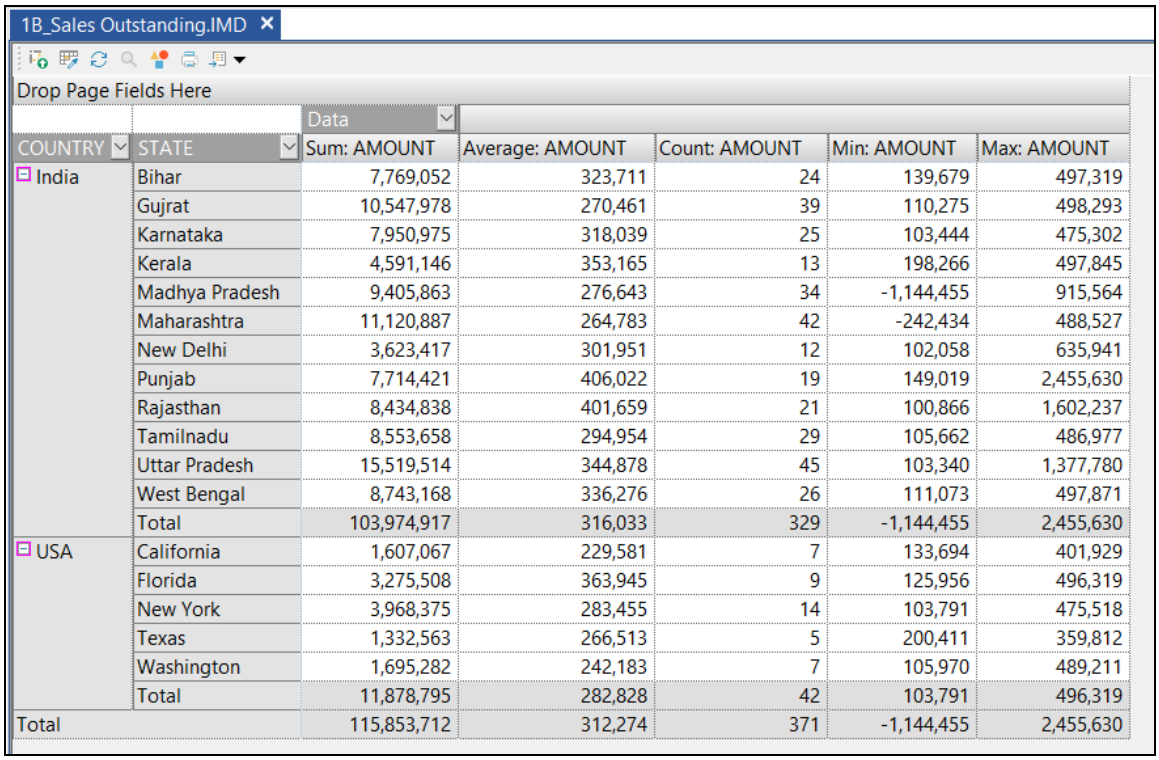


1B_Sales Outstanding.IMD

Drop Page Fields Here

COUNTRY	STATE	Data	Drop Column Fields Here
India	Bihar	Sum: AMOUNT	7,769,052
		Average: AMOUNT	323,711
		Count: AMOUNT	24
		Min: AMOUNT	139,679
		Max: AMOUNT	497,319
	Gujrat	Sum: AMOUNT	10,547,978
		Average: AMOUNT	270,461
		Count: AMOUNT	39
		Min: AMOUNT	110,275
		Max: AMOUNT	498,293
	Karnataka	Sum: AMOUNT	7,950,975
		Average: AMOUNT	318,039
		Count: AMOUNT	25
		Min: AMOUNT	103,444
		Max: AMOUNT	475,302
	Kerala	Sum: AMOUNT	4,591,146
		Average: AMOUNT	353,165

It can be seen that the Amount column is aligned horizontally. Now we want to place all the value in vertical form. We need to drag the Data column and drop in “Column field”.

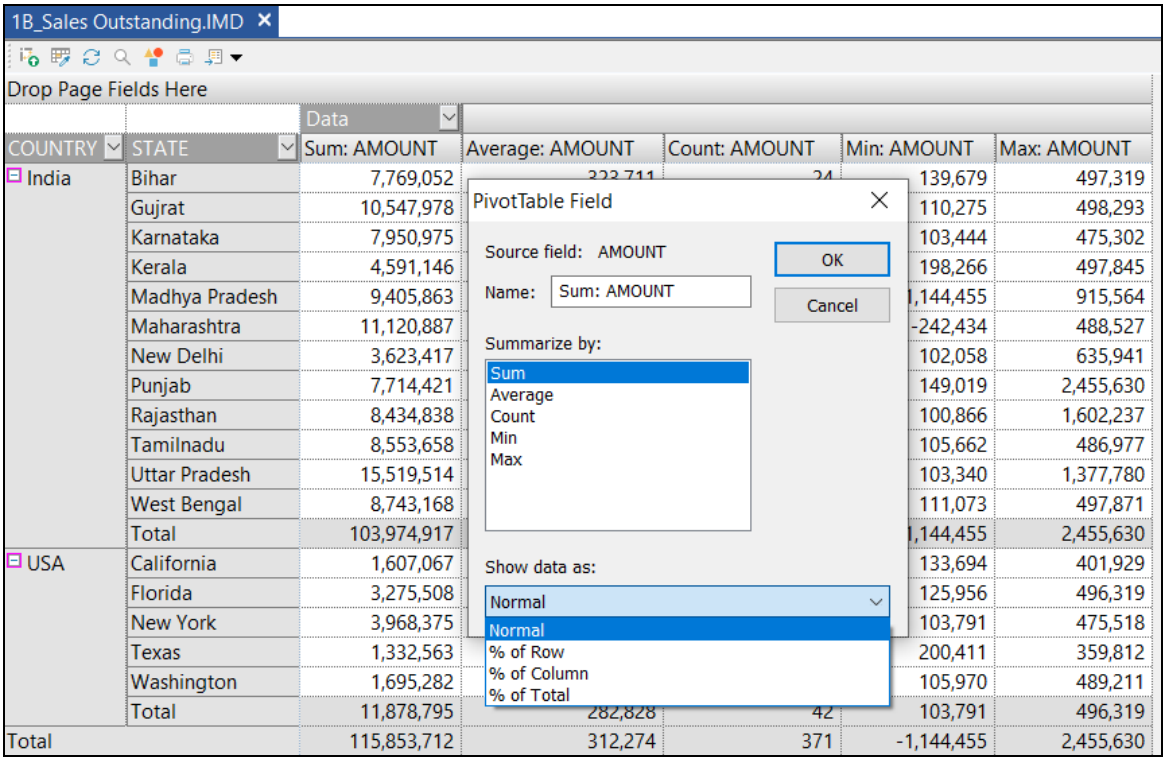


1B_Sales Outstanding.IMD

Drop Page Fields Here

		Data				
COUNTRY	STATE	Sum: AMOUNT	Average: AMOUNT	Count: AMOUNT	Min: AMOUNT	Max: AMOUNT
India	Bihar	7,769,052	323,711	24	139,679	497,319
	Gujrat	10,547,978	270,461	39	110,275	498,293
	Karnataka	7,950,975	318,039	25	103,444	475,302
	Kerala	4,591,146	353,165	13	198,266	497,845
	Madhya Pradesh	9,405,863	276,643	34	-1,144,455	915,564
	Maharashtra	11,120,887	264,783	42	-242,434	488,527
	New Delhi	3,623,417	301,951	12	102,058	635,941
	Punjab	7,714,421	406,022	19	149,019	2,455,630
	Rajasthan	8,434,838	401,659	21	100,866	1,602,237
	Tamilnadu	8,553,658	294,954	29	105,662	486,977
	Uttar Pradesh	15,519,514	344,878	45	103,340	1,377,780
	West Bengal	8,743,168	336,276	26	111,073	497,871
	Total	103,974,917	316,033	329	-1,144,455	2,455,630
USA	California	1,607,067	229,581	7	133,694	401,929
	Florida	3,275,508	363,945	9	125,956	496,319
	New York	3,968,375	283,455	14	103,791	475,518
	Texas	1,332,563	266,513	5	200,411	359,812
	Washington	1,695,282	242,183	7	105,970	489,211
	Total	11,878,795	282,828	42	103,791	496,319
Total		115,853,712	312,274	371	-1,144,455	2,455,630

We can change the display of Columns by right clicking on any value in any column (say Sum: Amount column). It shows various options as shown below. We can change display as required. It is to be noted that Pivot Table in IDEA does not allow 2 columns of same value.



10.6 CHART

This feature is available in **Analysis Tab > Categorize >Chart**. IDEA supports a number of chart and graph types to satisfy a variety of needs. The kind of chart or graph that we select determines the message that the output delivers. There are various type of charts and graphs:

- **Line:** A graph that provides a continuum of data that emphasizes the fluidity of information from one data point to the next without losing the individual data points.
- **Bar:** A chart that emphasizes distinct data points and emphasizes differences between data points.
- **Curve:** A graph that provides a continuum of data that emphasizes the data flow over individual data points. The curve may extend beyond data points in some cases to emphasize the flow.
- **Scatter:** A graph that places no emphasis on flow between data points and places the data points at the centre. Using this kind of graph can help the viewer come to their own conclusions about data patterns. A scatter graph can be good at illustrating clusters of data values.
- **Pie:** A chart that emphasizes the parts of a whole. For example, the monthly cost, as a percentage, for heat over an entire year. A pie chart normally includes a legend and can also include point labels.

- **Area:** A chart that emphasizes the amounts over a baseline (normally 0). The emphasis is on the area under the line, rather than the line itself.

It is always advisable to prepare a chart on Summarised data. Here, we are going to prepare a chart on State Summary.

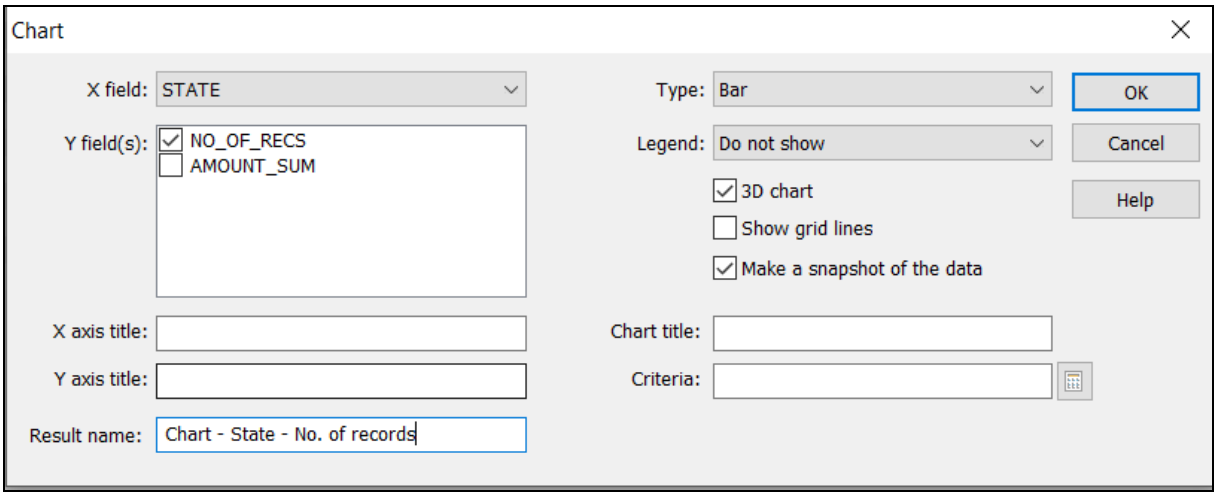
Example 1: State Summary

First we need to summarise the data by State.

1B_Sales Outstanding.IMD | State Summarization.IMD

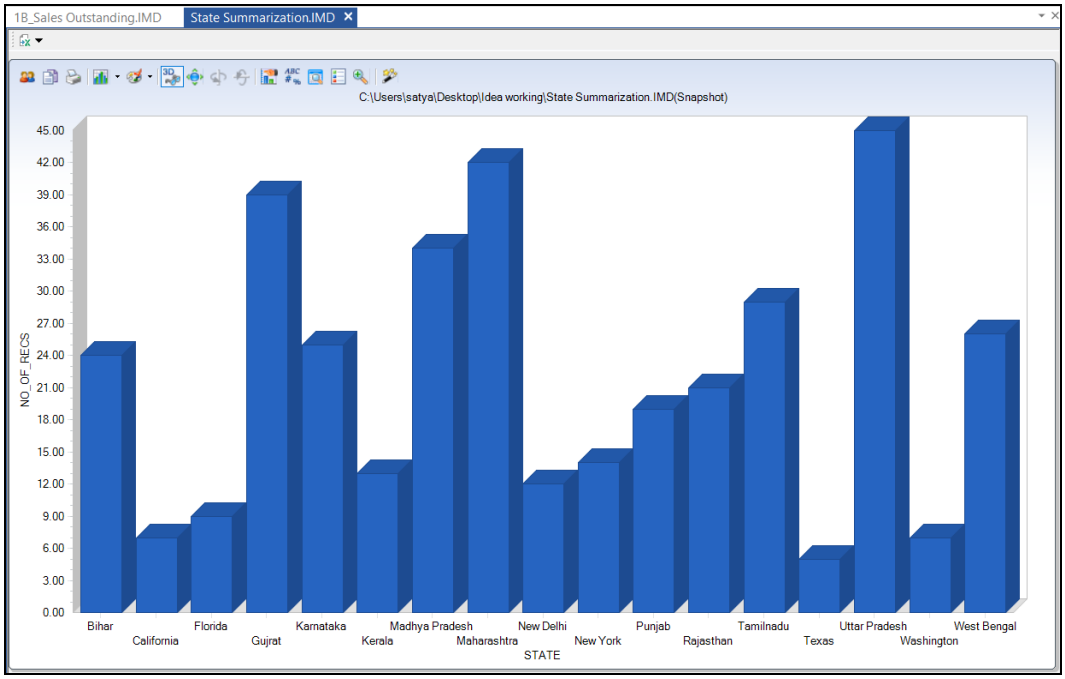
	STATE	NO_OF_RECS	AMOUNT_SUM
1	Bihar	24	7769052
2	California	7	1607067
3	Florida	9	3275508
4	Gujrat	39	10547978
5	Karnataka	25	7950975
6	Kerala	13	4591146
7	Madhya Pradesh	34	9405863
8	Maharashtra	42	11120887
9	New Delhi	12	3623417
10	New York	14	3968375
11	Punjab	19	7714421
12	Rajasthan	21	8434838
13	Tamilnadu	29	8553658
14	Texas	5	1332563
15	Uttar Pradesh	45	15519514
16	Washington	7	1695282
17	West Bengal	26	8743168

Now when we click on Chart option, it shows various options:

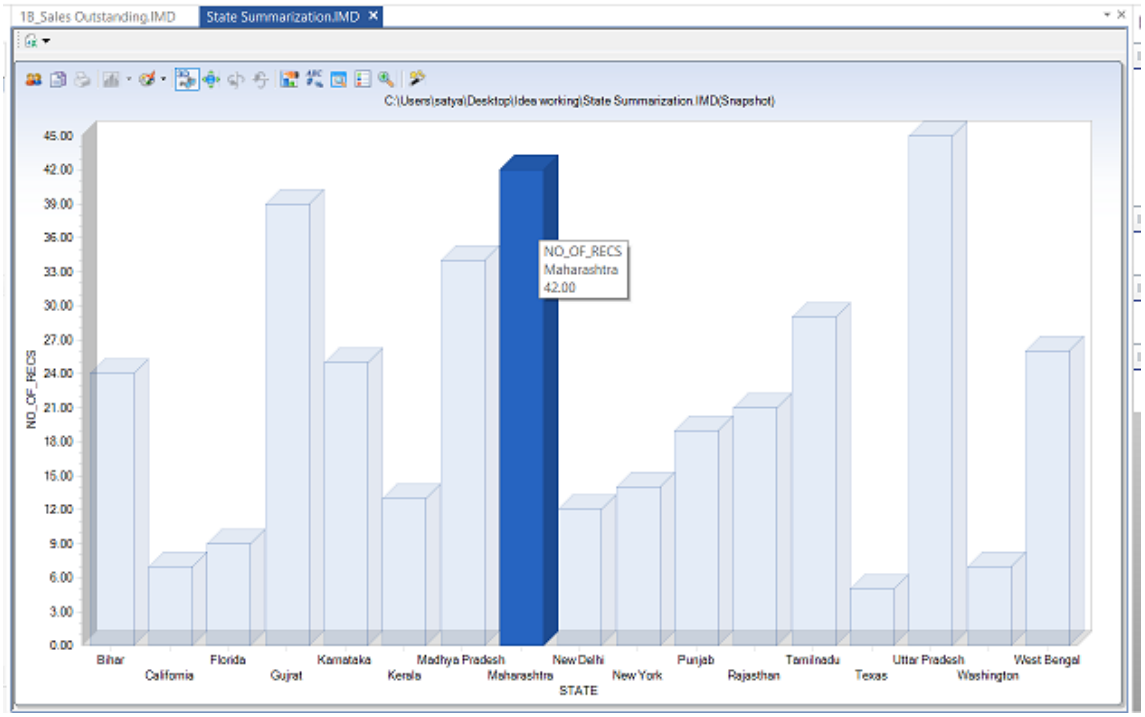


Here, we are preparing a Bar chart for number of records for State.

The output is as follows:



In case, we keep mouse at any State as shown below, it will highlight that particular state and if we left click on that state, it will give option to extract or display the records.



IDEA also supports the addition of a number of effects and additions to the charts and graphs. For example, we can add a 3D effect to the chart or graph. The 3Deffect looks more eye-popping and dramatic, but can reduce the effectiveness of the presentation by obscuring the data to a degree.

10.7 RELATE THE TABLES

The relate feature is like Lookup in Excel. There are 4 different types of Relate features in IDEA:

— Join

— Visual Connector

— Compare

— Append

- **Join Database**

This feature is available under **Analysis Tab > Relate > Join**. The Join Database is used to join two databases together based on a common field, or fields (called a key). Many DBMSs use multiple tables to store information—such as a Customer Outstanding table and a second table that contains Limits assigned to each Customer. The Join Database task could join these two tables together into a single table so that we can see not only the list of customers, but also all of the Outstanding Amount as well as their Limits. The common field, in this case, would likely be a customer name field or Cust ID field as the case may be.

To perform this task, we must consider the two databases, which we want to join. Here, in out example, the Customer Summarisation table will be Primary Table and Customer Limit table will be secondary database. It's important to keep the two tables separate or we will encounter problems later. The join relies on matching records in the primary database to those in the secondary database.

The inputs are as follows:

— Primary database: Customer Summarization (361 unique customers) – 3 columns

— Secondary database: Customer Credit Limit (347 unique Customers) – 2 columns

There are 5 types of Join that IDEA supports:

i) **Match Only:** Records that match the criteria in both databases only appear in output file. A customer in both the file (i.e. Sales Outstanding summary as well as Customer Limit), will only appear.

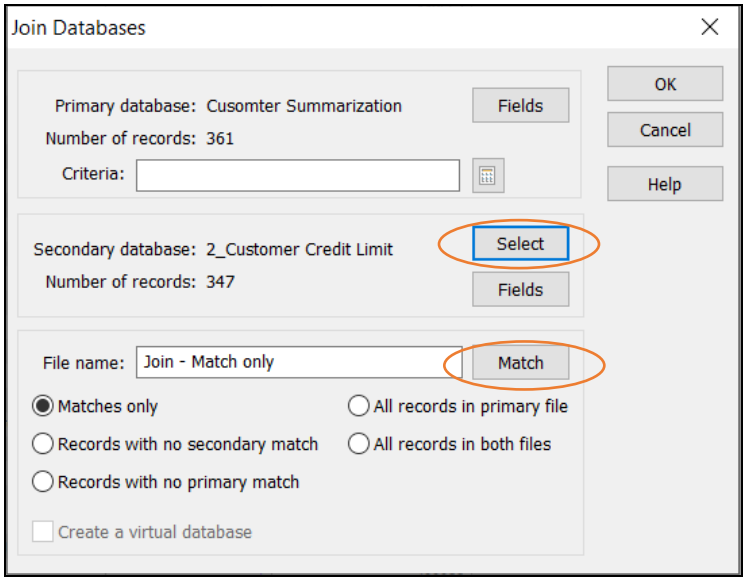
ii) **All records in Primary file:** All records in primary database will appear in the output file with their associated secondary database records. Orphaned secondary records will not appear in the output file. It means all customer will appear, even if there is no limit assigned to them.

iii) **Records with no secondary match:** Records with no secondary match appear in the output. This kind of join is used to find the Customers having no limit.

iv) **Records with no primary match:** Records with no primary match appear in the output. This kind of join is used to find orphaned records in the secondary database – such as limits without ant outstanding or say orders with no customer associated with them (clearly error in the secondary database)

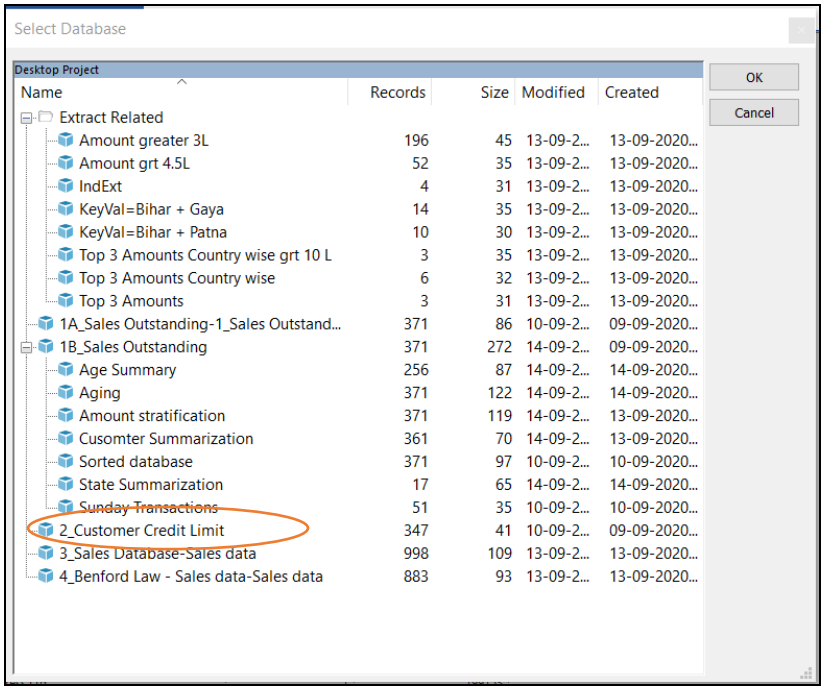
v) **All records in both files:** All records in both databases appear in the output, even if the secondary record is orphaned.

Example 1: Match only

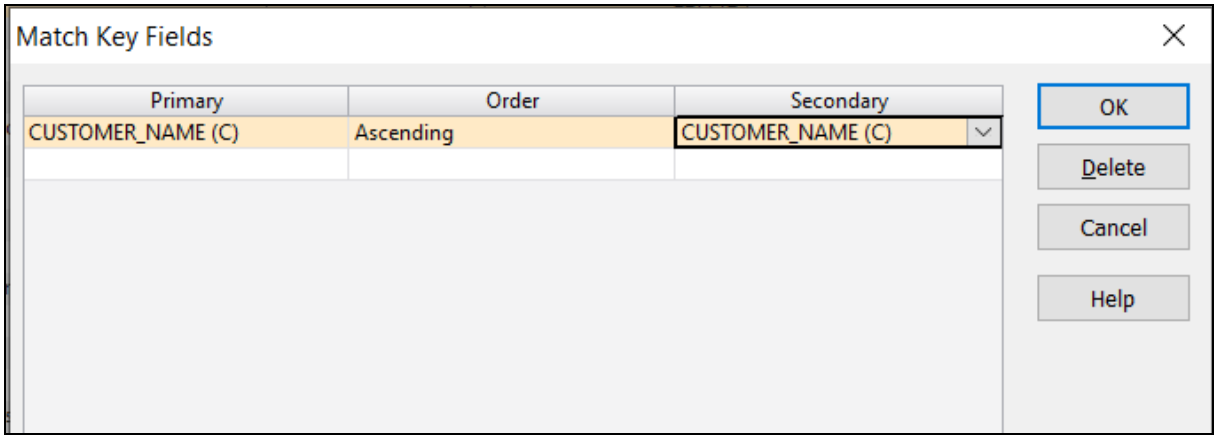


First we need to open the primary database i.e. “Customer Summarization” and then click on Join option.

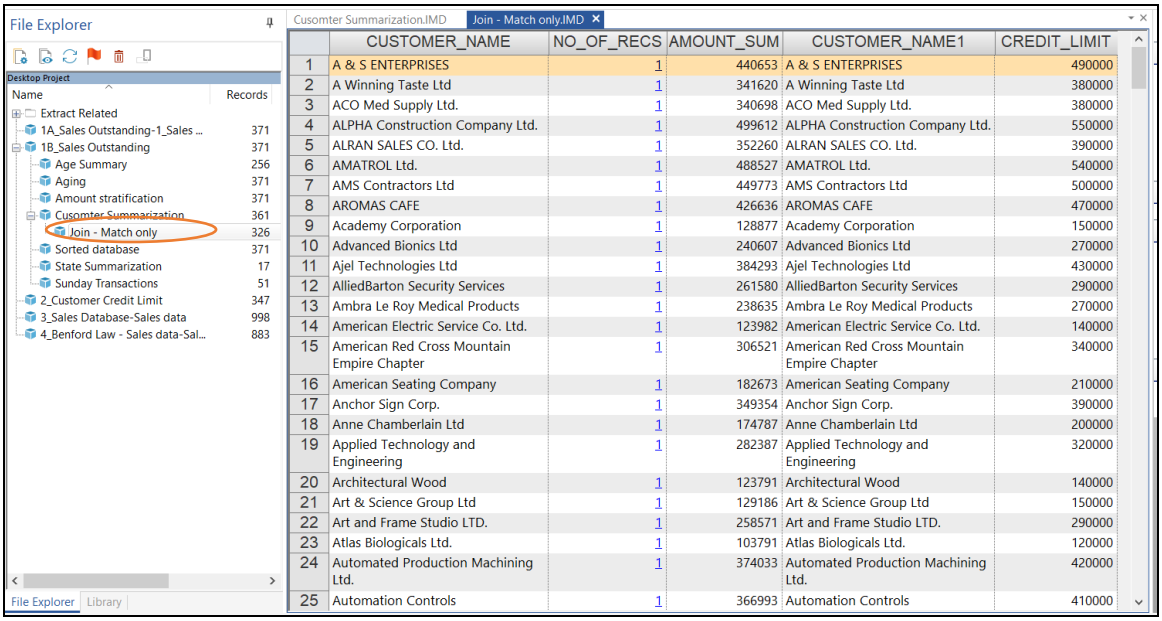
In the Join Database dialogue box, for Secondary database, we click on Select option. Then from the list of available databases, we need to select “Customer Credit Limit” database.



Then we need to select the common field in both databases using Match option. Here, Customer Name is the common field and it is to be ensure that the field type has to be same. Here it is character field. Then click ok.



Then we select option "Match only" and click OK. The output is as follows:



	CUSTOMER_NAME	NO_OF_RECS	AMOUNT_SUM	CUSTOMER_NAME1	CREDIT_LIMIT
1	A & S ENTERPRISES	1	440653	A & S ENTERPRISES	490000
2	A Winning Taste Ltd	1	341620	A Winning Taste Ltd	380000
3	ACO Med Supply Ltd.	1	340698	ACO Med Supply Ltd.	380000
4	ALPHA Construction Company Ltd.	1	499612	ALPHA Construction Company Ltd.	550000
5	ALRAN SALES CO. Ltd.	1	352260	ALRAN SALES CO. Ltd.	390000
6	AMATROL Ltd.	1	488527	AMATROL Ltd.	540000
7	AMS Contractors Ltd	1	449773	AMS Contractors Ltd	500000
8	AROMAS CAFE	1	426636	AROMAS CAFE	470000
9	Academy Corporation	1	128877	Academy Corporation	150000
10	Advanced Bionics Ltd	1	240607	Advanced Bionics Ltd	270000
11	Ajel Technologies Ltd	1	384293	Ajel Technologies Ltd	430000
12	AlliedBarton Security Services	1	261580	AlliedBarton Security Services	290000
13	Ambra Le Roy Medical Products	1	238635	Ambra Le Roy Medical Products	270000
14	American Electric Service Co. Ltd.	1	123982	American Electric Service Co. Ltd.	140000
15	American Red Cross Mountain Empire Chapter	1	306521	American Red Cross Mountain Empire Chapter	340000
16	American Seating Company	1	182673	American Seating Company	210000
17	Anchor Sign Corp.	1	349354	Anchor Sign Corp.	390000
18	Anne Chamberlain Ltd	1	174787	Anne Chamberlain Ltd	200000
19	Applied Technology and Engineering	1	282387	Applied Technology and Engineering	320000
20	Architectural Wood	1	123791	Architectural Wood	140000
21	Art & Science Group Ltd	1	129186	Art & Science Group Ltd	150000
22	Art and Frame Studio LTD.	1	258571	Art and Frame Studio LTD.	290000
23	Atlas Biologicals Ltd.	1	103791	Atlas Biologicals Ltd.	120000
24	Automated Production Machining Ltd.	1	374033	Automated Production Machining Ltd.	420000
25	Automation Controls	1	366993	Automation Controls	410000

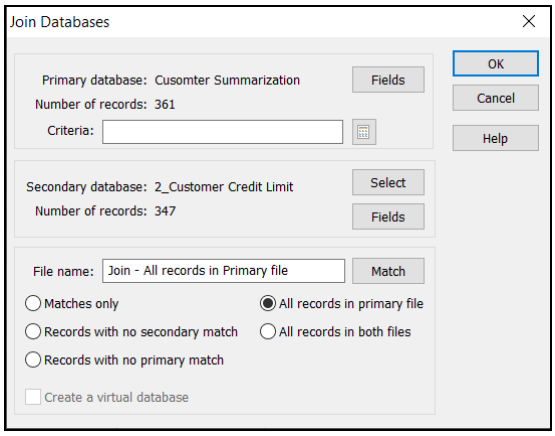
In the output, a new file is created ("Join – Match only" in the File Explorer under Customer Summarization. We can see there are 326 records which is matching in both the files. As we have not selected any specific field, it will show all 5 fields in the Output file (3 from Customer Summarisation and 2 from Customer Limit).

First 3 columns are from Primary Database, followed by columns from Secondary database. Against each common records of Primary database, records from secondary database has been plotted in the output file.

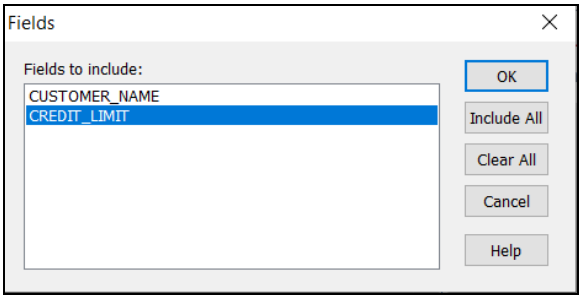
As the output file contains 326 records, it means there are 35 records (361 – 326) of Customer Summarisation (Primary database) not matching in Secondary database. Similarly, there are 21 records (347 – 326) of Customer Limit (Secondary database) not matching in Primary database.

Example 2: All records in Primary file

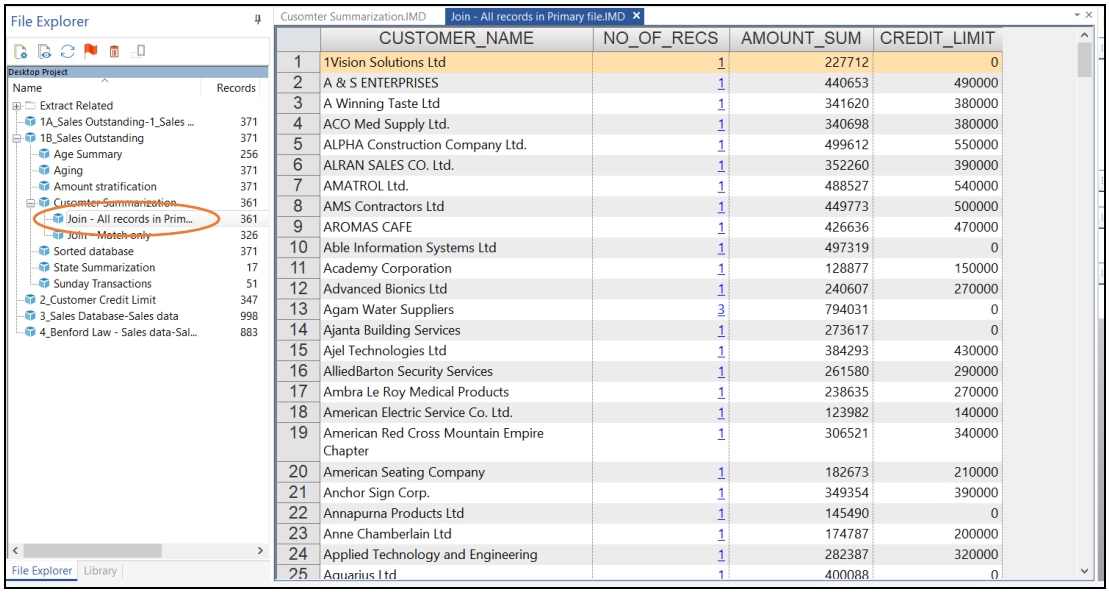
We click in Join option and click on option “All records in Primary file”.



As we want all records in primary, from the secondary database, we are not selecting Customer name in output.



The output file as follows:



	CUSTOMER_NAME	NO_OF_RECS	AMOUNT_SUM	CREDIT_LIMIT
1	1Vision Solutions Ltd	1	227712	0
2	A & S ENTERPRISES	1	440653	490000
3	A Winning Taste Ltd	1	341620	380000
4	ACO Med Supply Ltd.	1	340698	380000
5	ALPHA Construction Company Ltd.	1	499612	550000
6	ALRAN SALES CO. Ltd.	1	352260	390000
7	AMATROL Ltd.	1	488527	540000
8	AMS Contractors Ltd	1	449773	500000
9	AROMAS CAFE	1	426636	470000
10	Able Information Systems Ltd	1	497319	0
11	Academy Corporation	1	128877	150000
12	Advanced Bionics Ltd	1	240607	270000
13	Agam Water Suppliers	3	794031	0
14	Ajanta Building Services	1	273617	0
15	Ajel Technologies Ltd	1	384293	430000
16	AlliedBarton Security Services	1	261580	290000
17	Ambra Le Roy Medical Products	1	238635	270000
18	American Electric Service Co. Ltd.	1	123982	140000
19	American Red Cross Mountain Empire Chapter	1	306521	340000
20	American Seating Company	1	182673	210000
21	Anchor Sign Corp.	1	349354	390000
22	Annapurna Products Ltd	1	145490	0
23	Anne Chamberlain Ltd	1	174787	200000
24	Applied Technology and Engineering	1	282387	320000
25	Aquarius Ltd	1	400088	0

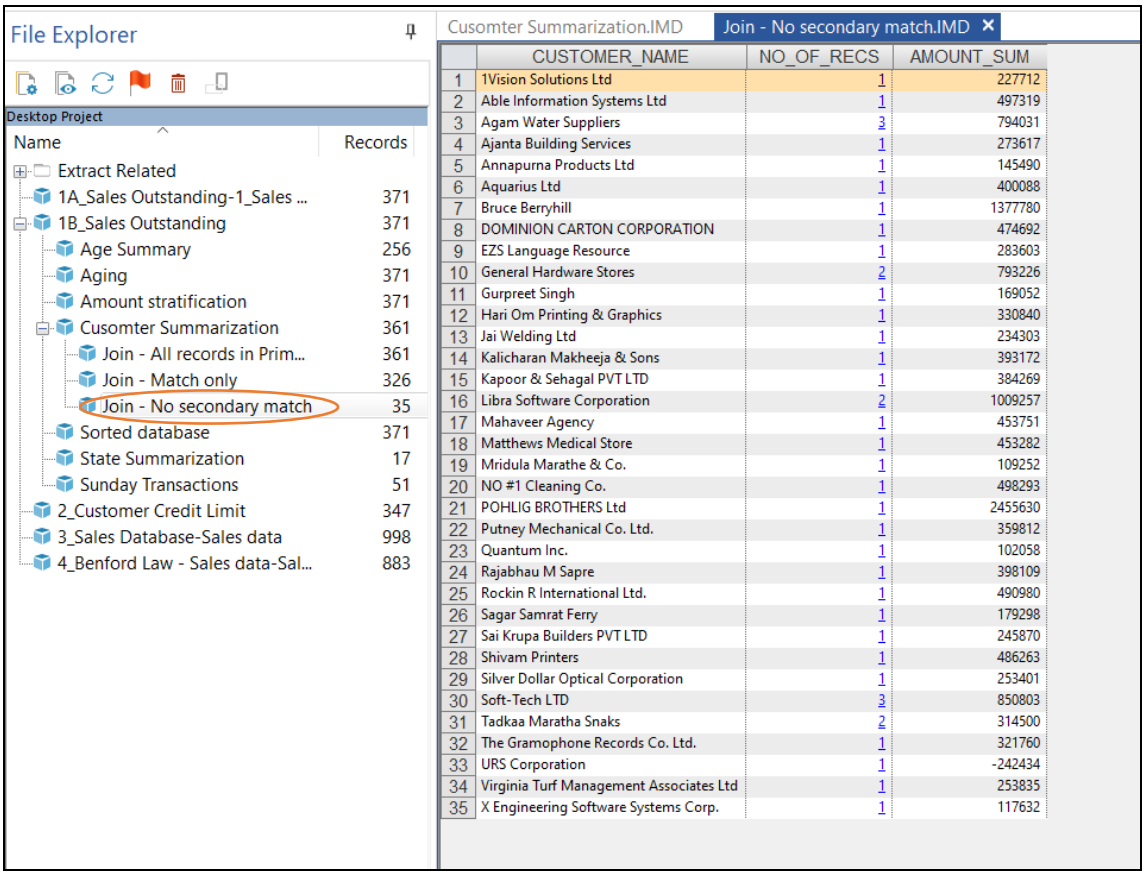
It can be seen that the output file contains 361 records i.e. all records from Primary database. The last column (Credit Limit) is from Secondary database. This column has some values as well as 0 for names not found in secondary database.

Example 3: Records with no secondary match

We click in Join option and click on option "Records with no secondary match" and write the file name as "Join – No Secondary match".

The output file has 35 records i.e. the records not found in Secondary database.

- Primary database – 361 records
- Match only database – 326 records
- Records missing in Secondary database – 35 records (361– 326)



	CUSTOMER_NAME	NO_OF_RECS	AMOUNT_SUM
1	1Vision Solutions Ltd	1	227712
2	Able Information Systems Ltd	1	497319
3	Agam Water Suppliers	3	794031
4	Ajanta Building Services	1	273617
5	Annapurna Products Ltd	1	145490
6	Aquarius Ltd	1	400088
7	Bruce Berryhill	1	1377780
8	DOMINION CARTON CORPORATION	1	474692
9	EZS Language Resource	1	283603
10	General Hardware Stores	2	793226
11	Gurpreet Singh	1	169052
12	Hari Om Printing & Graphics	1	330840
13	Jai Welding Ltd	1	234303
14	Kalicharan Makheeja & Sons	1	393172
15	Kapoor & Sehagal PVT LTD	1	384269
16	Libra Software Corporation	2	1009257
17	Mahaveer Agency	1	453751
18	Matthews Medical Store	1	453282
19	Mridula Marathe & Co.	1	109252
20	NO #1 Cleaning Co.	1	498293
21	POHLIG BROTHERS Ltd	1	2455630
22	Putney Mechanical Co. Ltd.	1	359812
23	Quantum Inc.	1	102058
24	Rajabhau M Sapre	1	398109
25	Rockin R International Ltd.	1	490980
26	Sagar Samrat Ferry	1	179298
27	Sai Krupa Builders PVT LTD	1	245870
28	Shivam Printers	1	486263
29	Silver Dollar Optical Corporation	1	253401
30	Soft-Tech LTD	3	850803
31	Tadkaa Maratha Snaks	2	314500
32	The Gramophone Records Co. Ltd.	1	321760
33	URS Corporation	1	-242434
34	Virginia Turf Management Associates Ltd	1	253835
35	X Engineering Software Systems Corp.	1	117632

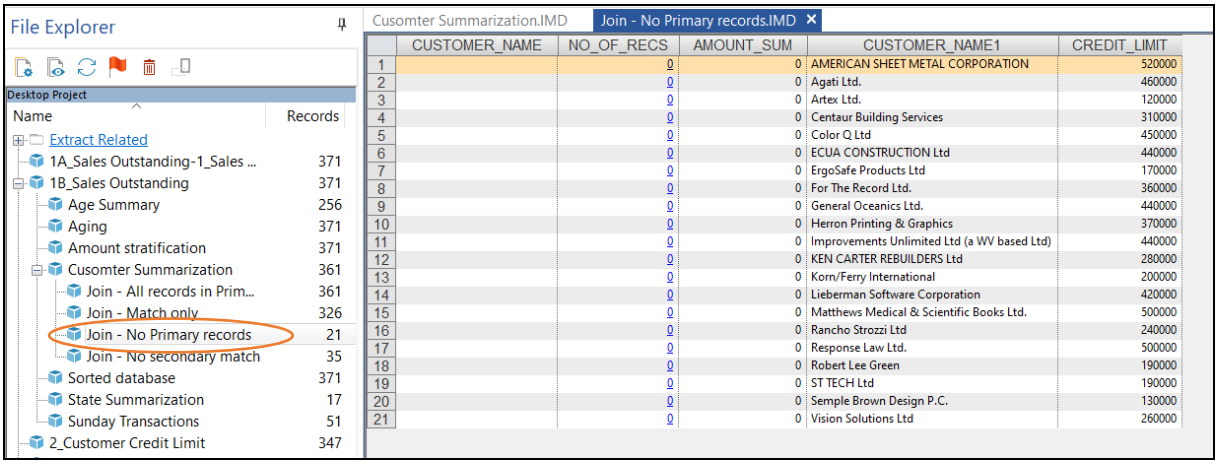
Note: As no records from secondary database, hence no column from secondary database in output file.

Example 4: Records with no primary match

We click in Join option and click on option "Records with no primary match" and write the file name as "Join – No Primary match".

The output file has 21 records i.e. the records not found in Primary database.

- Secondary database – 347 records
- Match only database – 326 records
- Records not found in Primary database – 21 records (347 – 326)



	CUSTOMER_NAME	NO_OF_RECS	AMOUNT_SUM	CUSTOMER_NAME1	CREDIT_LIMIT
1		0	0	AMERICAN SHEET METAL CORPORATION	520000
2		0	0	Agati Ltd.	460000
3		0	0	Artex Ltd.	120000
4		0	0	Centaur Building Services	310000
5		0	0	Color Q Ltd	450000
6		0	0	ECUA CONSTRUCTION Ltd	440000
7		0	0	ErgoSafe Products Ltd	170000
8		0	0	For The Record Ltd.	360000
9		0	0	General Oceanics Ltd.	440000
10		0	0	Herron Printing & Graphics	370000
11		0	0	Improvements Unlimited Ltd (a WV based Ltd)	440000
12		0	0	KEN CARTER REBUILDERS Ltd	280000
13		0	0	Korn/Ferry International	200000
14		0	0	Lieberman Software Corporation	420000
15		0	0	Matthews Medical & Scientific Books Ltd.	500000
16		0	0	Rancho Strozzi Ltd	240000
17		0	0	Response Law Ltd.	500000
18		0	0	Robert Lee Green	190000
19		0	0	ST TECH Ltd	190000
20		0	0	Semple Brown Design P.C.	130000
21		0	0	Vision Solutions Ltd	260000

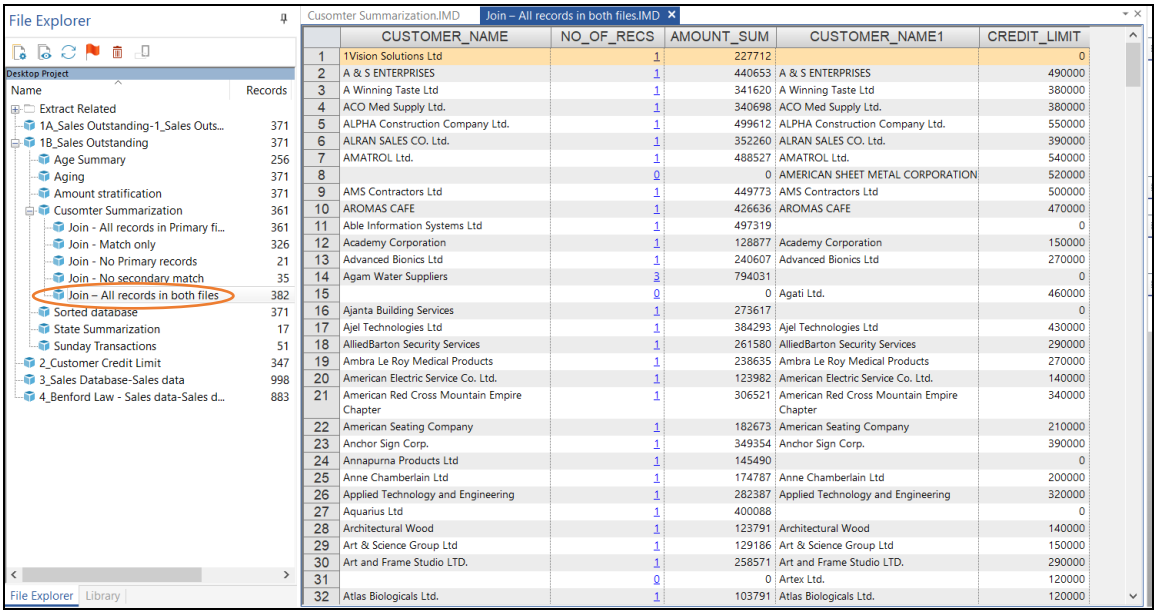
Note: As there are no matching names in Primary database, the first 3 columns are blank and last 2 columns are from Secondary database.

Example 5: All records in both files

We click in Join option and click on option "All records in both files" and write the file name as "Join – All records in both files".

The output file has 382 records i.e. all the records from Primary database and Secondary database.

- Primary database – 361 records
- Secondary database – 347 records
- Records not found in Primary database – 21 records
- All records count – 382 records (361 + 21)



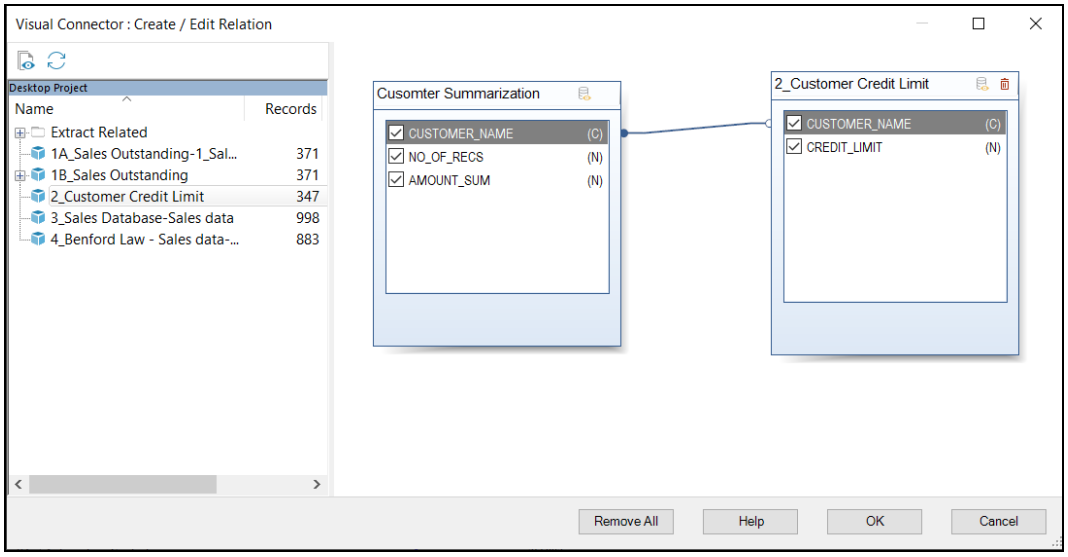
	CUSTOMER_NAME	NO_OF_RECS	AMOUNT_SUM	CUSTOMER_NAME1	CREDIT_LIMIT
1	1Vision Solutions Ltd	1	227712		0
2	A & S ENTERPRISES	1	440653	A & S ENTERPRISES	490000
3	A Winning Taste Ltd	1	341620	A Winning Taste Ltd	380000
4	ACO Med Supply Ltd.	1	340698	ACO Med Supply Ltd.	380000
5	ALPHA Construction Company Ltd.	1	499612	ALPHA Construction Company Ltd.	550000
6	ALRAN SALES CO. Ltd.	1	352260	ALRAN SALES CO. Ltd.	390000
7	AMATROL Ltd.	1	488527	AMATROL Ltd.	540000
8		0	0	AMERICAN SHEET METAL CORPORATION	520000
9	AMS Contractors Ltd	1	449773	AMS Contractors Ltd	500000
10	AROMAS CAFE	1	426636	AROMAS CAFE	470000
11	Able Information Systems Ltd	1	497319		0
12	Academy Corporation	1	128877	Academy Corporation	150000
13	Advanced Bionics Ltd	1	240607	Advanced Bionics Ltd	270000
14	Agam Water Suppliers	3	794031		0
15		0	0	Agati Ltd.	460000
16	Ajanta Building Services	1	273617		0
17	Ajel Technologies Ltd	1	384293	Ajel Technologies Ltd	430000
18	AlliedBarton Security Services	1	261580	AlliedBarton Security Services	290000
19	Ambra Le Roy Medical Products	1	238635	Ambra Le Roy Medical Products	270000
20	American Electric Service Co. Ltd.	1	123982	American Electric Service Co. Ltd.	140000
21	American Red Cross Mountain Empire Chapter	1	306521	American Red Cross Mountain Empire Chapter	340000
22	American Seating Company	1	182673	American Seating Company	210000
23	Anchor Sign Corp.	1	349354	Anchor Sign Corp.	390000
24	Annapurna Products Ltd	1	145490		0
25	Anne Chamberlain Ltd	1	174787	Anne Chamberlain Ltd	200000
26	Applied Technology and Engineering	1	282387	Applied Technology and Engineering	320000
27	Aquarius Ltd	1	400088		0
28	Architectural Wood	1	123791	Architectural Wood	140000
29	Art & Science Group Ltd	1	129186	Art & Science Group Ltd	150000
30	Art and Frame Studio LTD.	1	258571	Art and Frame Studio LTD.	290000
31		0	0	Artex Ltd.	120000
32	Atlas Biologicals Ltd.	1	103791	Atlas Biologicals Ltd.	120000

Note: The output file contains all the records from Primary database as well as Secondary database. Wherever it is not matching, the value is shown as blank in Name column and 0 in Amount / Credit Limit.

- **Visual Connector**

This feature is available under **Analysis Tab > Relate >Visual Connector**. We use the Visual Connector task to create a single database from two or more other databases. One of these databases must be the primary database. After we define the databases, we need to define the fields to include in the output data base.

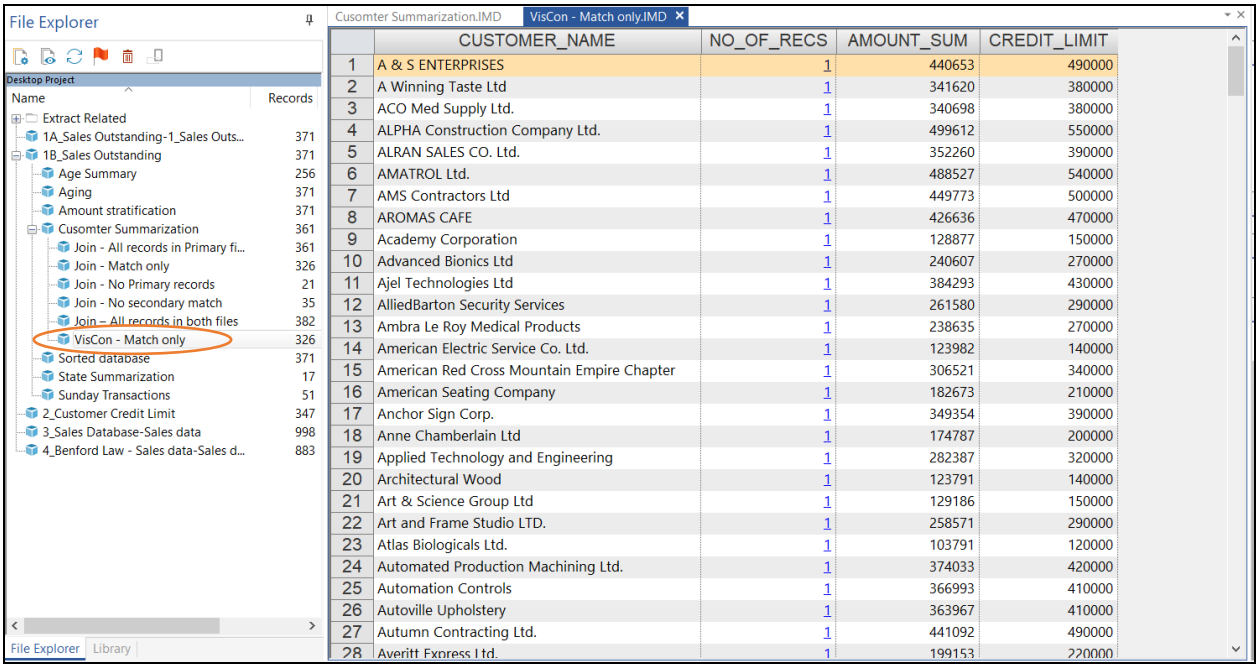
In the below example, first we open the Customer Summarisation (Primary database). Then click on Visual Connector. Then we select the secondary database and define the common field in both file/database i.e. Customer Name. We need to connect common field in both databases. This is similar to MS Access.



When we click on OK, it shows 2 options to select

- Matches only
- All records in the primary database, all matches

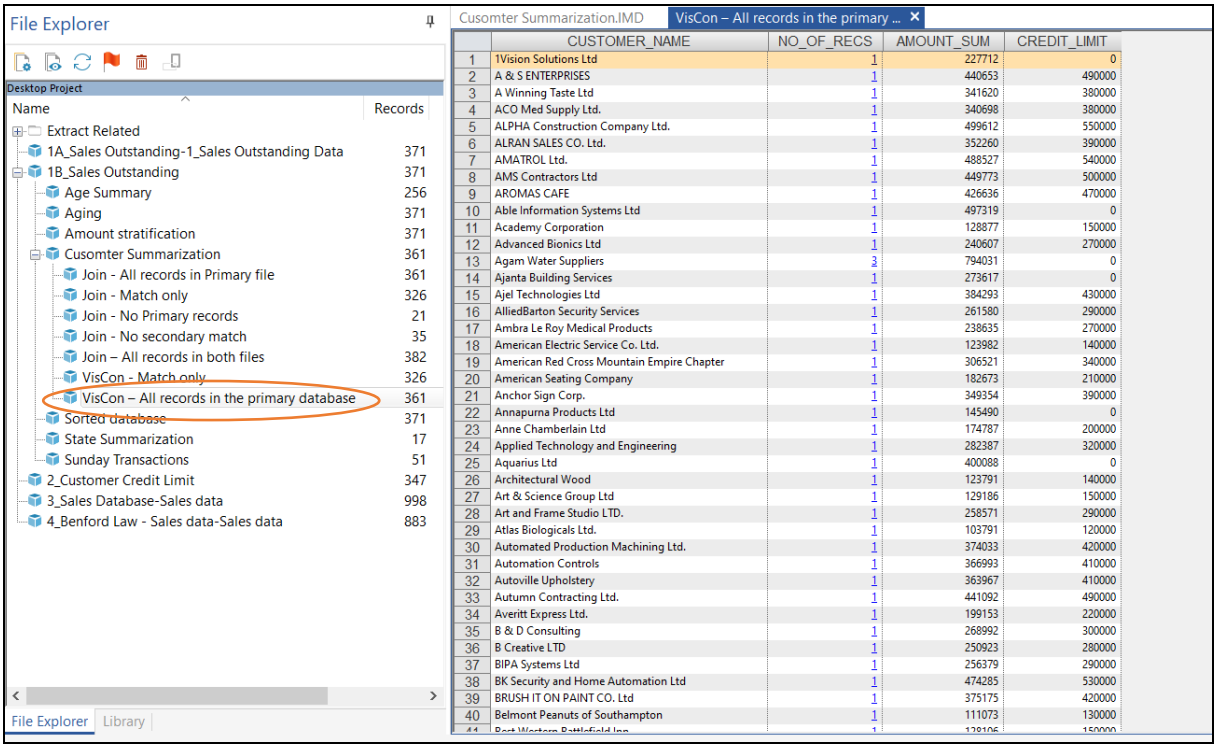
Example 1: The output file (VisCon – match only) is as follows:



	CUSTOMER_NAME	NO_OF_RECS	AMOUNT_SUM	CREDIT_LIMIT
1	A & S ENTERPRISES	1	440653	490000
2	A Winning Taste Ltd	1	341620	380000
3	ACO Med Supply Ltd.	1	340698	380000
4	ALPHA Construction Company Ltd.	1	499612	550000
5	ALRAN SALES CO. Ltd.	1	352260	390000
6	AMATROL Ltd.	1	488527	540000
7	AMS Contractors Ltd	1	449773	500000
8	AROMAS CAFE	1	426636	470000
9	Academy Corporation	1	128877	150000
10	Advanced Bionics Ltd	1	240607	270000
11	Ajel Technologies Ltd	1	384293	430000
12	AlliedBarton Security Services	1	261580	290000
13	Ambra Le Roy Medical Products	1	238635	270000
14	American Electric Service Co. Ltd.	1	123982	140000
15	American Red Cross Mountain Empire Chapter	1	306521	340000
16	American Seating Company	1	182673	210000
17	Anchor Sign Corp.	1	349354	390000
18	Anne Chamberlain Ltd	1	174787	200000
19	Applied Technology and Engineering	1	282387	320000
20	Architectural Wood	1	123791	140000
21	Art & Science Group Ltd	1	129186	150000
22	Art and Frame Studio LTD.	1	258571	290000
23	Atlas Biologicals Ltd.	1	103791	120000
24	Automated Production Machining Ltd.	1	374033	420000
25	Automation Controls	1	366993	410000
26	Autoville Upholstery	1	363967	410000
27	Autumn Contracting Ltd.	1	441092	490000
28	Averitt Express Ltd.	1	199153	220000

The Output file has 326 records, which is matching in both files. But here it has not shown Customer Name from Secondary database. As we have Amount Outstanding and Credit Limit side by side, we can compare the data.

Example 2: In this example, we select second option i.e. “All records in the primary database”. The output file (VisCon – All records in the primary database) is as follows:



It can be seen that the output file contains 361 records i.e. all records from Primary database. The last column (Credit Limit) is from Secondary database. This column has some values as well as 0 for names not found in secondary database.

- **Compare**

This feature is available under **Analysis Tab > Relate >Compare**.This feature helps to compare values from both the databases and gives the difference as an additional column by comparing Total field in both databases.

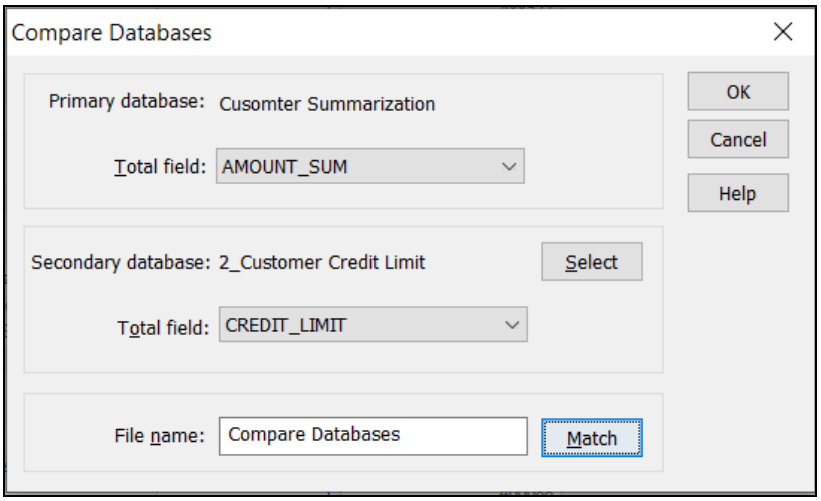
The inputs are:

— Primary database: Customer Summarisation
  - Total Field: AMOUNT_SUM

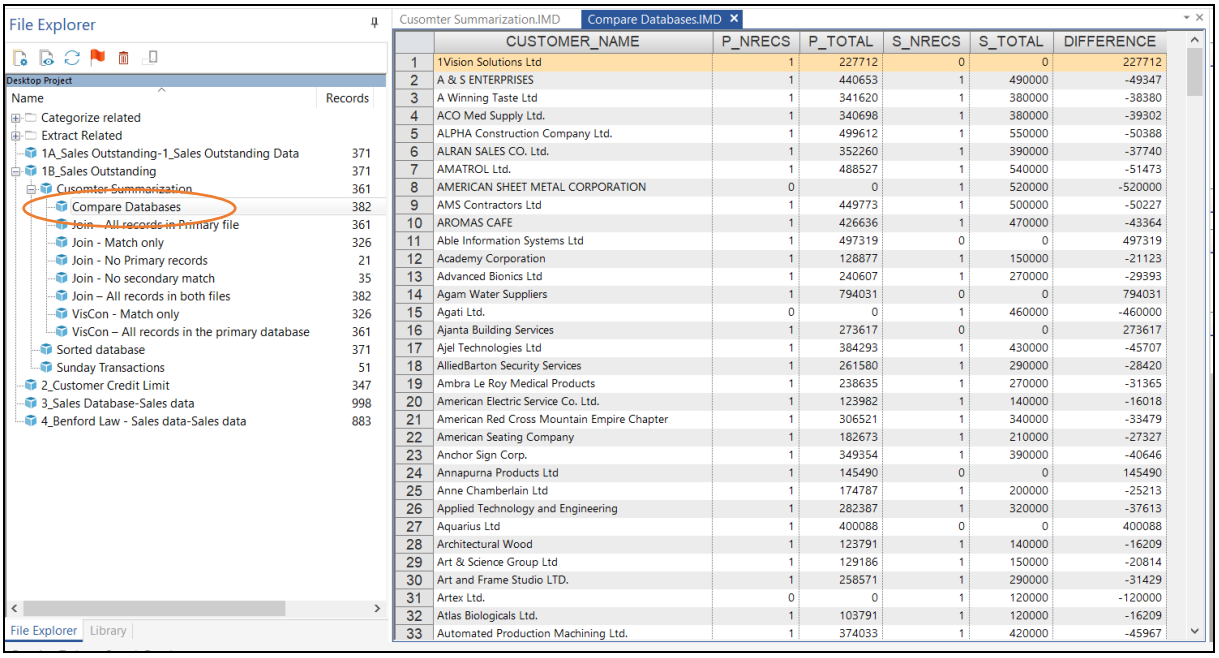
— Secondary database: Customer Credit Limit
  - Total Field: CREDIT_LIMIT

— Match field: Customer Name (common in both databases)

— File Name: Compare Databases



The output file is as follows:



	CUSTOMER_NAME	P_NRECS	P_TOTAL	S_NRECS	S_TOTAL	DIFFERENCE
1	1Vision Solutions Ltd	1	227712	0	0	227712
2	A & S ENTERPRISES	1	440653	1	490000	-49347
3	A Winning Taste Ltd	1	341620	1	380000	-38380
4	ACO Med Supply Ltd.	1	340698	1	380000	-39302
5	ALPHA Construction Company Ltd.	1	499612	1	550000	-50388
6	ALRAN SALES CO. Ltd.	1	352260	1	390000	-37740
7	AMATROL Ltd.	1	488527	1	540000	-51473
8	AMERICAN SHEET METAL CORPORATION	0	0	1	520000	-520000
9	AMS Contractors Ltd	1	449773	1	500000	-50227
10	AROMAS CAFE	1	426636	1	470000	-43364
11	Able Information Systems Ltd	1	497319	0	0	497319
12	Academy Corporation	1	128877	1	150000	-21123
13	Advanced Bionics Ltd	1	240607	1	270000	-29393
14	Agam Water Suppliers	1	794031	0	0	794031
15	Agati Ltd.	0	0	1	460000	-460000
16	Ajanta Building Services	1	273617	0	0	273617
17	Ajel Technologies Ltd	1	384293	1	430000	-45707
18	AlliedBarton Security Services	1	261580	1	290000	-28420
19	Ambra Le Roy Medical Products	1	238635	1	270000	-31365
20	American Electric Service Co. Ltd.	1	123982	1	140000	-16018
21	American Red Cross Mountain Empire Chapter	1	306521	1	340000	-33479
22	American Seating Company	1	182673	1	210000	-27327
23	Anchor Sign Corp.	1	349354	1	390000	-40646
24	Annapurna Products Ltd	1	145490	0	0	145490
25	Anne Chamberlain Ltd	1	174787	1	200000	-25213
26	Applied Technology and Engineering	1	282387	1	320000	-37613
27	Aquarius Ltd	1	400088	0	0	400088
28	Architectural Wood	1	123791	1	140000	-16209
29	Art & Science Group Ltd	1	129186	1	150000	-20814
30	Art and Frame Studio LTD.	1	258571	1	290000	-31429
31	Artex Ltd.	0	0	1	120000	-120000
32	Atlas Biologicals Ltd.	1	103791	1	120000	-16209
33	Automated Production Machining Ltd.	1	374033	1	420000	-45967

The output file has 382 records i.e. all the records from Primary database and Secondary database.

- Primary database – 361 records
- Secondary database – 347 records
- Records not found in Primary database – 21 records
- All records count – 382 records (361 + 21)

The output file contains all the records from Primary database as well as Secondary database. Wherever it is not matching, the value is shown as blank in Name column and 0 in Amount / Credit Limit.

The positive value in Difference column (Amount_Sum – Credit_Limit) shows that Outstanding Amount is more that the Credit Limit to that customer.

The negative value in Difference column (Amount_Sum – Credit_Limit) shows that Credit Limit is more than Outstanding Amount for that customer.

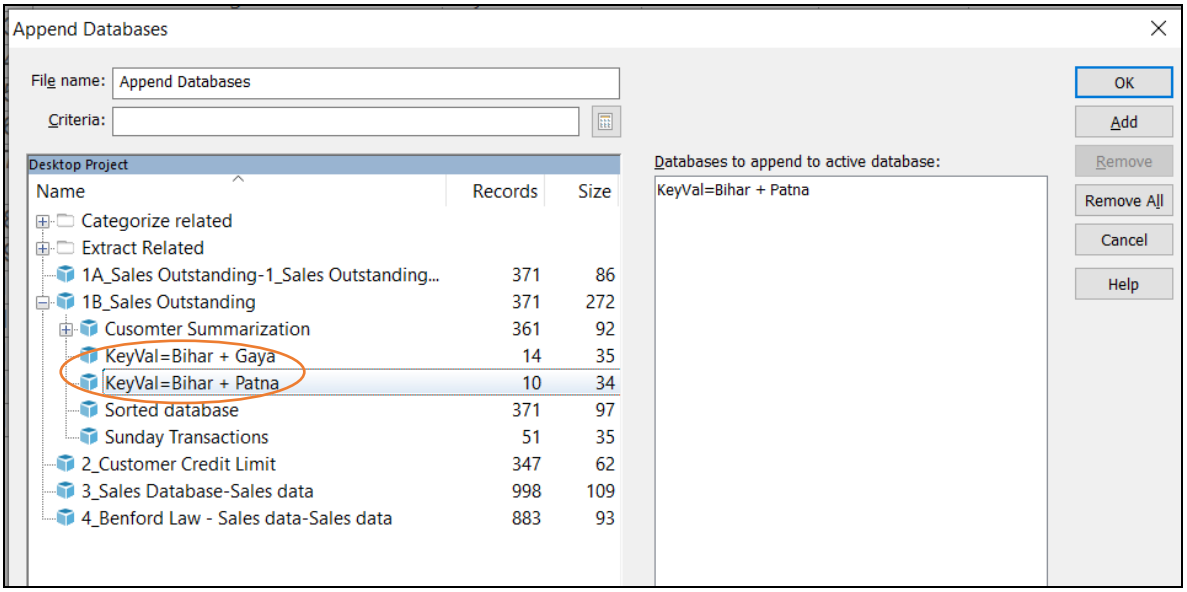
Both the scenarios need further review, especially the first one (positive values).

- **Append**

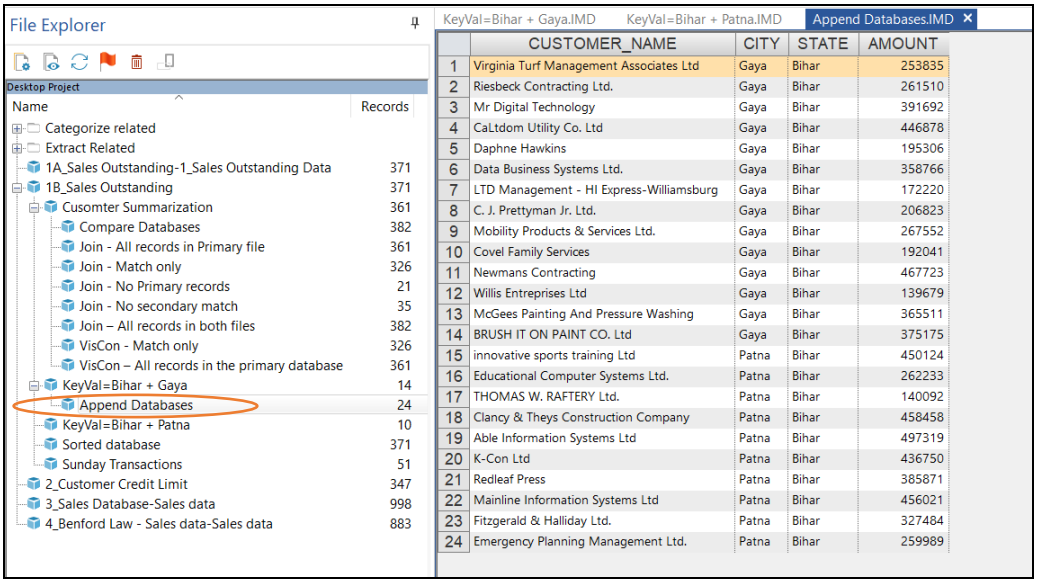
This feature is available under **Analysis Tab > Relate >Append**.This feature concatenate two or more IDEA databases into a new database.

Example: We have 2 databases i.e. "KeyVal=Bihar + Gaya" and "KeyVal=Bihar + Patna". First we need to open One database. We opened both the files and active file is "KeyVal=Bihar + Gaya"

We click on Append option. Then we need to double click on the second database ("KeyVal=Bihar + Patna") which we want to append in first file. The File Name is "Append Database".



The output file is as follows:



	CUSTOMER_NAME	CITY	STATE	AMOUNT
1	Virginia Turf Management Associates Ltd	Gaya	Bihar	253835
2	Riesbeck Contracting Ltd.	Gaya	Bihar	261510
3	Mr Digital Technology	Gaya	Bihar	391692
4	CaLtdom Utility Co. Ltd	Gaya	Bihar	446878
5	Daphne Hawkins	Gaya	Bihar	195306
6	Data Business Systems Ltd.	Gaya	Bihar	358766
7	LTD Management - HI Express-Williamsburg	Gaya	Bihar	172220
8	C. J. Prettyman Jr. Ltd.	Gaya	Bihar	206823
9	Mobility Products & Services Ltd.	Gaya	Bihar	267552
10	Covel Family Services	Gaya	Bihar	192041
11	Newmans Contracting	Gaya	Bihar	467723
12	Willis Entreprises Ltd	Gaya	Bihar	139679
13	McGees Painting And Pressure Washing	Gaya	Bihar	365511
14	BRUSH IT ON PAINT CO. Ltd	Gaya	Bihar	375175
15	innovative sports training Ltd	Patna	Bihar	450124
16	Educational Computer Systems Ltd.	Patna	Bihar	262233
17	THOMAS W. RAFTERY Ltd.	Patna	Bihar	140092
18	Clancy & Theys Construction Company	Patna	Bihar	458458
19	Able Information Systems Ltd	Patna	Bihar	497319
20	K-Con Ltd	Patna	Bihar	436750
21	Redleaf Press	Patna	Bihar	385871
22	Mainline Information Systems Ltd	Patna	Bihar	456021
23	Fitzgerald & Halliday Ltd.	Patna	Bihar	327484
24	Emergency Planning Management Ltd.	Patna	Bihar	259989

In the output file, we can see the second file is appended below the first file.

10.8 SAMPLE SELECTION USING IDEA

As a tool, IDEA has robust and stable capabilities to electronically audit 100% data in any application and database.

But CAATs do not replace traditional audit checks in terms of observation, inquiry, walk-through, substantive testing, and corroboration.

Because of this, there is need for sampling. Substantive testing for instance requires tests of details of transactions and balances. This involves specific manual activities which even though can be checked fully would be time consuming. It is activities like substantive testing that need sampling within IDEA. Public Accounting firms of international repute use IDEA primarily for sampling in their Financial Audits.

Sampling in IDEA is broadly statistical and probability-based.

The probability-based sampling techniques are:-

- Systematic
- Random
- Stratified Random

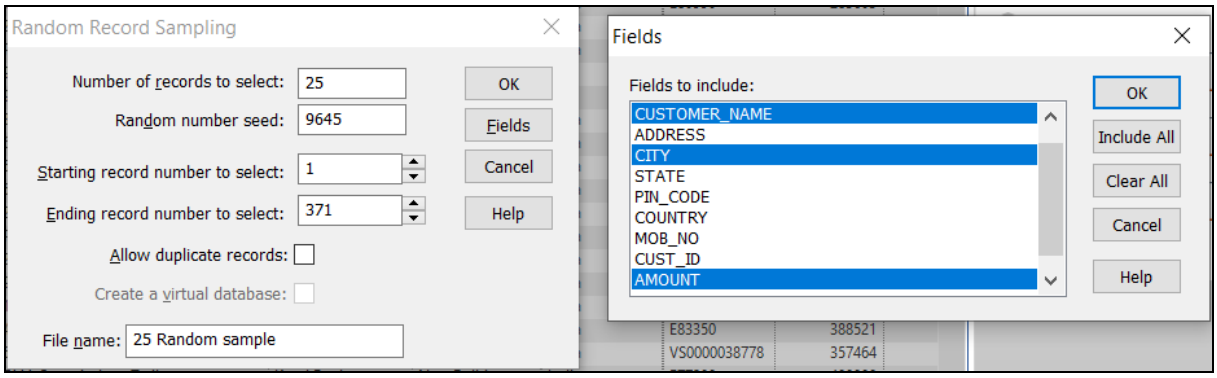
The statistical sampling techniques are:-

- Attribute
- Classical Variable
- Monetary Unit

- **Random sample:** This option is available in **Analysis Tab > Sample > Random**. We have Customer Sales Outstanding database (1B_Sales Outstanding). From this database, we want to select 25 random samples. The process is as follows:
  - On the Analysis tab, in the Sample group, click Random.
  - In the Number of records to select field, enter 25.
  - Accept the Random number seed value provided by IDEA.

    (IDEA uses the random number seed to start the algorithm for calculating the random numbers. If a sample needs to be extended, then entering the same random number seed but with a larger sample size produces the same original selection plus the required additional records.)
  - Accept the default values in the Starting record number to select and the Ending record number to select fields.

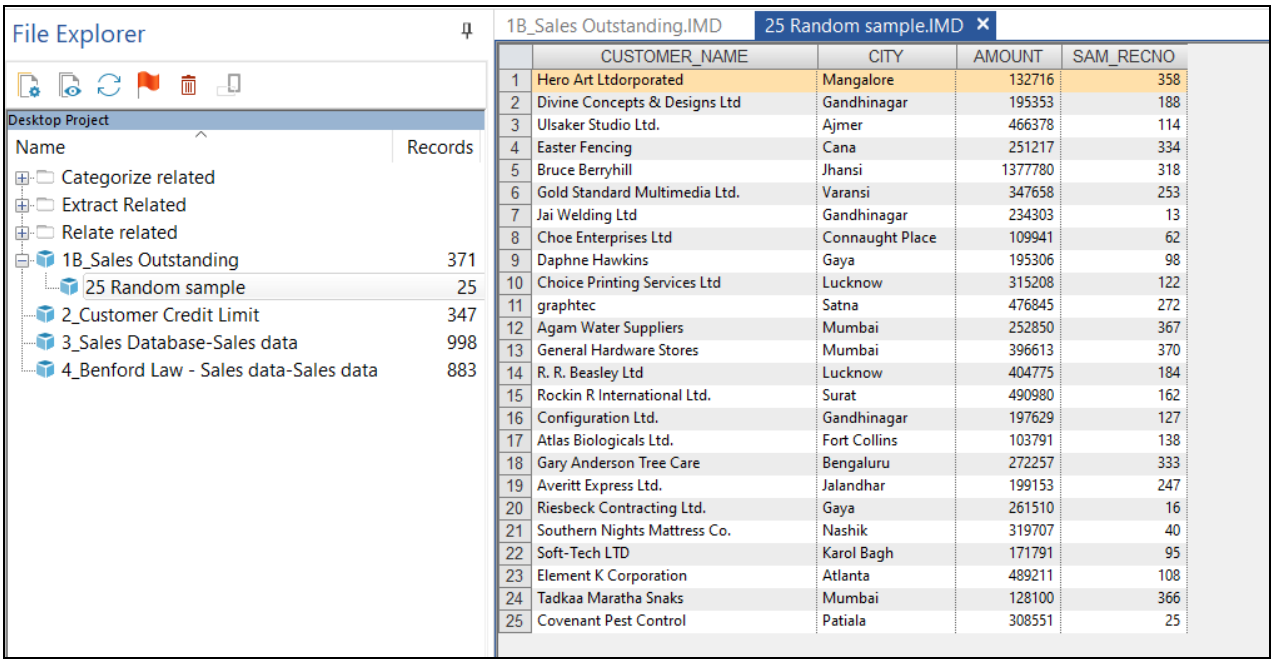
    (IDEA sets the defaults as the first and last records; in this case 1 and371.)
  - Leave the Allow duplicate records check box unselected.
  - In the File name field, enter "25 Random sample " as new file name.



  - Click Fields.

    (The Fields dialog box appears. Note that by default, IDEA selects all fields from the database to extract to the Sample of Customers database. Here, we are selecting 3 columns i.e. Customer Name, City and Amount)
  - Click OK.
  - In the Random Record Sampling dialog box, click OK.
  - View the output database and note the additional field (SAM_RECNO)that IDEA has added to the database as the right-most column. This contains the corresponding record numbers from the original database(1B_Sales Outstanding).

    Note that the database appears as a child database to its parent database (1B_Sales Outstanding) in the File Explorer window.



	CUSTOMER_NAME	CITY	AMOUNT	SAM_RECNO
1	Hero Art Ltdorporated	Mangalore	132716	358
2	Divine Concepts & Designs Ltd	Gandhinagar	195353	188
3	Ulsaker Studio Ltd.	Ajmer	466378	114
4	Easter Fencing	Cana	251217	334
5	Bruce Berryhill	Jhansi	1377780	318
6	Gold Standard Multimedia Ltd.	Varansi	347658	253
7	Jai Welding Ltd	Gandhinagar	234303	13
8	Choe Enterprises Ltd	Connaught Place	109941	62
9	Daphne Hawkins	Gaya	195306	98
10	Choice Printing Services Ltd	Lucknow	315208	122
11	graphtec	Satna	476845	272
12	Agam Water Suppliers	Mumbai	252850	367
13	General Hardware Stores	Mumbai	396613	370
14	R. R. Beasley Ltd	Lucknow	404775	184
15	Rockin R International Ltd.	Surat	490980	162
16	Configuration Ltd.	Gandhinagar	197629	127
17	Atlas Biologicals Ltd.	Fort Collins	103791	138
18	Gary Anderson Tree Care	Bengaluru	272257	333
19	Averitt Express Ltd.	Jalandhar	199153	247
20	Riesbeck Contracting Ltd.	Gaya	261510	16
21	Southern Nights Mattress Co.	Nashik	319707	40
22	Soft-Tech LTD	Karol Bagh	171791	95
23	Element K Corporation	Atlanta	489211	108
24	Tadkaa Maratha Snaks	Mumbai	128100	366
25	Covenant Pest Control	Patiala	308551	25

- To close the Sample of Customers database, right-click the Sample of Customers tab and select Close.

- **Systematic Sampling**: This feature is available under **Analysis Tab > Sample > Other > Systematic.**It is a method to extract a number of records from a database at equal intervals to a separate database. It is often referred to as interval sampling.

  There are two methods of determining the sample:

  - Entering the number of records, in which case IDEA computes the interval size.
  - Entering the selection interval, in which case IDEA computes the number of records.

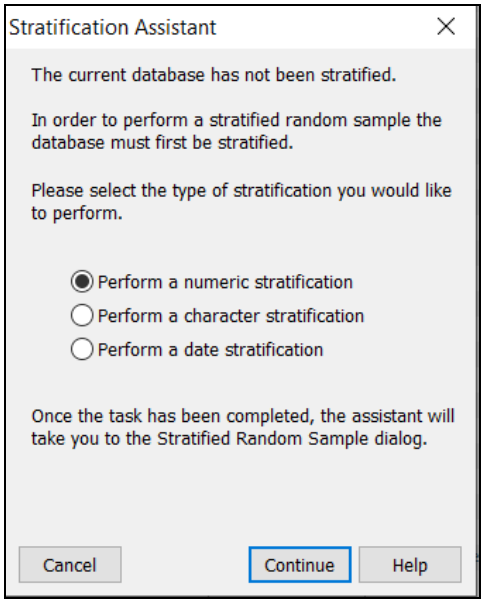
  IDEA calculates the above parameters on the number of records in the database and defaults to the first to last records. However, we can extract the sample from a range of records, if required.

  Note: Systematic sampling can be used where there is a need to cover samples evenly across the whole population, like in Payroll Audits. In Payroll activities the number of employees stays fairly uniform and even over a financial year. Here systematic sampling can be used for each quarter to cover different lots and intervals of Employees. Herein opinion can be expressed on a fairly representative lot of employees from the total employee base.

- **Stratified Random Sampling:** This feature is available under **Analysis Tab > Sample > Other >Stratified Random.** Stratified Random Record Sampling is based on the materiality of transactions. Here due emphasis is given to the 80-20 rule in business, wherein 80% of the value resides in 20% quantum and 20% value in 80% quantum. The 80% value representing the Class A is taken up for detailed scrutiny. The 20% value representing the Class C is taken up for limited random selection and scrutiny based on either rule of thumb or existing techniques based on business and data conditions.

Stratified Random Record Sampling maybe be used to extract a random sample with a specified number of records from each of a series of bands. This task requires then the database be first stratified into a series of numeric, character or date bands. IDEA then presents us with a table displaying the number of records within each band. We then enter the number of sample records we want IDEA to extract at random from each band.

IDEA provides a Stratification Assistant to guide us through the steps for stratifying and generating the sample.



It is not necessary to stratify the fie first, although if this has been done and a stratified database used, then the Assistant skips the initial stages and goes directly to the sample selection stage.

To ensure that we extract at least one transaction of each type in a random sample, we have to use Stratified Random Sampling: to stratify the database by the type field and then select at least one record from each band.

Stratified Random Sampling can be used in General Ledger Sampling as different values and risks are associated with different ledger items.

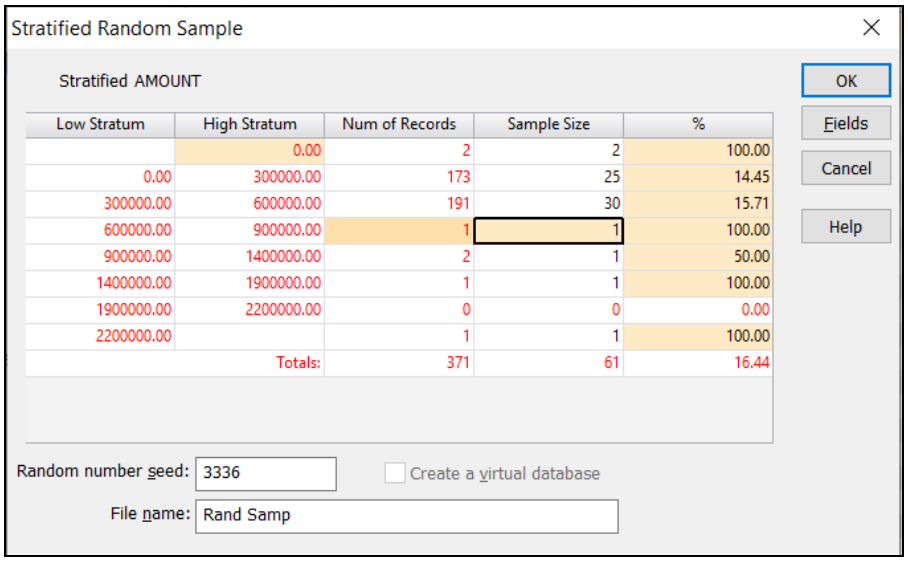
Example 1: Random sample from Numeric Stratification

When we select "Perform a numeric stratification", it open Stratification dialogue box, we are going to provide following inputs:

- Field to stratify: Amount
- Fields to total on: Amount
- Lower Limit: 0 (by default it picks lowest amount)
- Increment: first 3 upper limit – 300000, next 2 upper limit – 500000, last upper limit - 300000

Then we click OK.

It will open another dialogue box, which contains stratified Amount. Now we need to manually enter the number of samples we need in “Sample Size” column. Then select the fields. We are going ahead will all the fields.



In the above image, we can see that we have entered various numbers in Sample Size column as samples. Then we click on OK button.

New file is naming “Rand Samp” is created as a child of Stratification file in File Explorer, which contains 61 records as sample.

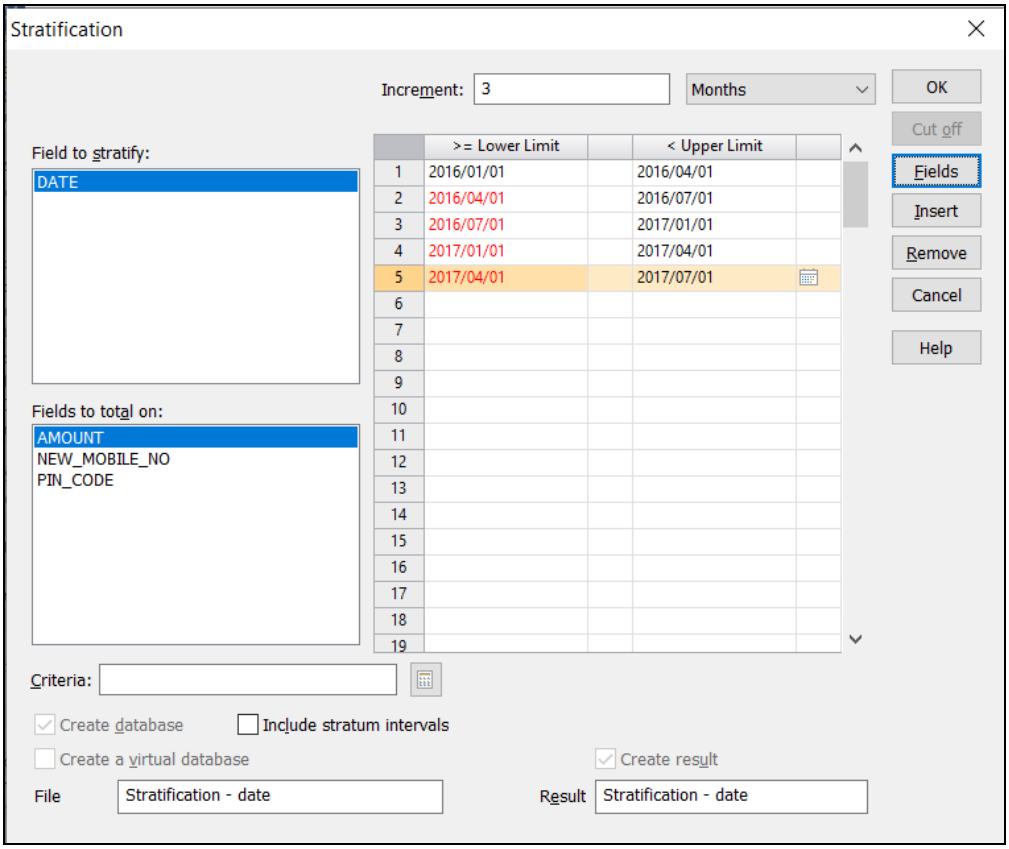
The second last column (STRATUM) contains stratum reference number from where the sample is picked. We can manually count the sample count from each stratum to cross check the output with input number given in Sample Size column.

The column "SAM_RECNO" contains the reference number of original data.

Example 2: Random sample from Date Stratification

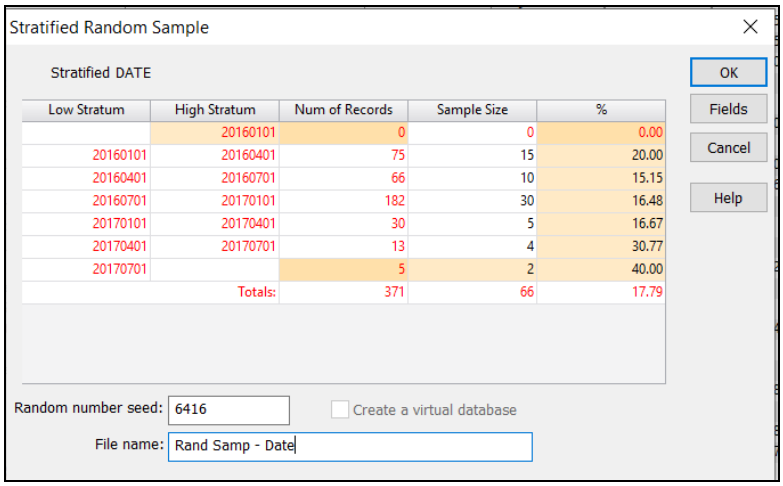
When we select "Perform a date stratification", it open Stratification dialogue box, we are going to provide following inputs:

- Field to stratify: Date
- Fields to total on: Amount
- Lower Limit: 2016/01/01 (by default it picks lowest date)
- Increment: first 3 upper limit – 3 months, next 1 upper limit – 6 month, last 3 upper limit – 3 month
- Fields: Date, Customer Name and Amount



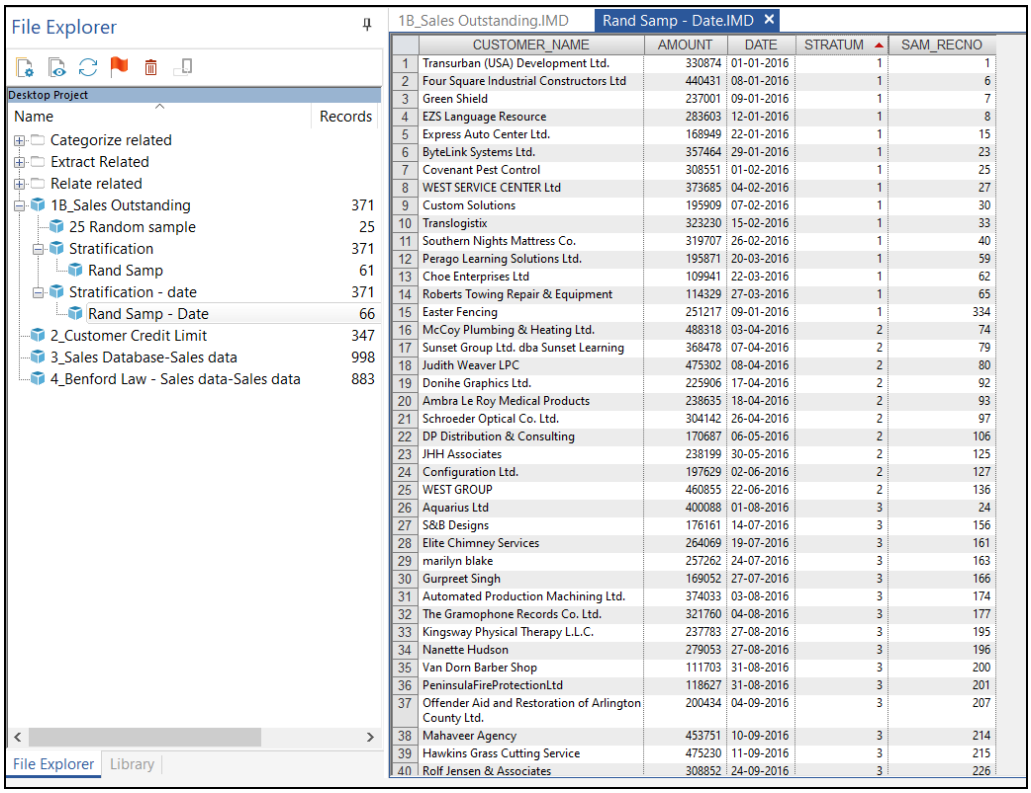
Then we click OK.

It will open another dialogue box, which contains stratified Amount. Now we need to manually enter the number of samples we need in "Sample Size" column.



In the above image, we can see that we have entered various numbers in Sample Size column as samples. Then we click on OK button.

New file is naming “Rand Samp - Date” is created as a child of “Stratification – date” file in File Explorer, which contains 66 records as sample.



The second last column (STRATUM) contains stratum reference number from where the sample is picked. We can manually count the sample count from each stratum to cross check the output with input number given in Sample Size column.

The column "SAM_RECNO" contains the reference number of original data.

- **Attribute Sampling:** This feature is available under **Analysis Tab > Sample >Attribute.** Attribute Sampling refers to the examination of a subset of a population (a sample) in order to assess how frequently a particular event or attribute occurs in the population as a whole. An attribute has only two possible values: true or false. In auditing, typical attributes are whether an item is in error or not, whether a particular control has been exercised or not, or whether the entity complied with a particular law or not.

  The Attribute Planning and Evaluation task is meat to plan attribute samples and evaluate results. The Planning dialog boxes are used to determine the minimum sample size required to meet specified audit objectives, as well as the critical or maximum allowable number of sample deviations in the sample.

  The two Planning options, Beta Risk Control and Beta and Alpha Risk Control, allow us to control the risk of one or two types of incorrect decisions. Some auditors do not set the risk of not relying on controls when they are effective (Alpha Risk). They set only the risk of relying on controls when they are not effective (Beta Risk) and the tolerable deviation rate. They set an expected deviation rate to partially control for Alpha Risk. To use this approach, we select the Planning (Beta Risk Control) tab in the Attribute Sampling dialog box.

  In some cases, we may want to consider two risks: the risk of relying on controls when they are not effective (Beta Risk) and the risk of relying on controls when they are effective (Alpha Risk). Each risk relates to a rate of deviation (errors) which are necessary while arriving at the sample selection. Beta Risk relates to the tolerable deviation rate; Alpha Risk relates to the lower threshold rate or the acceptable deviation rate.

  For example, we might set the risk of relying on controls when they are not effective (Beta Risk) at 10%, the tolerable deviation rate at 6%, the risk of not relying on controls when they are effective (Alpha Risk) at 20% and the acceptable deviation rate at 1%. IDEA calculates a sample size so that if we find the indicated number of deviations, there is a 10% risk (or 90% confidence level to control Beta Risk) that the population contains a deviation rate greater than 6%; if we find more than the acceptable number of deviations, there is a 20% risk (or 80% confidence level to control Alpha Risk) that the deviation rate is less than 1%.

  If we are not sure of what Planning task is appropriate for our situation, we may need to consult an expert statistician.

  After planning, the sample is obtained and tested, and the number od deviations in the sample determined. If the number of deviations found are less than or equal to the critical number of deviations determined when planning the sample, the test is successful. The sample results can be analysed precisely by using the evaluation task.

  Evaluation of the sample allows us to make an inference, with some degree of confidence, about whether the true rate of deviations in the population exceeds some tolerable amount or not.

IDEA's Attribute Sampling programs are used to plan for and evaluate samples obtained by Random Record Sampling. If the sample was obtained by Stratified Random Sampling, then we do not use this evaluation program; instead, we seek the guidance of a statistician.

- **Monetary Unit Sampling:** This feature is available under **Analysis Tab > Sample >Monetary Unit.**A numeric amount field is the basis of the sample in Monetary Unit Sampling, in which the monetary unit is sampled. Having chosen the monetary units to test on an interval (or random item within each interval) basis, the items which contain these units are identified. To do this, it is necessary to calculate the cumulative value of the relevant field for each item.

Record	Ref	Amount	Cumulative
1	A123	50	50
2	A124	75	125
3	A125	40	165

With an interval of 100 starting at a random start point of 35, the first selection would be at 35, and the second at 135. It is the cumulative field that is used for the selection.

Any items greater than the sampling interval are forced to be chosen by the Auditor for detailed substantive testing. However, these key or high value items have to be extracted from the population and subjected to separate testing, leaving the items below a specified high value amount (which can, if desired, be set to a different limit t the interval) to be evaluated statistically.

IDEA offers options for determining the sample size for both compliance and substantive tests, based on confidence level, population size, materiality and certain options on the weighting statistics used.

The output file contains an Editable field called AUDIT_AMT. if desired, this can be updated from the results of audit testing by using the Field Manipulation task and ten a detailed evaluation of the results can be performed against book value.

The Monetary Unit Sampling technique is used extensively by the Comptroller and Auditor General of India.

- **Classical Variables Sampling:** This feature is available under **Analysis Tab > Sample >Variables.**We use Classical Variables Sampling to provide estimates of quantities that are not recorded in the main database (also known as the population), by projecting results obtained for a sample drawn from the population.

A common use of the Classical Variables Sampling is to estimate the total error and the total corrected, or audited, amount in the population, using an audited sample from the population.

Classical Variables Sampling is also useful for making accounting estimates. For example, we may estimate how many rupees of the inventory, based on a sample for which the portion of each entry that is obsolete has been determined from manual records.

The sampling unit used in Classical Variables Sampling is a single record.The sample database contains the key field from the population database, that is, the Numeric field whose entries are being checked by the audit, as well as an Editable field in which corrected amounts, determine through audit, can be entered.

All other fields from the population database are included in the sample by default; however, it is optional to omit them.

Classical Variables Sampling techniques for projecting errors are most useful for sample auditing of accounts in which there may be errors or misstatements. If we wish to determine limits on the error in the population though we expect no errors or very few errors in the population, and if the principal risk is overstatement, we should consider using IDEA's Monetary Unit Sampling instead of Classical Variables Sampling.

If we do not wish to project errors between the total book value and audited amounts, and require only limits on the total book value and audited amounts, and require only limits on the total audited amount based on a sample of audited amounts, the mean estimator supplied by Classical Variables Sampling may be a good choice. The mean estimator can be computed for samples having no errors or very few errors.

Classical Variables Sampling is performed in three steps: sample preparation, sample audit, and sample evaluation. In sample preparation, the population database is stratified, the required sample size is determined based on our audit objectives, and the sample is extracted. In auditing the sample, we must update the field containing audited amounts with correct values determined through the audit procedure. Finally, the audited sample is evaluated to obtain projections of the most likely total audited amount and error in the population, and precision limits on this quantity. IDEA automates the sample preparation and san sample evaluation phases of the Classical Variables Sampling process.

Classical Variables Sampling may use stratified estimates, in which case the population and sample databases must be stratified, or un-stratified estimates, in which case they are not. Generally, stratification reduces the sample size needed to achieve a desired result but requires a larger number of errors in the sample. Various estimation techniques exist in Classical Variables Sampling theory. The most efficient technique generally depends upon the characteristics of the population database. IDEA supports six Classical Variables Sampling estimation techniques.

- Mean
- Difference
- Combined Ratio
- Separate Ratio
- Combines Regression
- Separate Regression.

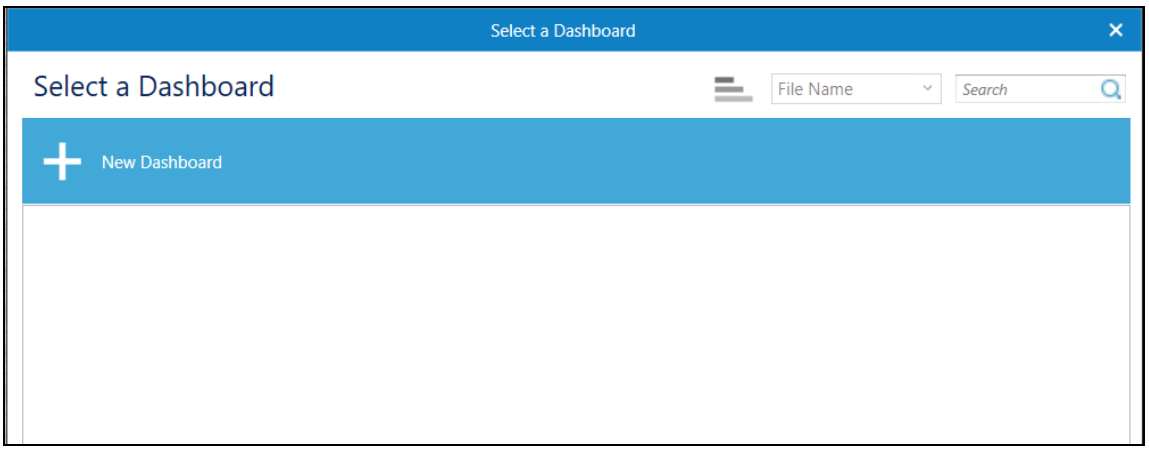
Mean estimation is the last efficient of all and generally not used, except for making comparisons, or estimating quantities for which no recorded values exist prior to the audit.

10.9 VISUALISATION OF DATA IN IDEA

In the IDEA, other than Charts, there are two more options for Visualisation of data –

- Visualize
- Discover

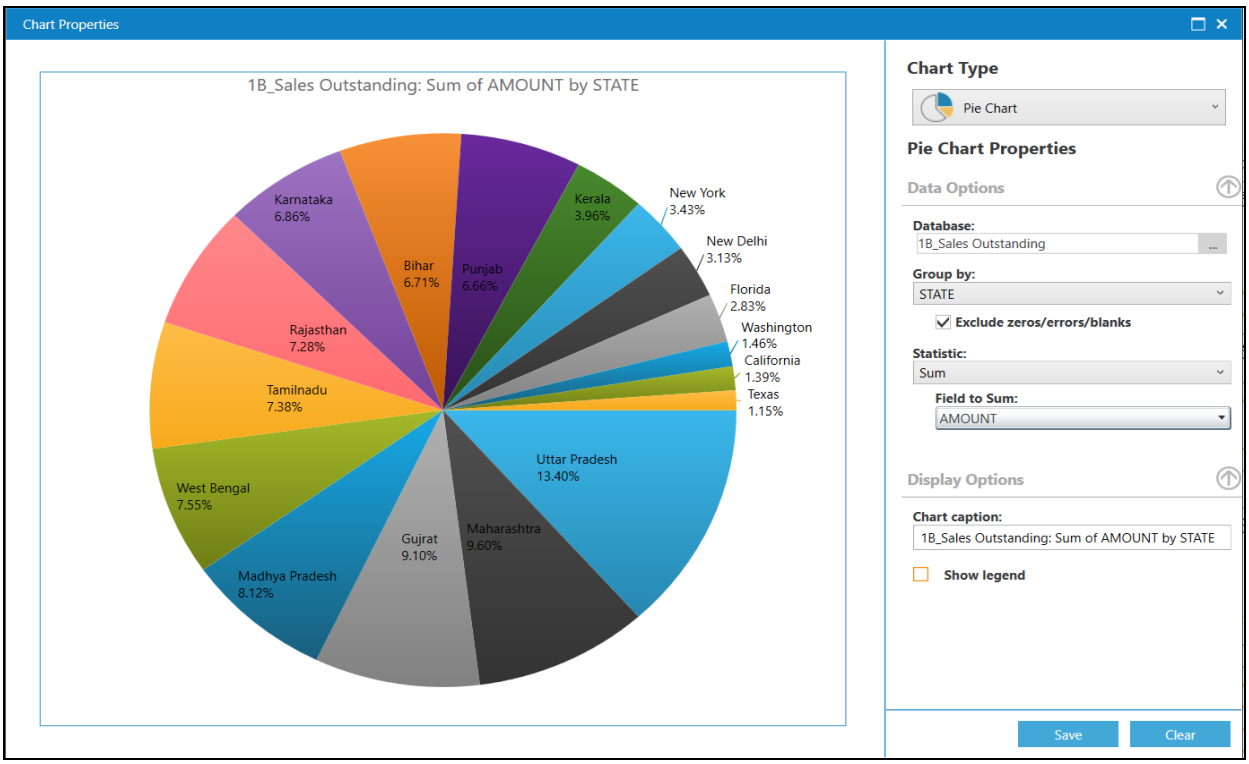
- **Visualize:** This feature is available in **Analysis tab > Visualization > Visualize**. This helps to explore and visualise the data. When we click on Visualize option, it opens a dialogue box for Dashboard.



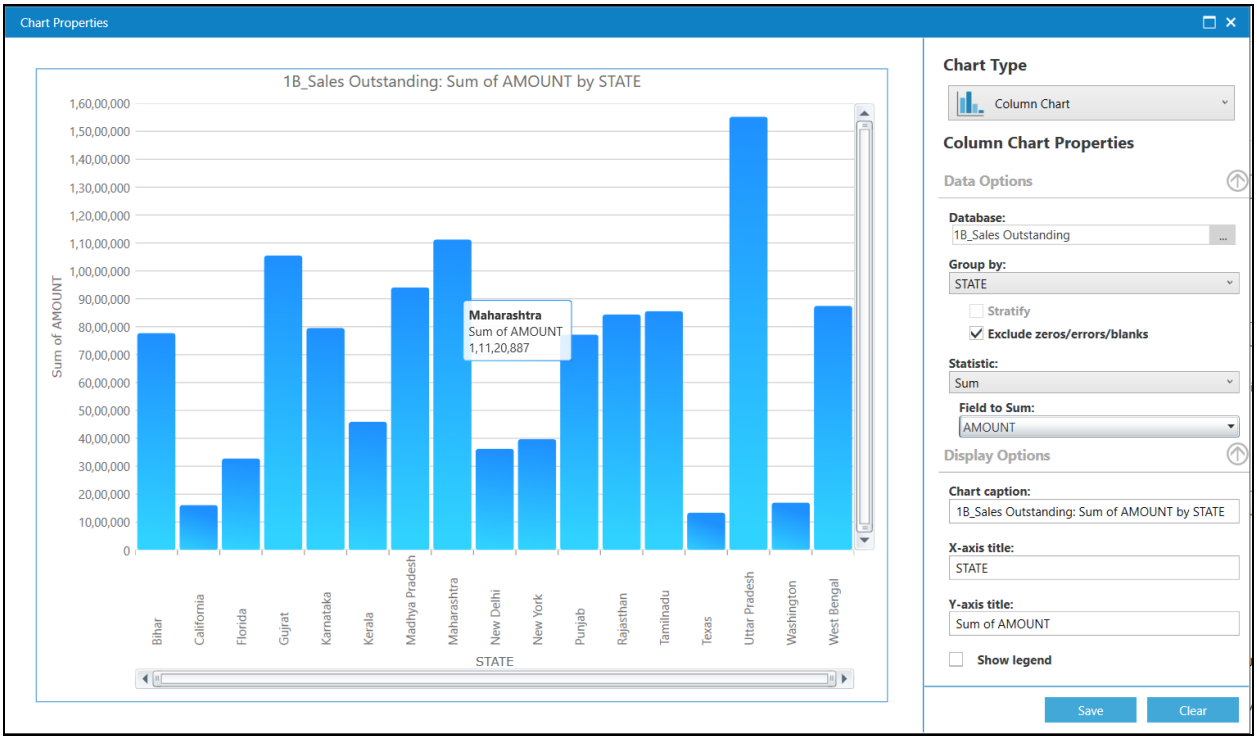
We need to click on "New Dashboard". It will open a blank Chart Properties dialogue box. Here, we need to select the chart type first based on database. We have 6 chart types available – Column Chart, Bar Chart, Pie Chart, Line Chart, Scatter Chart and Treemap.

In this example, the inputs are:

- **Chart Type:** Pie chart.
- **Database:** We need to select the database which we want to visualise. Here, we are selecting "1B_sales Outstanding"
- **Group by:** State column
- **Exclude zeros/errors/blanks:** Tick box to avoid any blanks/ errors / zeros
- **Statistic:** SUM (3 option available – Sum, Count, Average)
- **Field to sum:** Amount column

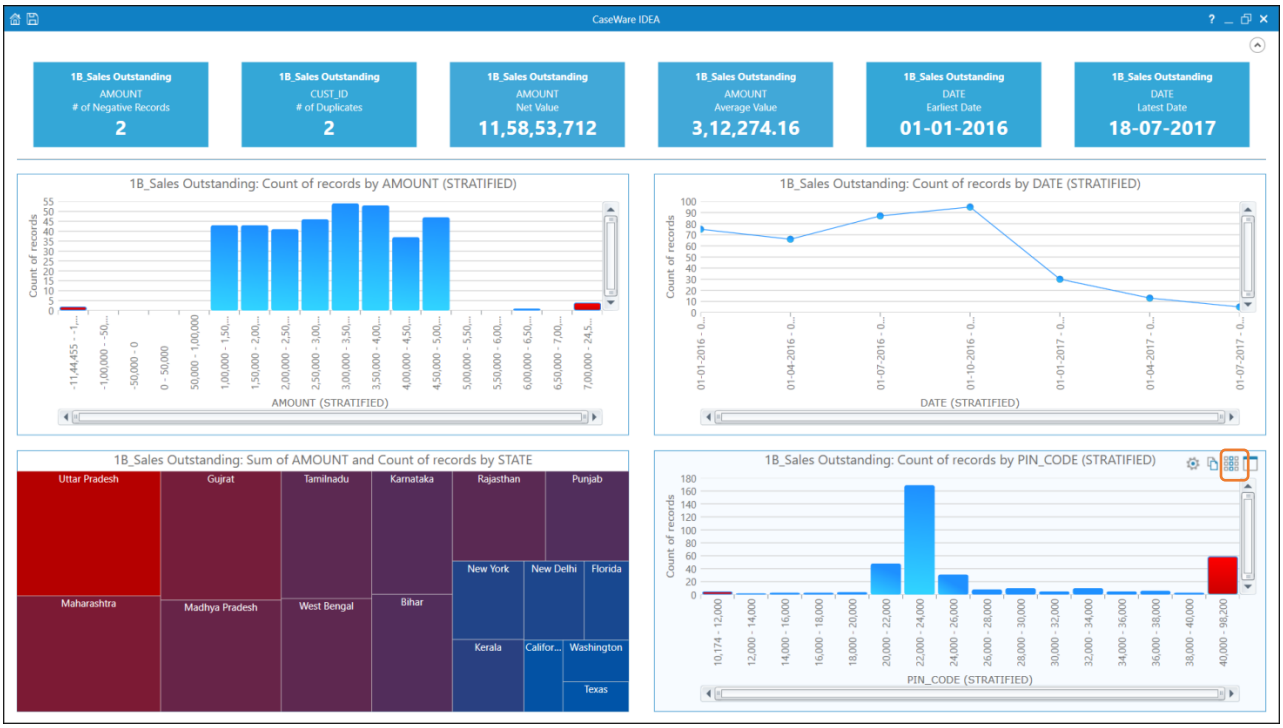


In case, we change the chart type to **Column Chart**, the visualisation appears as follows:



This way it becomes easy to visualise the data and make the inferences. We can save the Chart and it will be saved in the Visualization folder in Idea working project.

- **Discover:** This feature is available in **Analysis tab > Visualization >Discover**. This helps to discover insights on the active database. When we click at the Discover option, it analyses the database and automatically creates charts based on various parameters. It also identifies (in any) the key summary from the database as shown below:



In case, we want we can change the charts by clicking at the Properties button on Top right side of each chart and field statistics.

10.10 GAP DETECTION

This feature is available in **Analysis Tab > Explore > Gap Detection**. We use the Gaps task to detect missing elements in a sequence of numbers, characters, or dates. To start performing this task, we must specify the field we want to use in the "Field To Use" option. If the field chosen is a character field, we must also provide a value for the Mask property that specifies the mask we want IDEA to use to search for gaps. For example, if character field consists of four numbers, then provide a mask of NNNN where each N represents a number. We can use C to specify a character in place of a number.

Here, for the purpose of Gap detection, we are using Sales Database. It has 2 years of data, with following columns:

- SALES DATE
- INVOICE NO.

- CUSTOMER NAME
- CUST_ID
- AMOUNT

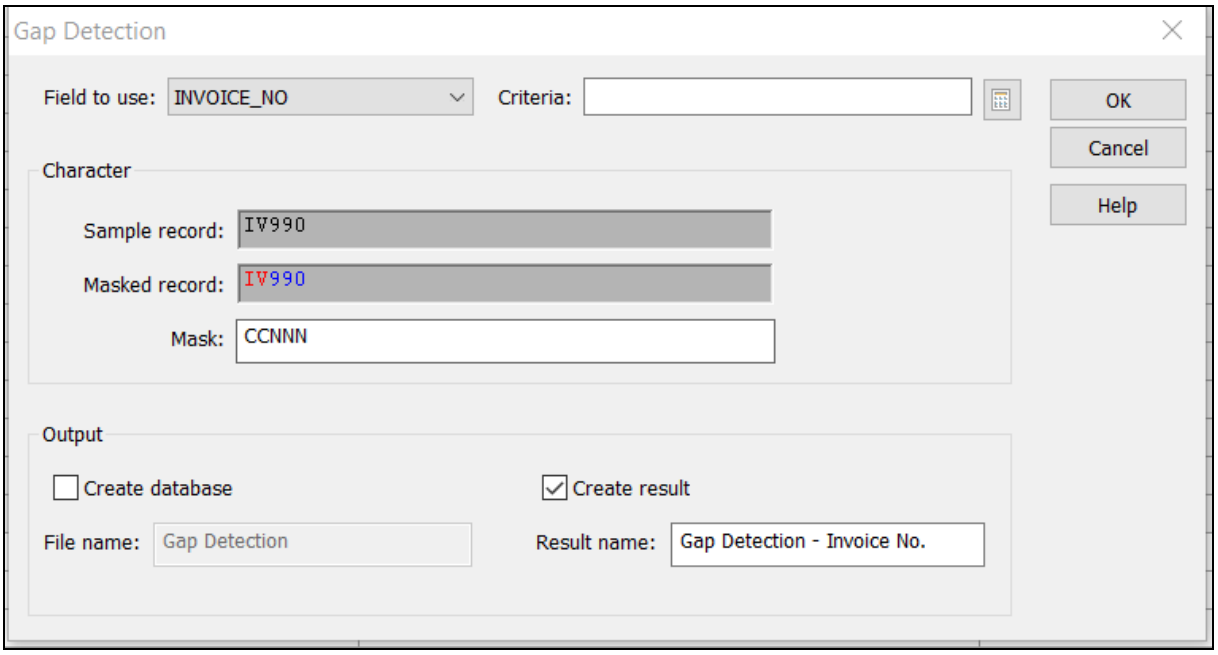
In this example, we want to identify the Gap in Dates and Invoices.

Example 1: Gap in invoices (Missing invoices)

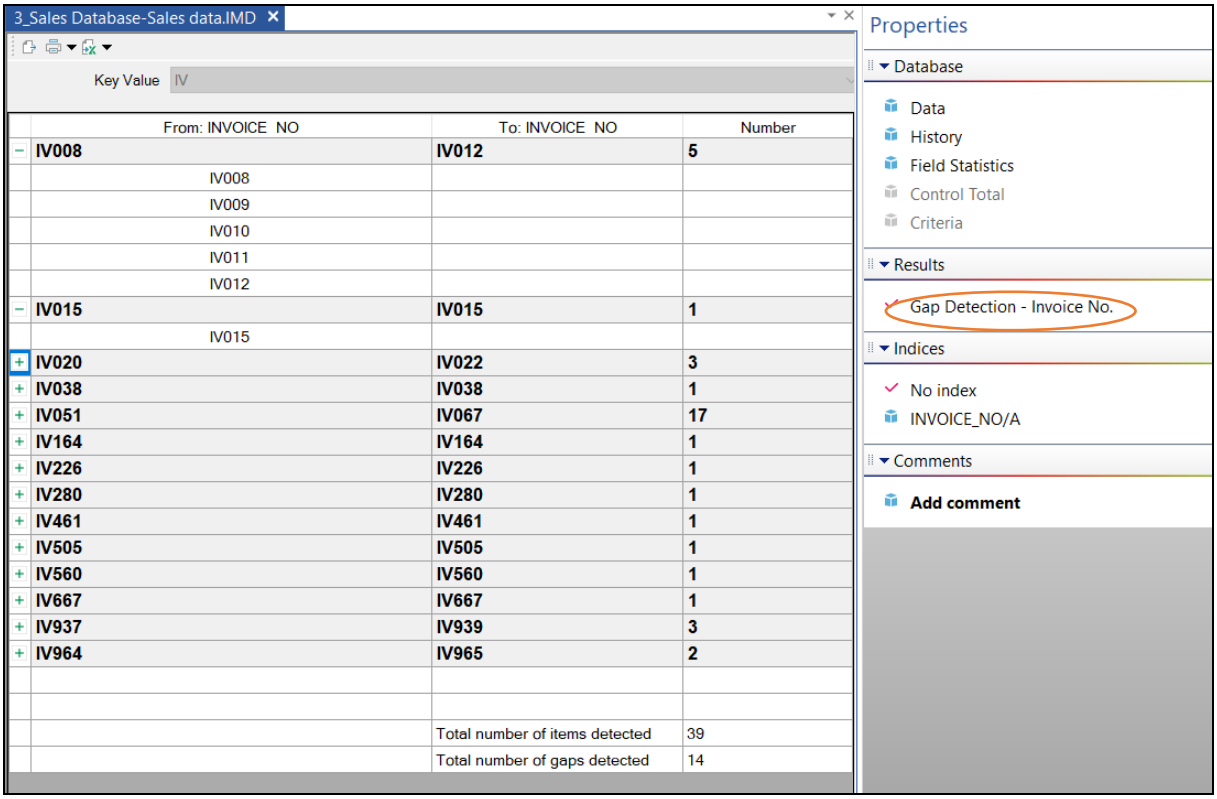
We open the Sales Database file and click on Gap Detection option. It opens "Gap Detection" dialogue box. We provide following inputs:

- **Field to use:** INVOICE_NO
- **Criteria:** blank (No criteria)
- **Sample record:** default value
- **Masked record:** default value
- **Mask:** CCNNN (default value – picked up by IDEA)
- **Output:** Not required; hence not ticked
- **Create Result:** Ticked
- **Result Name:** Gap Detection - Invoice No.

Note: The inputs will change as per the field we use for Gap detection.



The output result is as follows:



It can be seen that the result contains a summary list of invoices (From and To) and number of invoices missing. When we expand each row (say IV008 to IV012), we find that there are 5 missing invoices and which are numbered as IV008, IV009, IV010, IV011 and IV012.

In the same manner, when we expand, second line i.e. IV015 to IV015, it shows that there is only 1 invoice missing i.e. IV015.

Like this, 14 gaps were detected, which contain list of 39 missing invoices.

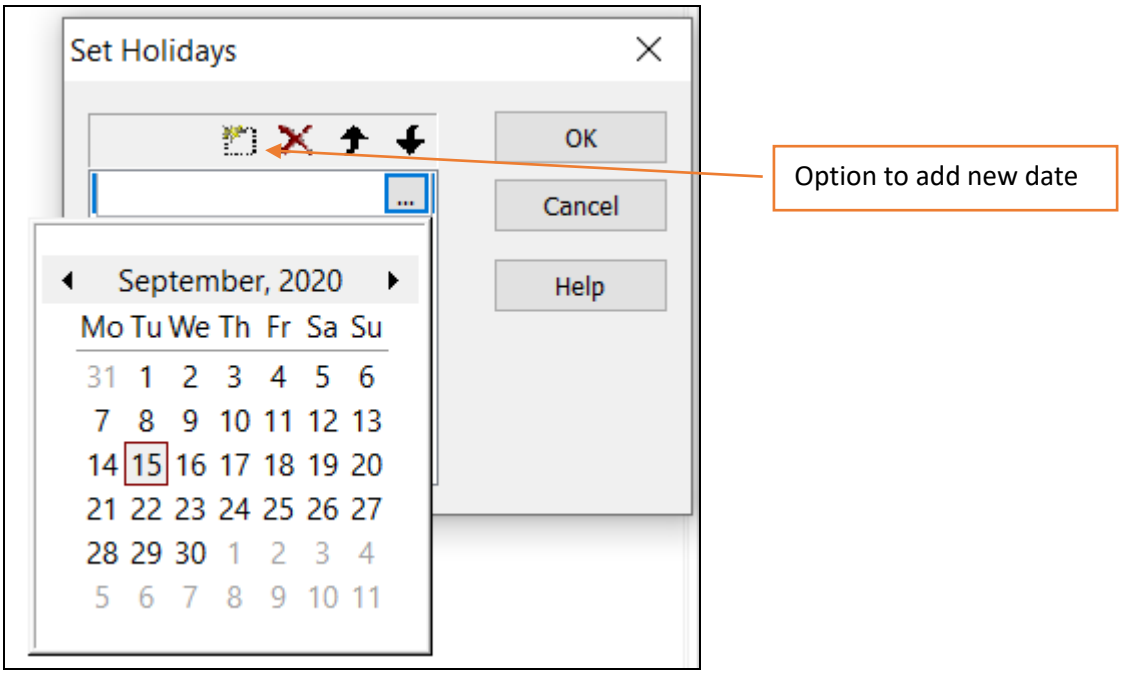
Example 2: Gaps in Sales Dates (No transaction dates)

We open the Sales Database file and click on Gap Detection option. It opens "Gap Detection" dialogue box. We provide following inputs:

- **Field to use:** SALES_DATE
- **Criteria:** blank (No criteria)
- **Date (All or range):** All
- **Ignore Weekends:** Tick (It tells IDEA to ignore weekends when set to True.)
- **Ignore Holidays:** Tick (It tells IDEA to ignore holidays we set with the Holidays property when set to True.)

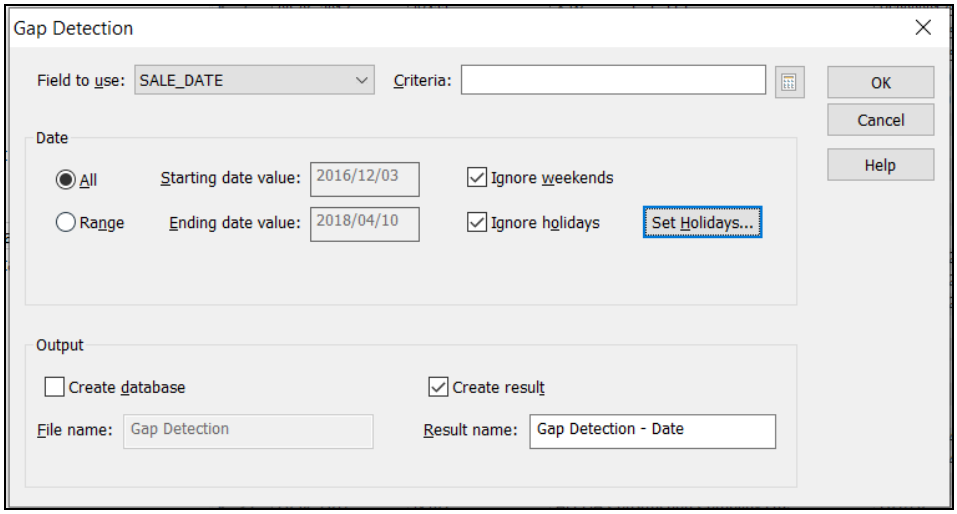
Holidays: Defines holiday dates. You provide the dates as a contiguous list, which can look pretty confusing if you run them all together:

Say Holiday was on 13/01, 18/1. 26/1, 12/7 every year. Then we need to use the calendar and select the individual dates for holiday.

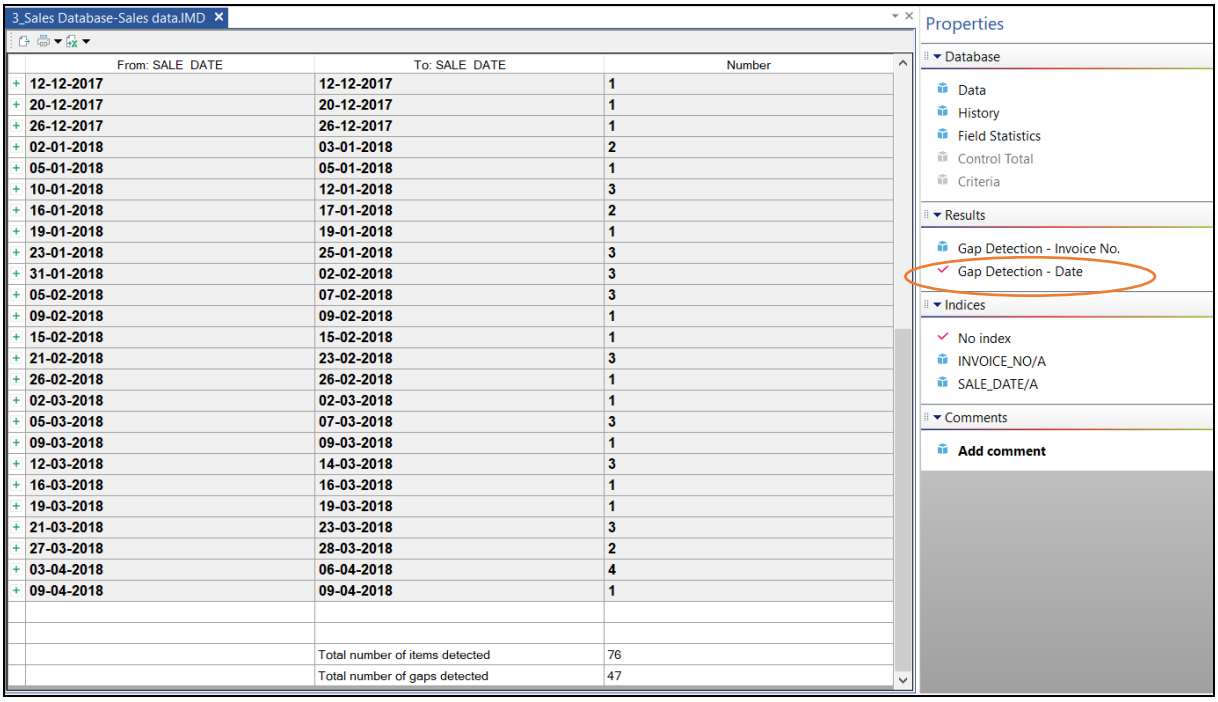


- **Output:** Not required; hence not ticked
- **Create Result:** Ticked
- **Result Name:** Gap Detection - Date

Note: The inputs will change as per the field we use for Gap detection.



The output file for gaps in dates:



3_Sales Database-Sales data.IMD

From: SALE_DATE	To: SALE_DATE	Number
12-12-2017	12-12-2017	1
20-12-2017	20-12-2017	1
26-12-2017	26-12-2017	1
02-01-2018	03-01-2018	2
05-01-2018	05-01-2018	1
10-01-2018	12-01-2018	3
16-01-2018	17-01-2018	2
19-01-2018	19-01-2018	1
23-01-2018	25-01-2018	3
31-01-2018	02-02-2018	3
05-02-2018	07-02-2018	3
09-02-2018	09-02-2018	1
15-02-2018	15-02-2018	1
21-02-2018	23-02-2018	3
26-02-2018	26-02-2018	1
02-03-2018	02-03-2018	1
05-03-2018	07-03-2018	3
09-03-2018	09-03-2018	1
12-03-2018	14-03-2018	3
16-03-2018	16-03-2018	1
19-03-2018	19-03-2018	1
21-03-2018	23-03-2018	3
27-03-2018	28-03-2018	2
03-04-2018	06-04-2018	4
09-04-2018	09-04-2018	1
	Total number of items detected	76
	Total number of gaps detected	47

It shows that there are 47 gaps detected, with no transactions on 76 dates in the entire Sales database.

It is to be remembered that the objective of Gap detection is to find the gaps in the specified fields. So it is going to find gaps in 2 dates or invoice numbers. But if there is any transaction on holiday, that is not material from Gap detection perspective.

10.11 DUPLICATE KEY

This feature is available in **Analysis Tab > Explore > Duplicate Key**. It helps to identify duplicate items within a database. It has 3 options i.e. Detection, Exclusion and Fuzzy. We are going to review each option one by one below.

Example 1: Detection of Duplicate invoices in "Sales Database"(Same-Same-Same (SSD) Test)

The Duplicate Key Detection option helps to perform Same-Same-Same (SSS) test. The purpose of the Same-Same-Same (SSS) test is to identify abnormal duplications as potential indicators of errors or fraud. Application of this test assists in detecting duplicate expenses claimed, occurrences of the same payment to vendors made in error, multiple warranty claims, or duplicated service fees paid by private or government health plans.

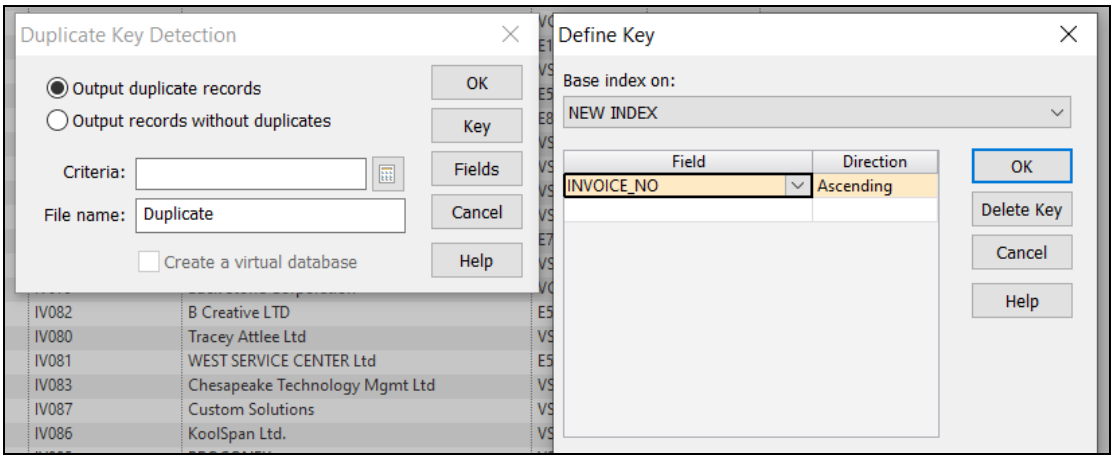
This option can output duplicate records or its inverse of outputting records without duplicates. Up to eight fields may be selected by the auditor to match.

The same-same-same test can be used with various combinations of fields that all must match before outputting the results. In our example test, we wish to see if there are any duplicate invoices in Sales Database. We have option to select multiple variables.

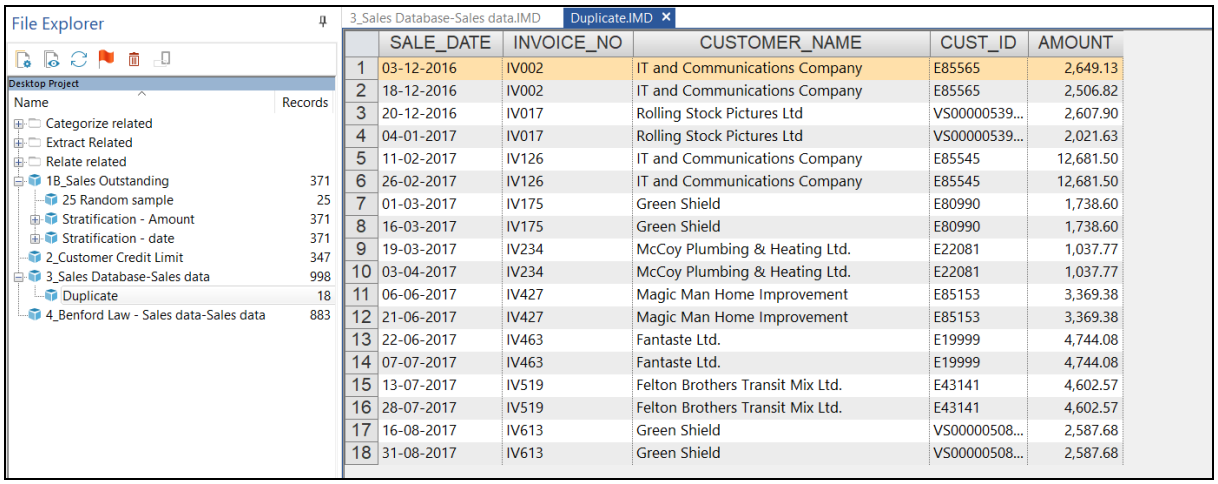
First we open Sales Database. Then we click on Detection option in Duplicate Key option. It opens "Duplicate Key Detection" dialogue box. We select the option "output duplicate records".

- Then we click on Key button to define Invoice number field as the Key for duplicate detection.
- The Criteria option kept blank (no criteria). In case, we need we can add a criteria.
- Output file name: Duplicate

Then we click on OK button.



The output file is as follows:



	SALE_DATE	INVOICE_NO	CUSTOMER_NAME	CUST_ID	AMOUNT
1	03-12-2016	IV002	IT and Communications Company	E85565	2,649.13
2	18-12-2016	IV002	IT and Communications Company	E85565	2,506.82
3	20-12-2016	IV017	Rolling Stock Pictures Ltd	VS00000539...	2,607.90
4	04-01-2017	IV017	Rolling Stock Pictures Ltd	VS00000539...	2,021.63
5	11-02-2017	IV126	IT and Communications Company	E85545	12,681.50
6	26-02-2017	IV126	IT and Communications Company	E85545	12,681.50
7	01-03-2017	IV175	Green Shield	E80990	1,738.60
8	16-03-2017	IV175	Green Shield	E80990	1,738.60
9	19-03-2017	IV234	McCoy Plumbing & Heating Ltd.	E22081	1,037.77
10	03-04-2017	IV234	McCoy Plumbing & Heating Ltd.	E22081	1,037.77
11	06-06-2017	IV427	Magic Man Home Improvement	E85153	3,369.38
12	21-06-2017	IV427	Magic Man Home Improvement	E85153	3,369.38
13	22-06-2017	IV463	Fantaste Ltd.	E19999	4,744.08
14	07-07-2017	IV463	Fantaste Ltd.	E19999	4,744.08
15	13-07-2017	IV519	Felton Brothers Transit Mix Ltd.	E43141	4,602.57
16	28-07-2017	IV519	Felton Brothers Transit Mix Ltd.	E43141	4,602.57
17	16-08-2017	IV613	Green Shield	VS00000508...	2,587.68
18	31-08-2017	IV613	Green Shield	VS00000508...	2,587.68

The output file has 18 records of Duplicate invoices. The list of duplicate invoices shown in Invoice_No. column.

While inputting data in "Duplicate Key Detection" dialogue box, if we select the option "output records without duplicates", then as the option suggests, it will return the output file without any duplicate invoice number.

Example 2: Duplicate Key - Exclusion option (Same-Same-Different (SSD) Test)

The same-same-different (SSD) test is used to identify records with near duplicates for fields selected by the auditor. The auditor may select up to eight fields to match and one field that is excluded from the matching.

Mark Nigrini states, "The same-same-different test is a powerful test for errors and fraud. This test should be considered for every forensic analytics project." His experience has shown that, "This test always detects errors in accounts payable data" and "The longer the time period, the higher the chances of SSD detecting errors."

One example that demonstrates the value of this test is in the detection of instances where Customer Name and Mobile number is same but Customer ID is different (same Customer Name, same Mobile Number, but different Cust ID). The auditor needs to follow-up whether same Customer has been assigned 2 different Cust ID.

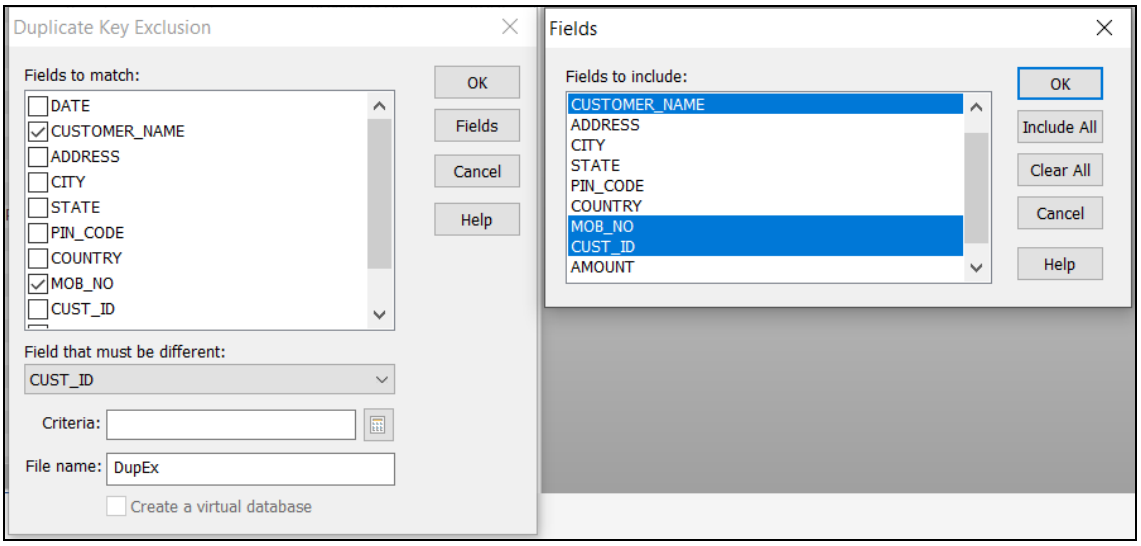
The built-in Duplicate Key Exclusion feature of IDEA is simple and easy to use. However, it is limited, as it does not display more than one exclusion.

First we need to open "Sales Outstanding" Database. While preparing field statistics, we observed that there are many duplicate names, Cust ID, Mobile Number and Address. The duplicate names, mobile number, Cust ID, etc, we can find using Duplicate option.

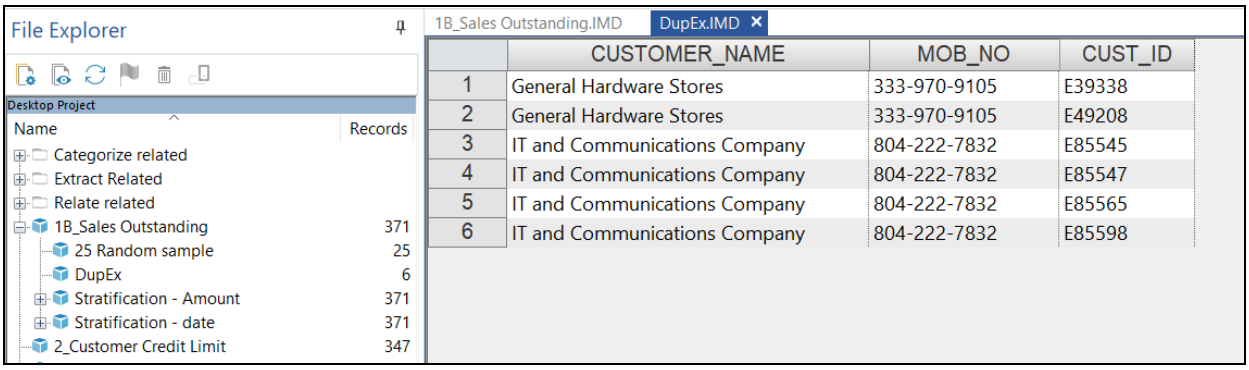
But here, we want to find the cases where Customer Name and Mobile Number is same but the Cust ID is different.

The inputs are:

- **Field to match:** we select Customer Name and Mobile No. column
- **Field that must be different:** CUST_ID column
- In **fields** for output, we are selecting only 3 columns i.e. Customer Name, Mobile No. and Cust ID.
- **File name:** DupEx (default name)



The output file is as follows:



	CUSTOMER_NAME	MOB_NO	CUST_ID
1	General Hardware Stores	333-970-9105	E39338
2	General Hardware Stores	333-970-9105	E49208
3	IT and Communications Company	804-222-7832	E85545
4	IT and Communications Company	804-222-7832	E85547
5	IT and Communications Company	804-222-7832	E85565
6	IT and Communications Company	804-222-7832	E85598

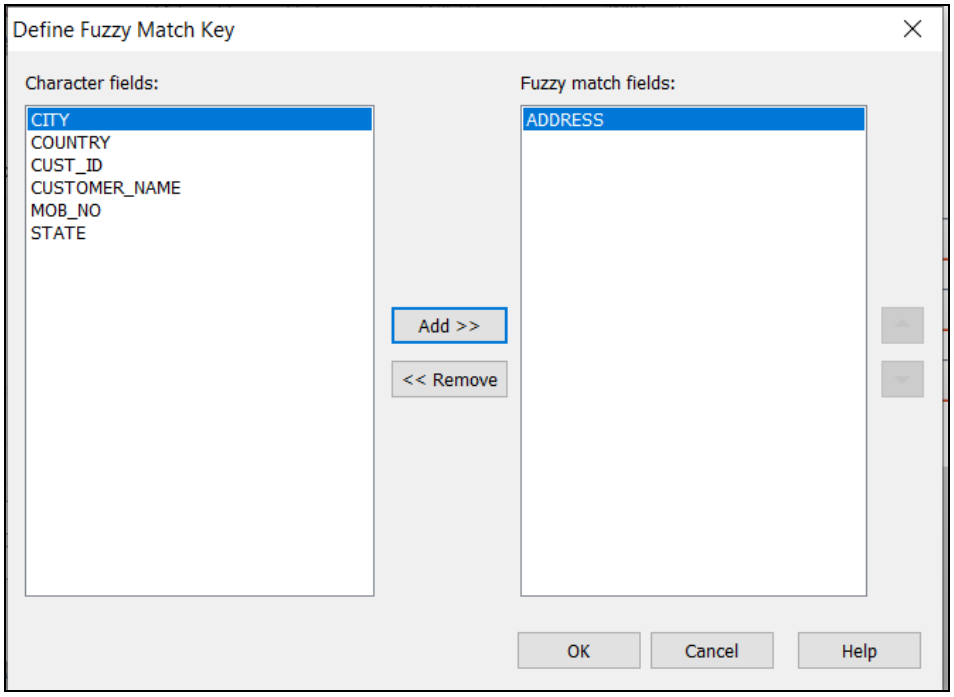
It shows that there are 2 parties which common mobile number but different Cust ID, which needs explanation from the company.

Example 3: Duplicate detection using Fuzzy logic in Sales Outstanding

As the name suggest, the Fuzzy logic, matches values based on number of characters used. It will give result for <=100% match values.

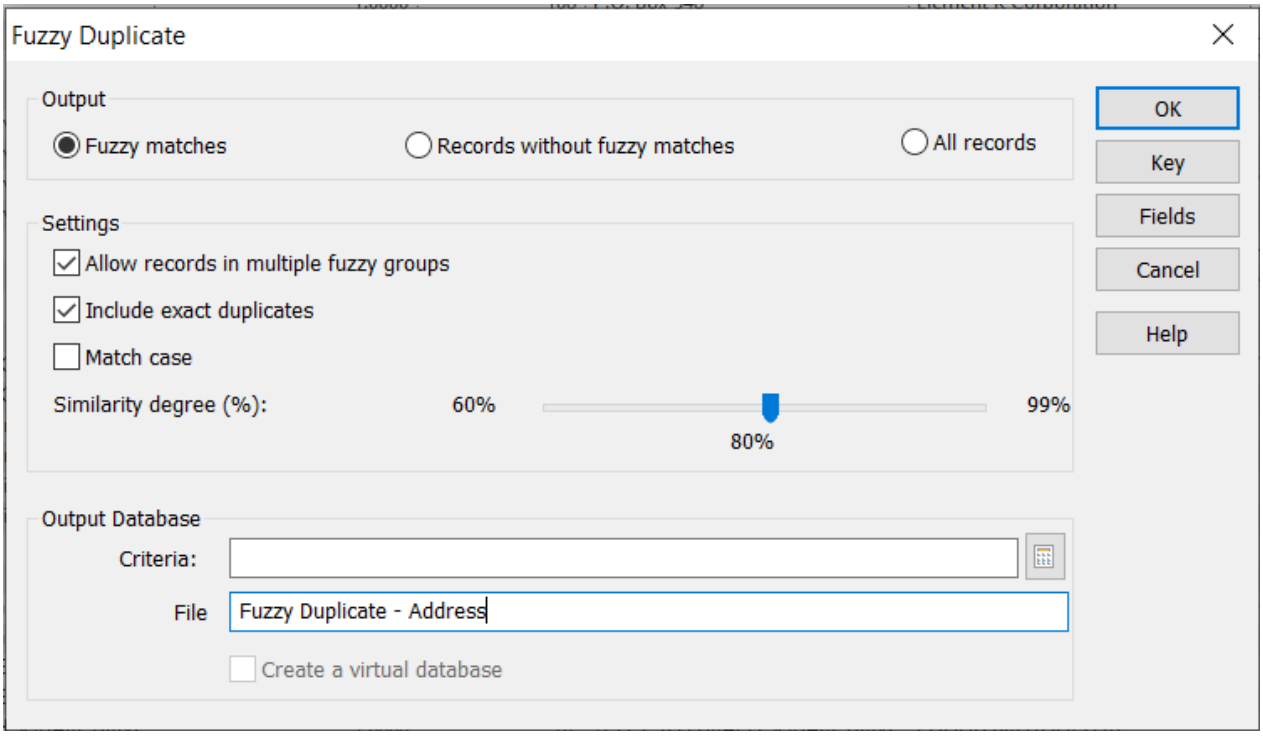
In our example of Sales Outstanding, we want to apply Fuzzy Logic on Address column. The inputs are:

- **Output** (Fuzzy matches / Records without fuzzy matches / All records): Fuzzy matches
- In **Setting**:
  - All records in multiple fuzzy groups: Tick (default)
  - Include exact duplicates: Tick (default)
  - Match case: Not ticked (address can be written in various pattern)
- **Similarity degree:**80% (default value; we can move the slider bar to change the similarity degree. The higher the degree of similarity, the more **similar** the records will be in the Fuzzy groups)
- **Criteria:** Blank (not defined)
- **File Name:** Fuzzy Duplicate - Address
- In the Key option, we select Address column.

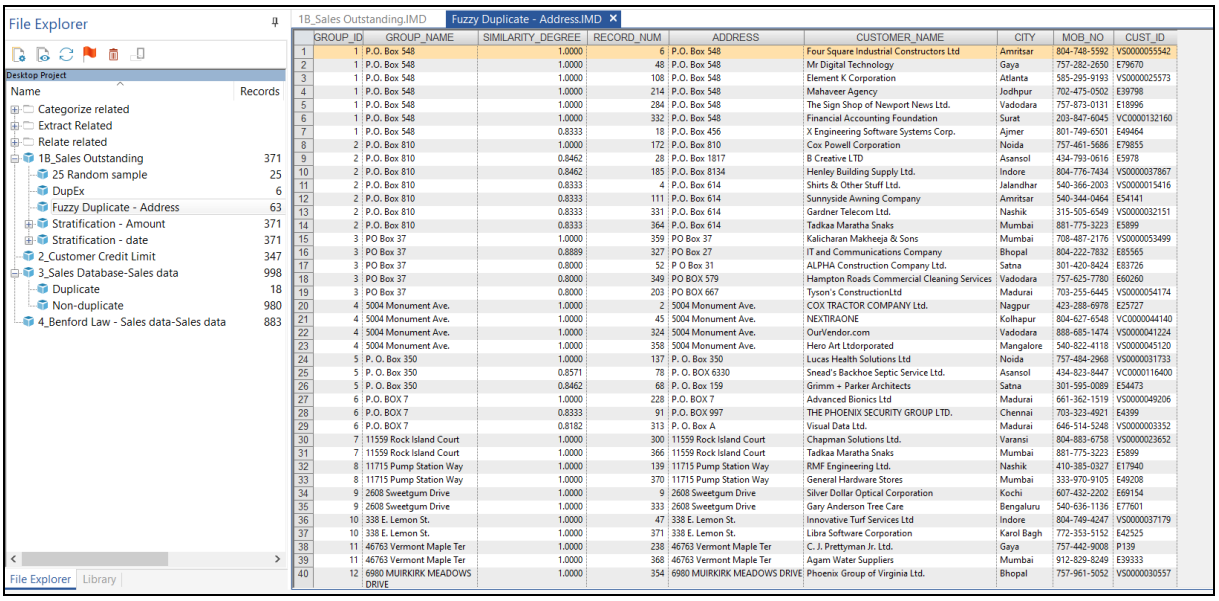


- In the fields option, we select 5 fields i.e. Customer Name, Address, City, Mobile No. and Cust ID.

The final input screen is as follows:



The output file is as follows:



	GROUP_ID	GROUP_NAME	SIMILARITY_DEGREE	RECORD_NUM	ADDRESS	CUSTOMER_NAME	CITY	MOB_NO	CUST_ID
1	1	P.O. Box 548	1.0000	6	P.O. Box 548	Four Square Industrial Constructors Ltd	Amritsar	804-748-5592	VS0000055542
2	1	P.O. Box 548	1.0000	48	P.O. Box 548	Mr Digital Technology	Gaya	757-282-2650	E79670
3	1	P.O. Box 548	1.0000	108	P.O. Box 548	Element K Corporation	Atlanta	585-295-9193	VS0000025573
4	1	P.O. Box 548	1.0000	214	P.O. Box 548	Mahaveer Agency	Jodhpur	702-475-0502	E39798
5	1	P.O. Box 548	1.0000	284	P.O. Box 548	The Sign Shop of Newport News Ltd.	Vadodara	757-873-0131	E18996
6	1	P.O. Box 548	1.0000	332	P.O. Box 548	Financial Accounting Foundation	Surat	203-847-6045	VC0000132160
7	1	P.O. Box 548	0.8333	18	P.O. Box 456	X Engineering Software Systems Corp.	Ajmer	801-749-6501	E49464
8	2	P.O. Box 810	1.0000	172	P.O. Box 810	Cox Powell Corporation	Noida	757-461-5686	E79855
9	2	P.O. Box 810	0.8462	28	P.O. Box 1817	B Creative LTD	Asansol	434-793-0616	E5978
10	2	P.O. Box 810	0.8462	185	P.O. Box 8134	Henley Building Supply Ltd.	Indore	804-776-7434	VS0000037867
11	2	P.O. Box 810	0.8333	4	P.O. Box 614	Shirts & Other Stuff Ltd.	Jalandhar	540-366-2003	VS0000015416
12	2	P.O. Box 810	0.8333	111	P.O. Box 614	Sunnyside Awning Company	Amritsar	540-344-0464	E54141
13	2	P.O. Box 810	0.8333	331	P.O. Box 614	Gardner Telecom Ltd.	Nashik	315-505-6549	VS0000032151
14	2	P.O. Box 810	0.8333	364	P.O. Box 614	Tadkaa Maratha Snaks	Mumbai	881-775-3223	E5899
15	3	PO Box 37	1.0000	359	PO Box 37	Kalicharan Makheeja & Sons	Mumbai	708-487-2176	VS0000053499
16	3	PO Box 37	0.8889	327	PO Box 27	IT and Communications Company	Bhopal	804-222-7832	E85565
17	3	PO Box 37	0.8000	52	P O Box 31	ALPHA Construction Company Ltd.	Satna	301-420-8424	E83726
18	3	PO Box 37	0.8000	349	PO BOX 579	Hampton Roads Commercial Cleaning Services	Vadodara	757-625-7780	E60260
19	3	PO Box 37	0.8000	203	PO BOX 667	Tyson's ConstructionLtd	Madurai	703-255-6445	VS0000054174
20	4	5004 Monument Ave.	1.0000	2	5004 Monument Ave.	COX TRACTOR COMPANY Ltd.	Nagpur	423-288-6978	E25727
21	4	5004 Monument Ave.	1.0000	45	5004 Monument Ave.	NEXTIRAONE	Kolhapur	804-627-6548	VC0000044140
22	4	5004 Monument Ave.	1.0000	324	5004 Monument Ave.	OurVendor.com	Vadodara	888-685-1474	VS0000041224
23	4	5004 Monument Ave.	1.0000	358	5004 Monument Ave.	Hero Art Ltdorporated	Mangalore	540-822-4118	VS0000045120
24	5	P. O. Box 350	1.0000	137	P. O. Box 350	Lucas Health Solutions Ltd	Noida	757-484-2968	VS0000031733
25	5	P. O. Box 350	0.8571	78	P. O. BOX 6330	Snead's Backhoe Septic Service Ltd.	Asansol	434-823-8447	VC0000116400
26	5	P. O. Box 350	0.8462	68	P. O. Box 159	Grimm + Parker Architects	Satna	301-595-0089	E54473
27	6	P.O. BOX 7	1.0000	228	P.O. BOX 7	Advanced Bionics Ltd	Madurai	661-362-1519	VS0000049206
28	6	P.O. BOX 7	0.8333	91	P.O. BOX 997	THE PHOENIX SECURITY GROUP LTD.	Chennai	703-323-4921	E4399
29	6	P.O. BOX 7	0.8182	313	P. O. Box A	Visual Data Ltd.	Madurai	646-514-5248	VS0000003352
30	7	11559 Rock Island Court	1.0000	300	11559 Rock Island Court	Chapman Solutions Ltd.	Varansi	804-883-6758	VS0000023652
31	7	11559 Rock Island Court	1.0000	366	11559 Rock Island Court	Tadkaa Maratha Snaks	Mumbai	881-775-3223	E5899
32	8	11715 Pump Station Way	1.0000	139	11715 Pump Station Way	RMF Engineering Ltd.	Nashik	410-385-0327	E17940
33	8	11715 Pump Station Way	1.0000	370	11715 Pump Station Way	General Hardware Stores	Mumbai	333-970-9105	E49208
34	9	2608 Sweetgum Drive	1.0000	9	2608 Sweetgum Drive	Silver Dollar Optical Corporation	Kochi	607-432-2202	E69154
35	9	2608 Sweetgum Drive	1.0000	333	2608 Sweetgum Drive	Gary Anderson Tree Care	Bengaluru	540-636-1136	E77601
36	10	338 E. Lemon St.	1.0000	47	338 E. Lemon St.	Innovative Turf Services Ltd	Indore	804-749-4247	VS0000037179
37	10	338 E. Lemon St.	1.0000	371	338 E. Lemon St.	Libra Software Corporation	Karol Bagh	772-353-5152	E42525
38	11	46763 Vermont Maple Ter	1.0000	238	46763 Vermont Maple Ter	C. J. Prettyman Jr. Ltd.	Gaya	757-442-9008	P139
39	11	46763 Vermont Maple Ter	1.0000	368	46763 Vermont Maple Ter	Agam Water Suppliers	Mumbai	912-829-8249	E39333
40	12	6980 MUIRKIRK MEADOWS DRIVE	1.0000	354	6980 MUIRKIRK MEADOWS DRIVE	Phoenix Group of Virginia Ltd.	Bhopal	757-961-5052	VS0000030557

In the output file, we can see there are 63 records, where Address is matching (<= 100%match) with other address. We can see 3rd column i.e. Similarity Degree to see the percentage of match. We need to closely review these addresses to find any common party.

10.12 BENFORD'S LAW

Benford's law allows auditors to find irregularities in the numerical data by analyzing the digits. The primary objective of using the Benford's law function is to enable auditors to turn raw facts and figures into information and to develop a better understanding of the data to perform subsequent actions..

The Benford's Law feature is available in **Analysis Tab > Explore > Benford's Law**. As per this law, number consistently fall into a pattern with low digits occurring more frequently in the first position than larger digits. The mathematical tenet defining the frequency of digits is known as Benford's Law. In Benford's Law, number patterns are used to detect potential fraud, possible errors, manipulative biases, irregularities, etc.

Benford's law states that in a table of statistics or in a listing of various data such as an invoice amount, cost, and worked hours, each digit has a certain probability of occurrence. For example, if the reimbursement claims do not follow Benford's law, then it may indicate fictitious receipts generated to make a reimbursement claim by the employee.

Using IDEA, we can analyse following Digit Tests:

- 1st Digit Test
- 2nd Digit Test
- First Two Digit Test
- First Three Digit Test

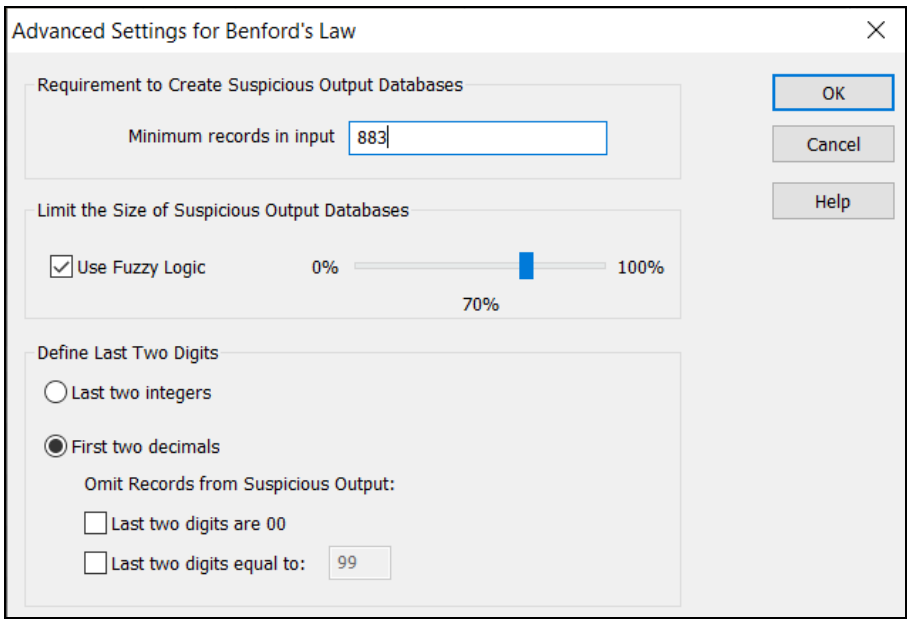
- Last Two Digit Test
- Second Order Test

The Benford's law can be applied on –

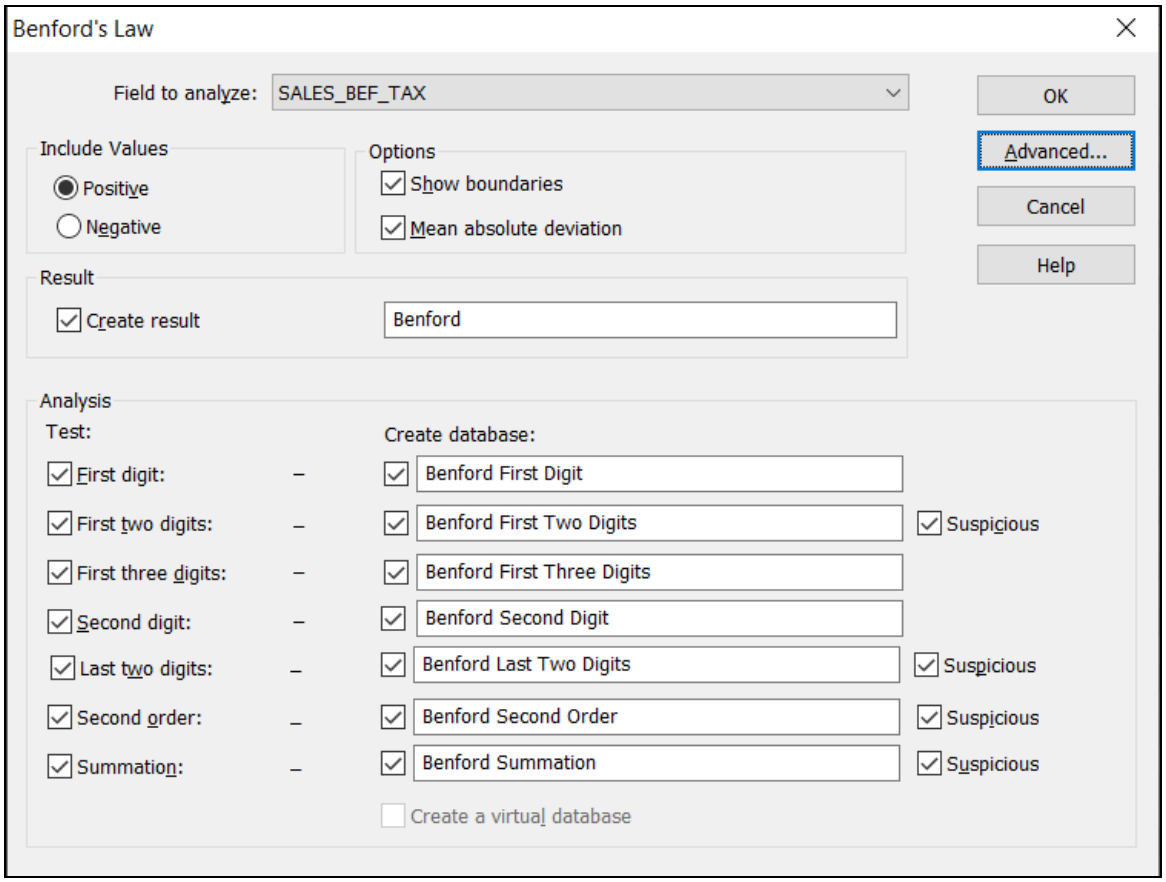
- Sales / Purchases
- Accounts payable (Expenses) data
- Estimations (accruals) in General Ledger
- Non-arm's-length transactions
- Customer refunds
- Bad debts

In this example, we are using "Benford Law – Sales data" (a file created for Benford's Law). The inputs are as follows:

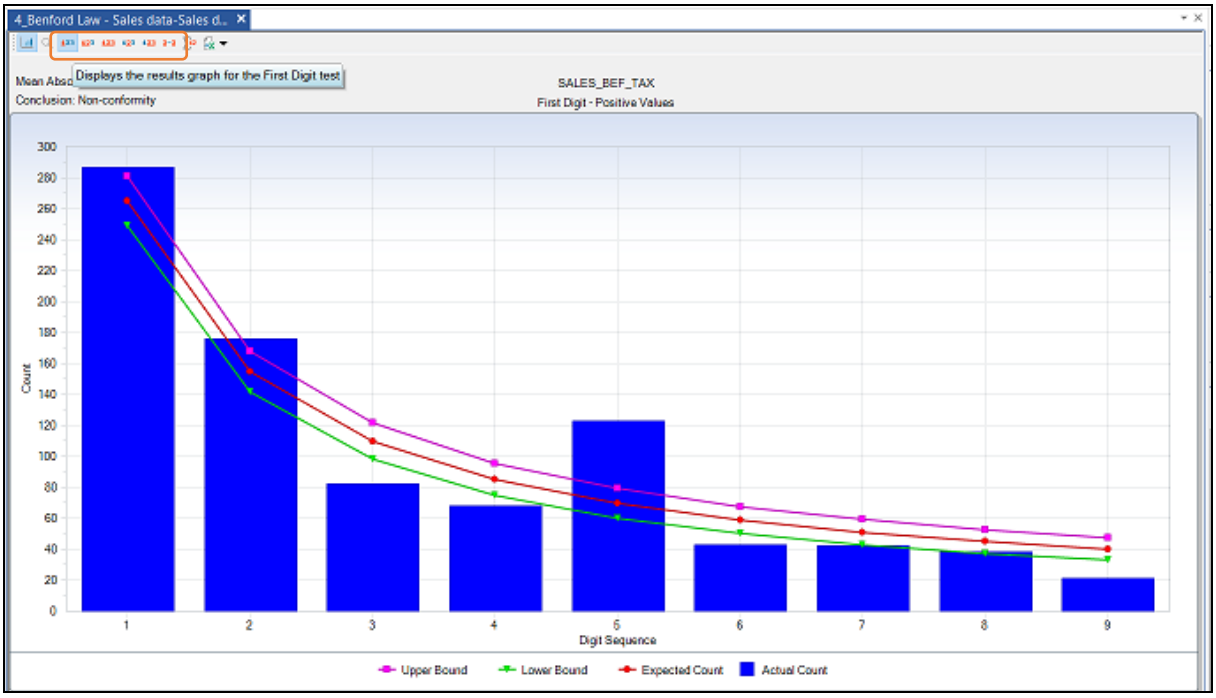
- Field to analyze: SALES_BEF_TAX
- Include Values (Positive / Negative): Positive
- Options:
  - Show boundaries: Tick (default)
  - Mean absolute deviation: Tick (default)
- Create Result: Tick (default)
- Result Name: Benford
- Analysis:
  - First digit: Tick (default)
  - First two digits: Tick (default)
  - First three digits: Tick (default)
  - Second digit: Tick (default)
  - Last two digits: Tick (default)
  - Second order: Tick (default)
  - Summation: Tick (default)
- Create database: : Tick (default) on all boxes
- Advanced button: Default value is 1000 for minimum records, but here we have only 883 records



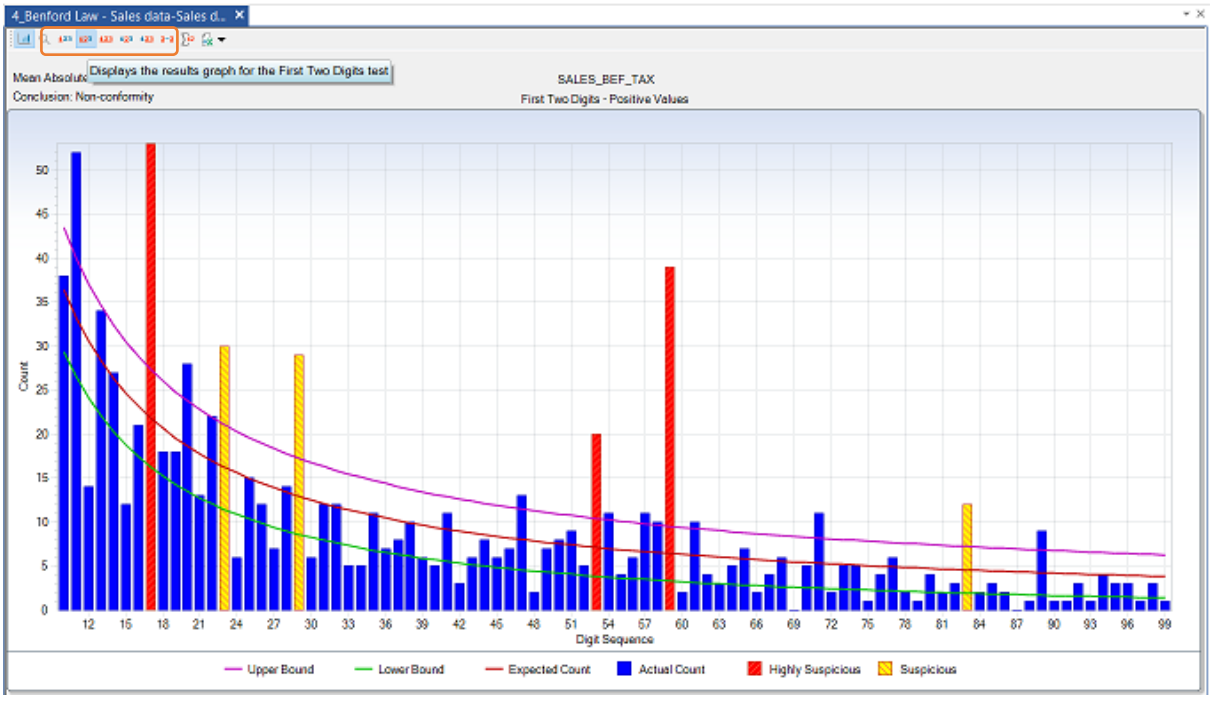
Then we click OK.



The output file is as follows:

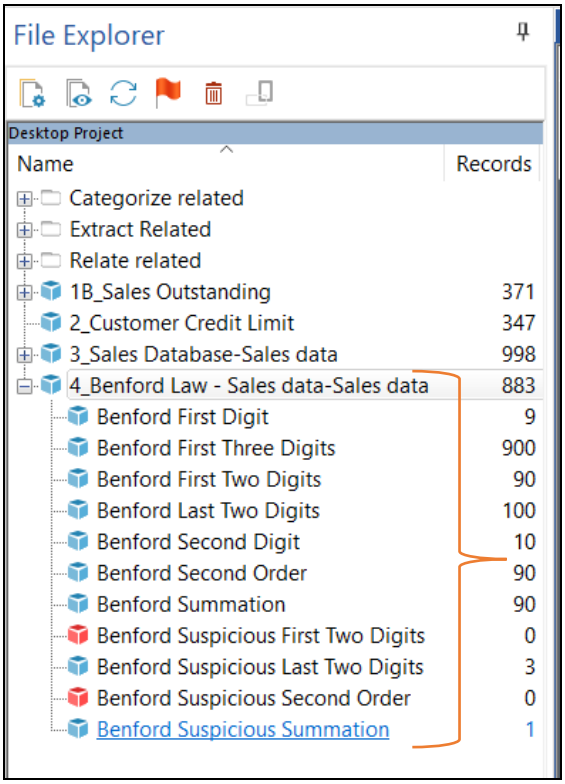


The first output screen is showing results for single digit usage. In this, digit 5 has been used than expected as per Benford's Law. Now we need to analyse all the numbers starting with digit 5 to find manipulation, if any.

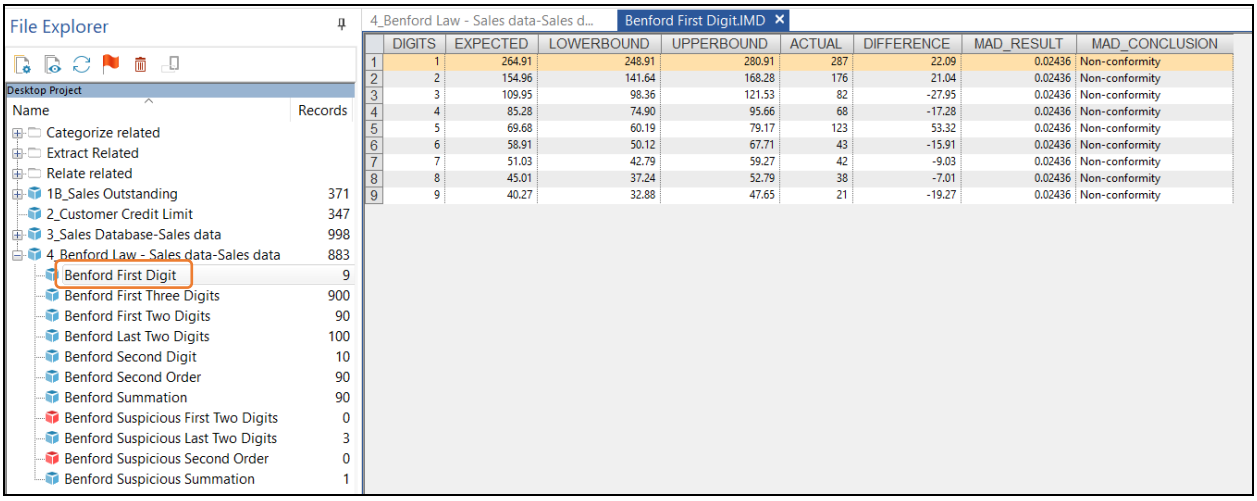


In the above image, we can see usage of first 2 digits from left. It also highlights the suspicious digits as well. Like this, we can review result for various digits and its occurrences.

In the File Explorer, it has created various files as shown below, which we can use for the purpose of analysis.



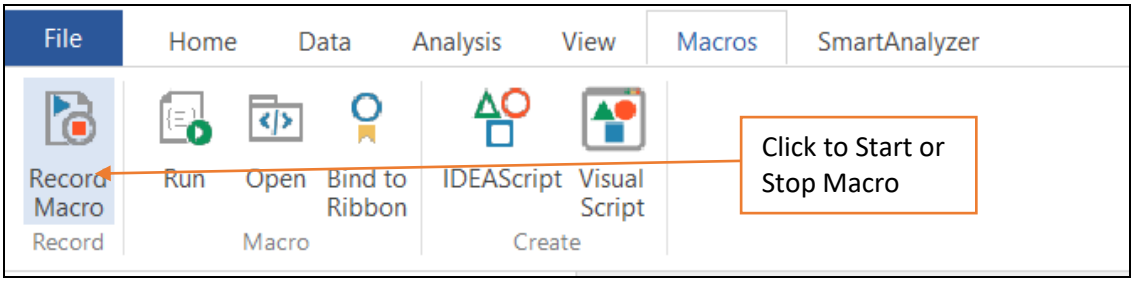
The summary table as per Benford's Law for first digit is as follows:



10.13 MACROS

If we have any tasks that we do repeatedly, we can record a macro to automate those tasks. A macro is an action or a set of actions that we can run as many times as we want. When we create a macro, we are recording your mouse clicks and keystrokes. After we create a macro, we can edit it to make minor changes to the way it works.

This automation can be done in 2 methods, either by Macro (recording of steps) or by VBA / IDEA Script (writing code).

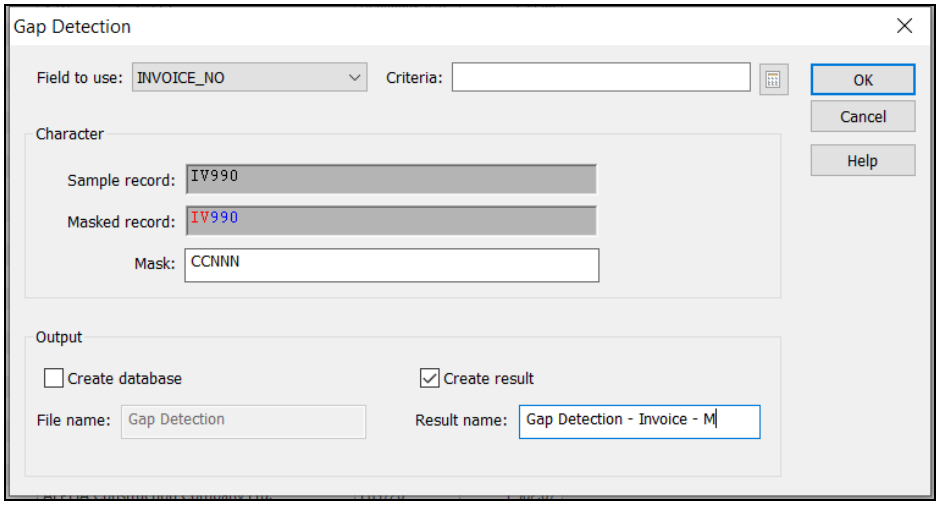


Here we are going to see sample for Macro i.e. recording of steps.

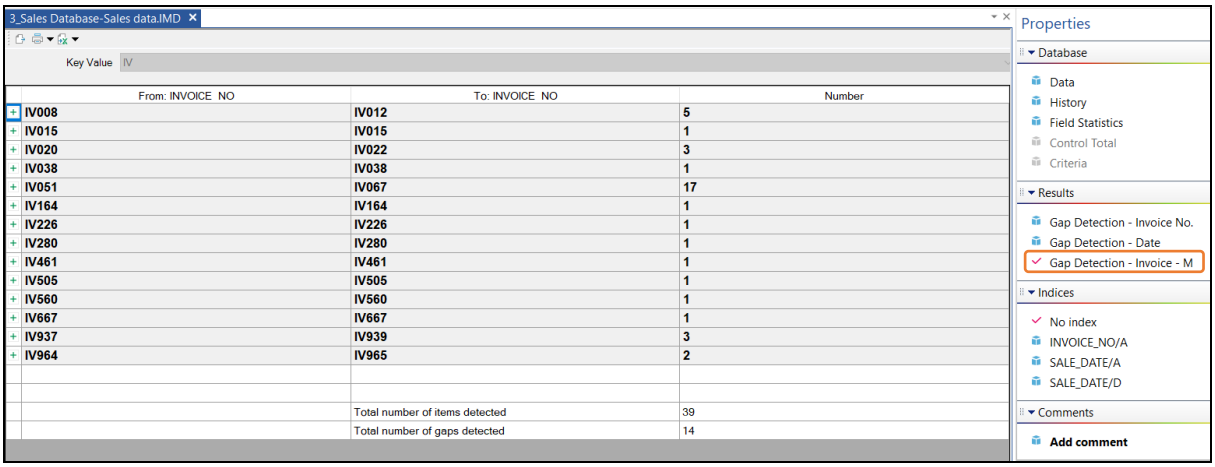
Suppose that every month, we create a report on Gap detection on Invoices for our sales department. We can create and then run a macro for this task.

The steps are as follows:

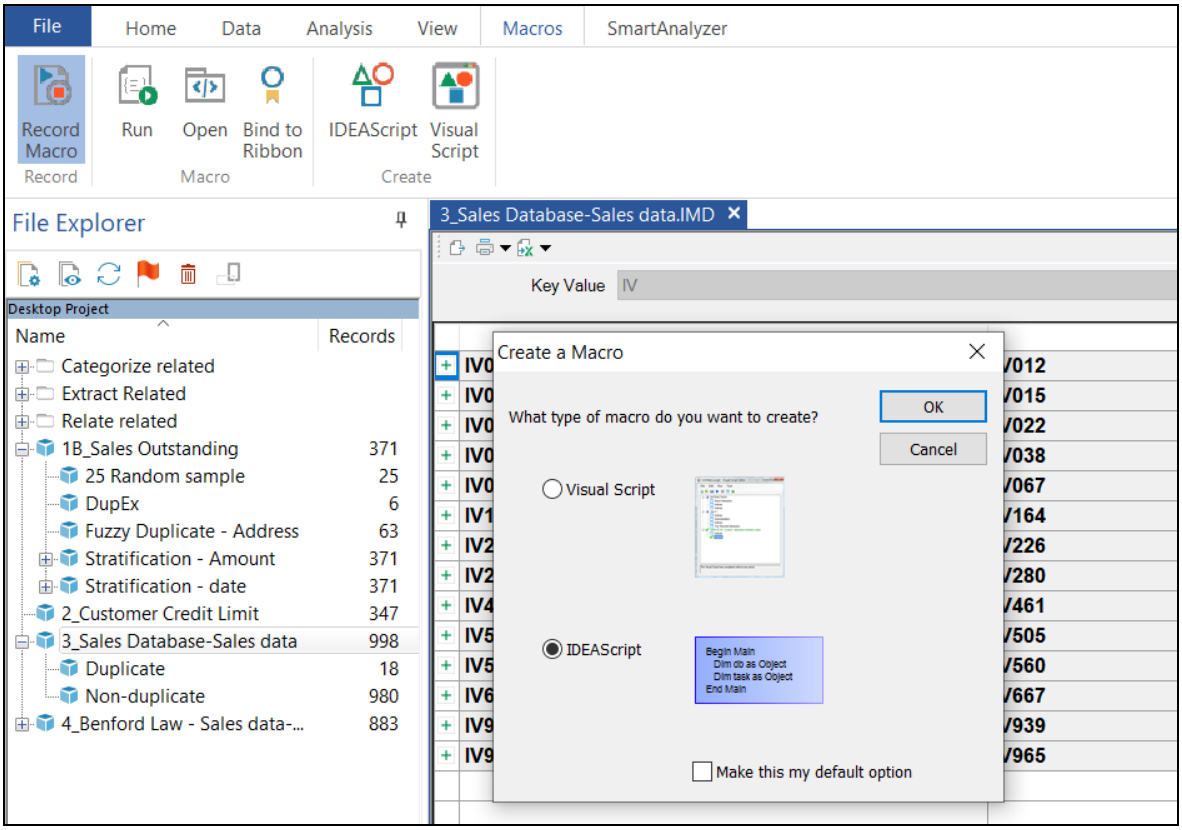
- Start IDEA
- The click in Record Marco button (short cut key – “Ctrl + R”)
- Open the Sales Database file
- Click on Gap Detection option in Explorer in Analysis Tab.
- Give the required inputs as shown below:



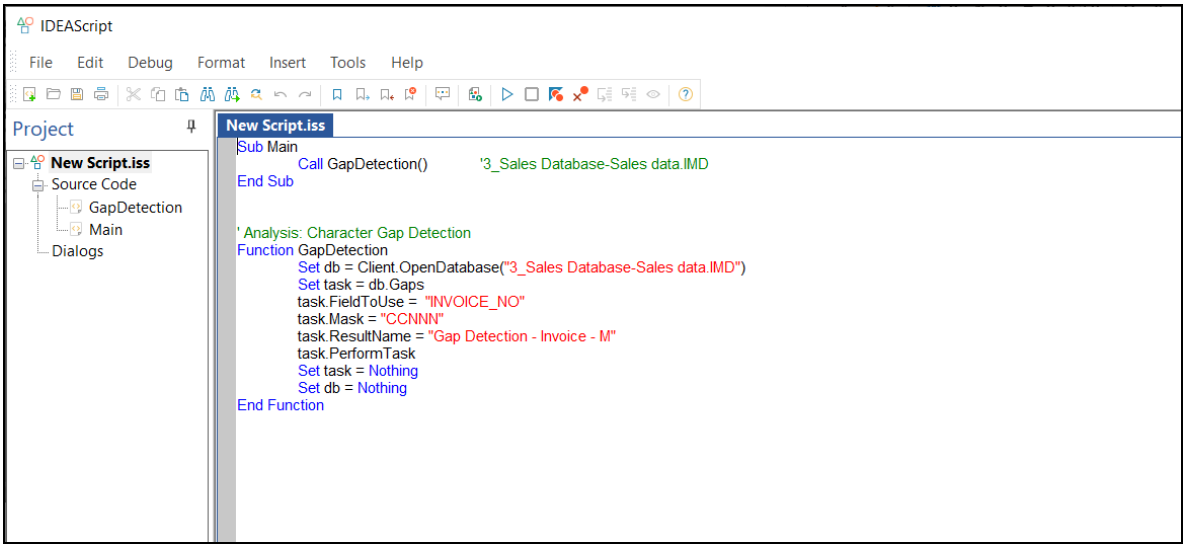
- Click on OK
- The Output on missing invoices using Gap detection will be on screen.



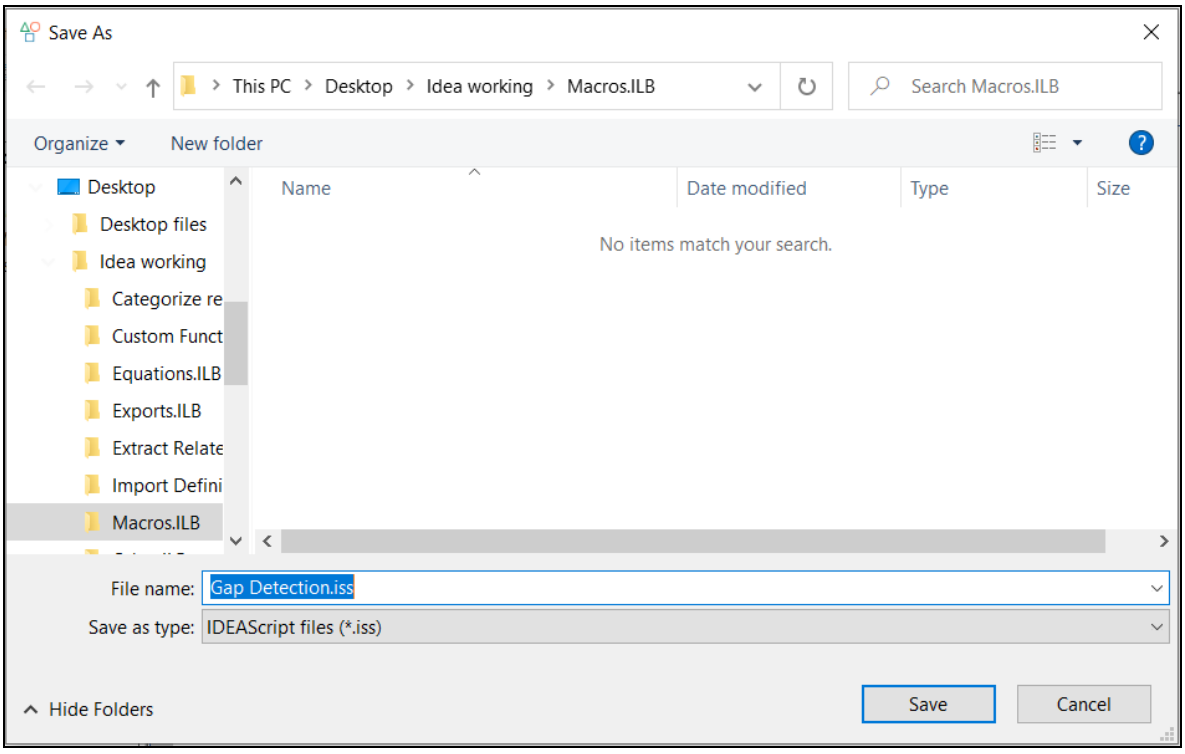
- Stop recording of macro. It will ask to create a Macro either in Visual Script or IDEAScript.
- We select IDEAScript and click OK



- IDEA script for gap detection will be generated.

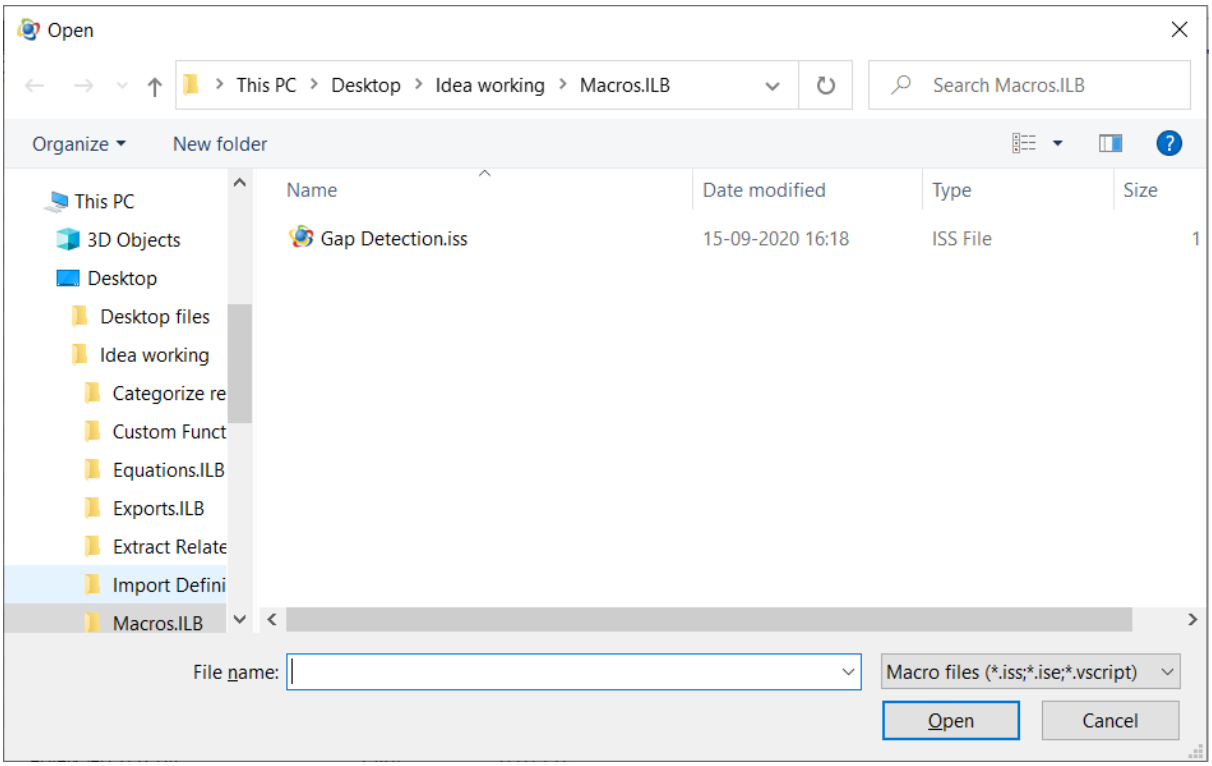


- Save the script for future use. It will be automatically stored in Macros.ILB folder of Idea working project.



Now we can run this macro again and it will perform same action and will give result.

For running an existing macro, click on Run option in Macros. It will open the Macro folder. Select the macro, we want to run and click on OPEN.



Then Gap Detection Macro will run and output will be displayed.

Note: Never execute macros unless we know what task the macro performs and that the macro is safe to use. Executing a macro that we don't know about can cause damage to our data or produce unreliable results. In addition, macros can cause significant problems on our machine, such as installing a virus. The macros that are completely safe are the ones we understand and obtain from a reliable source.

10.14 PRINT REPORT

IDEA makes it possible to output reports in two different forms: PDF and Microsoft Word (the Word output is actually in Rich Text Format, RTF, so any application that can read RTF can use this output). Each output format has advantages and disadvantages. For example, using PDF files makes it easy for everyone to view the report but making changes will be difficult.

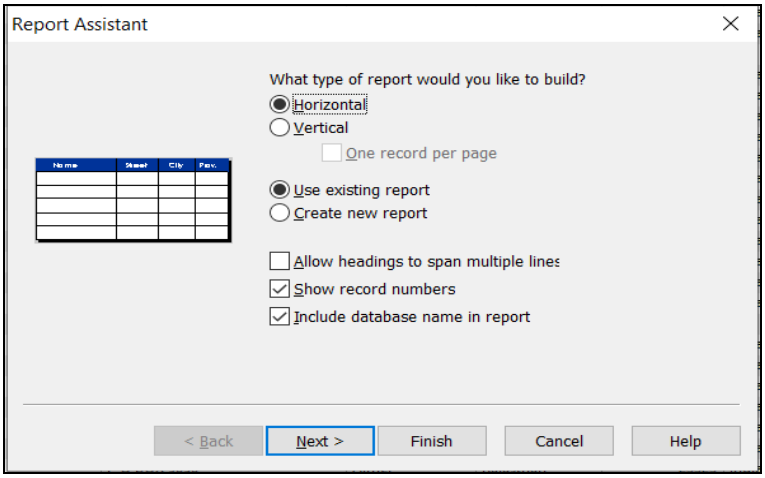
The examples in this section show how to create output directly from a database. Of course, we may want to create customized output, which means writing code to output the custom material.

Defining a Report

Before we can output a report, we need to create one. This is one part of the report creation process that we must perform manually—there isn't any way to code reports. The following steps show how to create the report used for the examples. The same procedure works for any report we want to create.

i) Open the Sample-Customers database by double-clicking its entry in the File Explorer. We see the Idea working database opened in IDEA.

ii) Select File > Print > Create Report. We see the Report Assistant dialog box shown below.

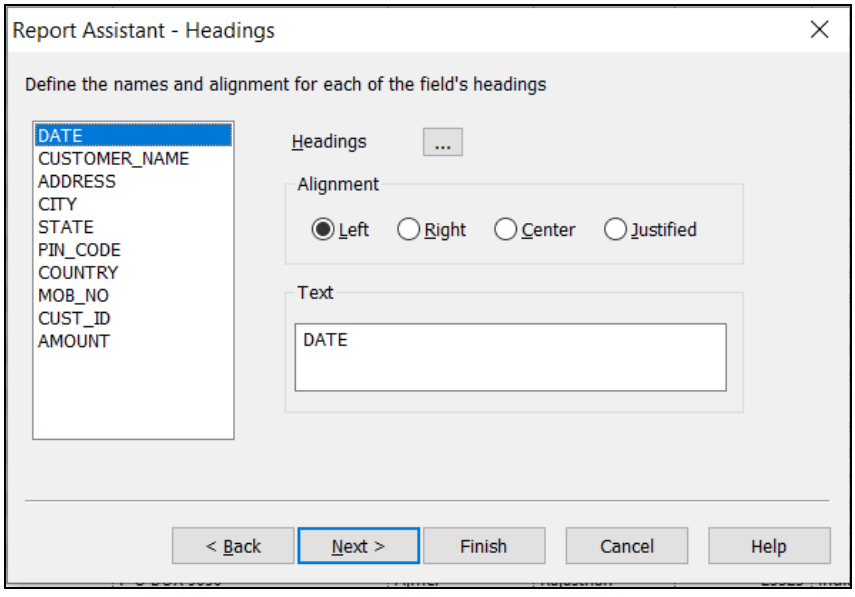


iii) Select an orientation for the report. The example uses horizontal.

iv) Choose between a new and existing report. The example is a new report.

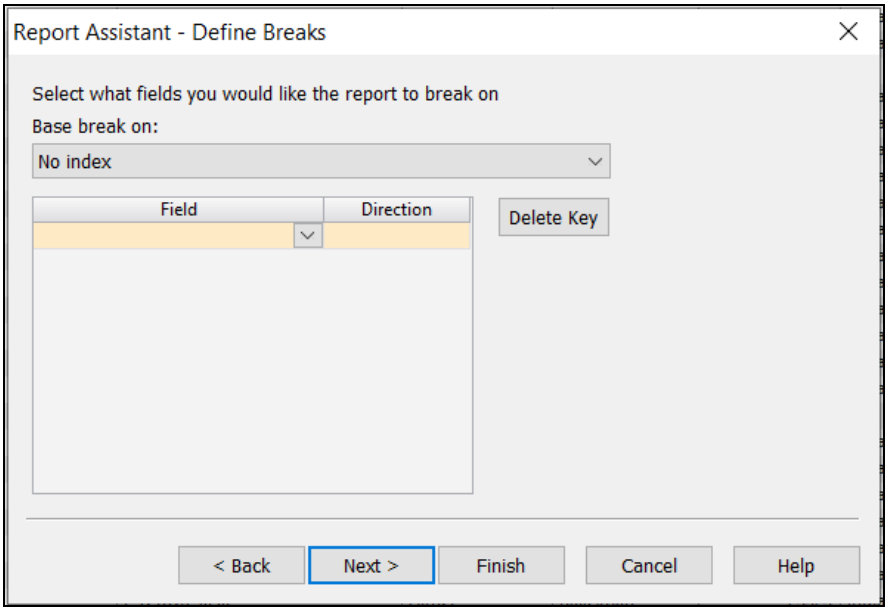
v) Select options for the report. The example uses the default options of not allowing headers to span multiple lines, showing record numbers, and showing the database name in the report.

vi) Click Next. We will see the Headings page of the Report Assistant shown below.



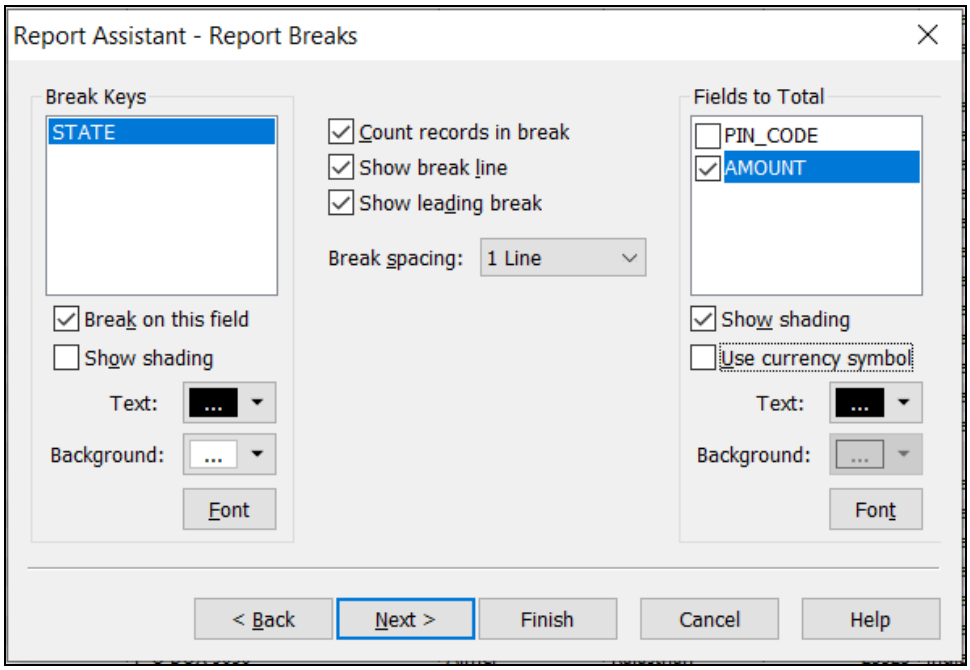
vii) Define a font, alignment, and title for each of the headings that will appear in the report. The example uses the default font and alignment for each of the fields. It uses the following heading titles (in order of the fields shown in the figure above): Date, Customer Name, Address, City, State, Pin Code, Country, Mobile no., Cust ID and Amount.

viii) Click Next. We will see the Define Breaks page of the Report Assistant as shown below. The figure shows the break used for the example report.



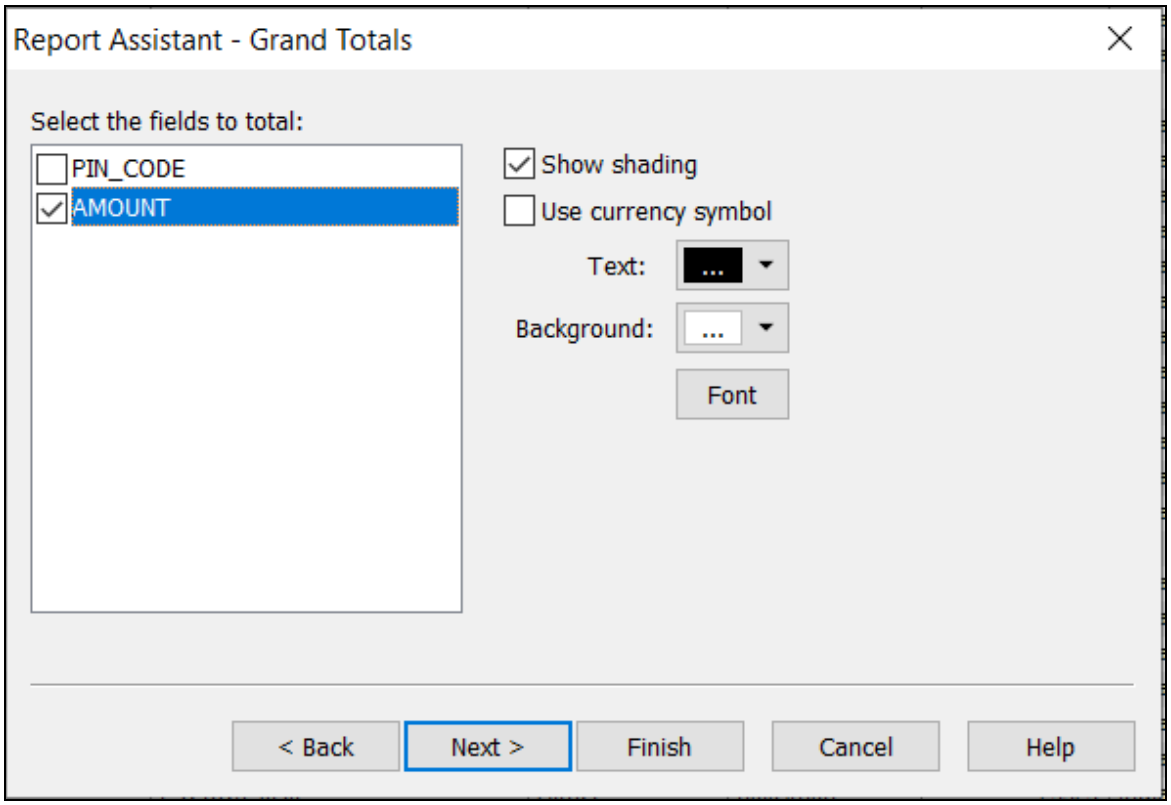
ix) Select a field name and sort order for each of the breaks we want to include in the report.

x) Click Next. We will see the Report Breaks page of the Report Assistant shown below. The figure shows the report break configuration for the example report.



xi) Specify how we want each of the break keys to appear in the report. We can define whether we want the break field to actually produce a break, the break colouring, and the break font. Each break can also use options such as counting the number of records in the break. If we want, we can also total the break and configure how the total appears.

xii) Click Next. You see the Grand Totals page of the Report Assistant shown below. IDEA automatically selects fields that could provide a grand total, even if these fields aren't part of the breaks we supplied earlier. The figure shows the grand totals configuration for the example report.

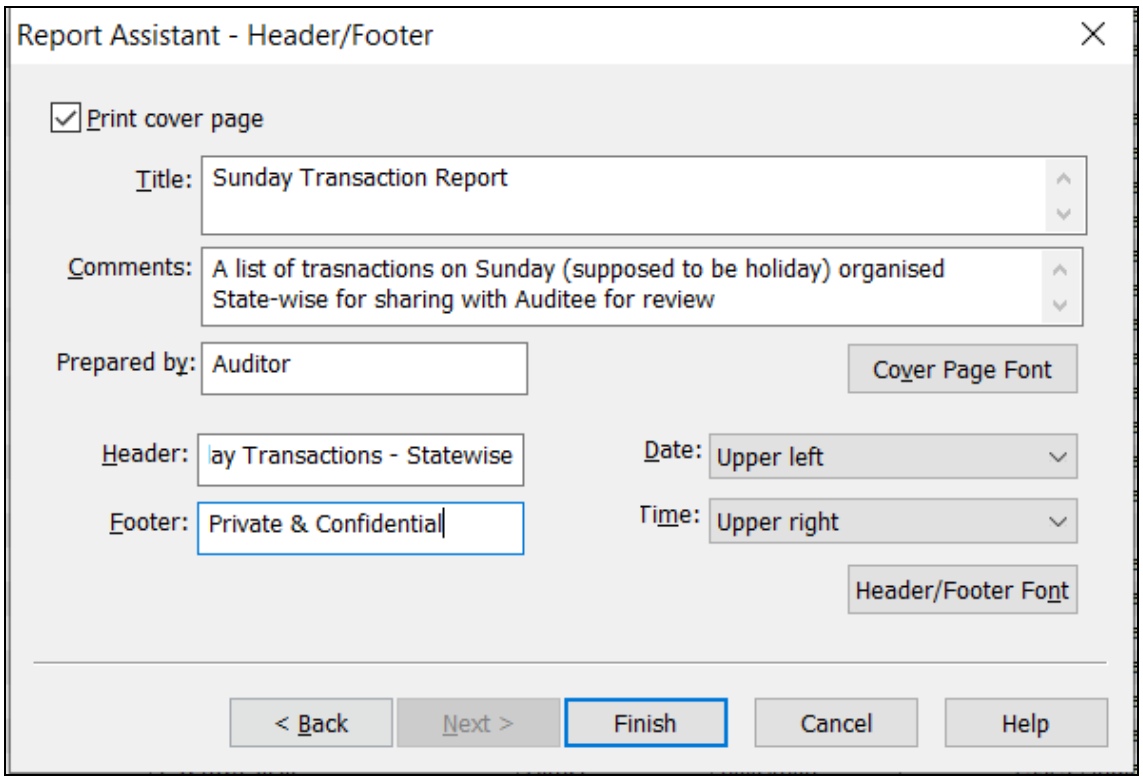


xiii) Provide any required grand totals for the report by checking the associated field. We can configure the grand totals to use shading and the currency symbol. It's also possible to change the text colouring (both foreground and background) and font.

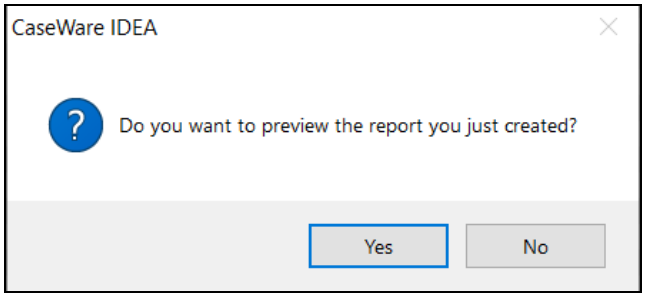
xiv) Click Next. We see the Header/Footer page of the Report Assistant as shown below. The figure shows the configuration for the example report.

xv) Determine whether we want a cover page. If so, select the Print cover page option. Define values for the Title, Comments, and Prepared by fields. Modify the cover page font if desired.

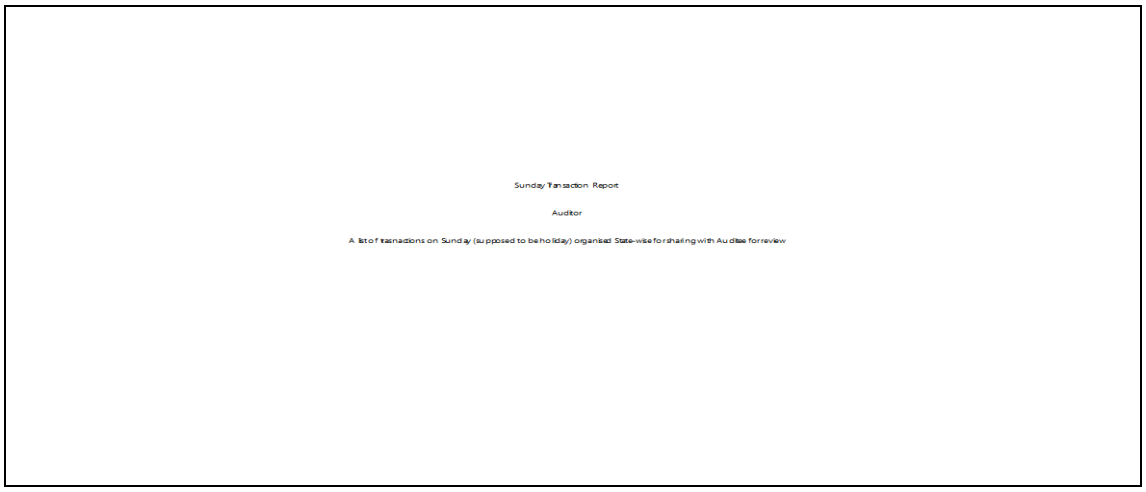
xvi) Add header and footer details as desired by modifying the Header, Footer, Date, and Time fields. You can modify the header and footer font, if desired.



xvii) Click Finish. IDEA asks whether we want to preview the report we created. If we preview the report, we can ensure it contains everything we want.



xviii) The figure belowshows how the example report appears in preview mode. We can use the single page view to see more details. It's also possible to zoom in and zoom out as needed.

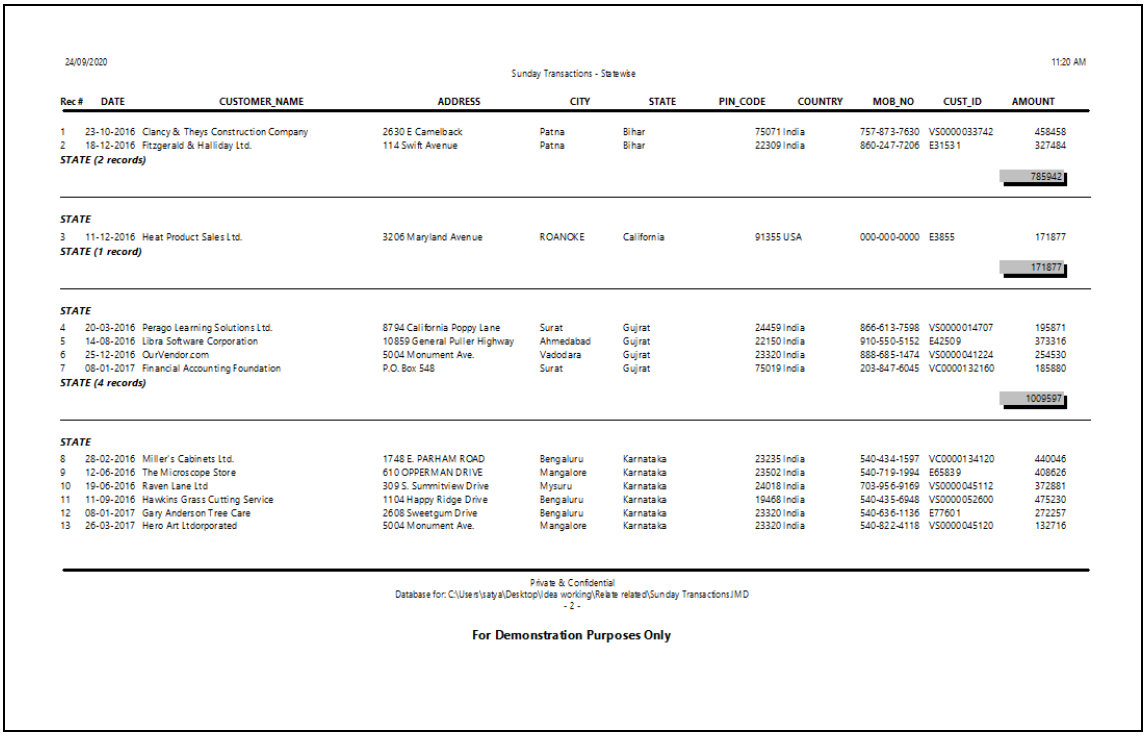


Sunday Transaction Report

Auditor

A list of trasnactions on Sunday (supposed to be holiday) organised State-wise for sharing with Auditee for review

(Cover page)



24/09/2020 Sunday Transactions - Statewise 11:20 AM

Rec #	DATE	CUSTOMER_NAME	ADDRESS	CITY	STATE	PIN_CODE	COUNTRY	MOB_NO	CUST_ID	AMOUNT
1	23-10-2016	Clancy & Theys Construction Company	2630 E Camelback	Patna	Bihar	75071	India	757-873-7630	VS0000033742	458458
2	18-12-2016	Fitzgerald & Halliday Ltd.	114 Swift Avenue	Patna	Bihar	22309	India	860-247-7206	E31531	327484
STATE (2 records)										785942
STATE										
3	11-12-2016	Heat Product Sales Ltd.	3206 Maryland Avenue	ROANOKE	California	91355	USA	000-000-0000	E3855	171877
STATE (1 record)										171877
STATE										
4	20-03-2016	Perago Learning Solutions Ltd.	8794 California Poppy Lane	Surat	Gujrat	24459	India	866-613-7598	VS0000014707	195871
5	14-08-2016	Libra Software Corporation	10859 General Puller Highway	Ahmedabad	Gujrat	22150	India	910-550-5152	E42509	373316
6	25-12-2016	OurVendor.com	5004 Monument Ave.	Vadodara	Gujrat	23320	India	888-685-1474	VS0000041224	254530
7	08-01-2017	Financial Accounting Foundation	P.O. Box 548	Surat	Gujrat	75019	India	203-847-6045	VC0000132160	185880
STATE (4 records)										1009597
STATE										
8	28-02-2016	Miller's Cabinets Ltd.	1748 E. PARHAM ROAD	Bengaluru	Karnataka	23235	India	540-434-1597	VC0000134120	440046
9	12-06-2016	The Microscope Store	610 OPPERMAN DRIVE	Mangalore	Karnataka	23502	India	540-719-1994	E65839	408626
10	19-06-2016	Raven Lane Ltd	309 S. Summitview Drive	Mysuru	Karnataka	24018	India	703-956-9169	VS0000045112	372881
11	11-09-2016	Hawkins Grass Cutting Service	1104 Happy Ridge Drive	Bengaluru	Karnataka	19468	India	540-435-6948	VS0000052600	475230
12	08-01-2017	Gary Anderson Tree Care	2608 Sweetgum Drive	Bengaluru	Karnataka	23320	India	540-636-1136	E77601	272257
13	26-03-2017	Hero Art Ltdorporated	5004 Monument Ave.	Mangalore	Karnataka	23320	India	540-822-4118	VS0000045120	132716

Private & Confidential
Database for: C:\Users\satya\Desktop\Idea working\Relate related\Sunday Transactions.MD
- 2 -

For Demonstration Purposes Only

(Page 2 of report)

At this point, we can save the report as a view by selecting **View > Views > Save View**. Enter a name for the view in the File name field of the Save As dialog box and click Save. Saving the report as a view makes it possible to select from multiple reports later. Reports aren't saved in the Properties window, so we do need to save them as views if we want to recall them later.

There isn't any way to programmatically select a report. To use a report later, select View > Views > Open View. Select the view you want to use in the Open dialog box and click Open. If we want to see the report before we print it, select File > Print> Print Preview. After we finish using the report, return IDEA to the default view by selecting View > Views > Reset.

Outputting Data in PDF Format

Using PDF format for the report, it has a number of advantages. The biggest advantage is that we can find a version of Adobe Acrobat for just about any platform made, which means that it doesn't matter what kind of computer the viewer uses. A PDF file tends to provide a more professional view than other kinds of report formats. It's relatively easy to add content if we have the full version of Adobe Acrobat (but not as easy as using Word).

We can also protect the finished document from changes. A PDF file also tends to be smaller than a Word document, so it requires less time to transfer using e-mail and other techniques. There are many other reasons to use PDF files, but these are the main reasons to use it for an IDEA database.

Outputting Data in Word Format

Reports created in Word tend to provide the best means of sharing information that we want others to comment on and possibly modify. It's a great way to create a report for a group scenario because Word provides a flexible way for multiple authors to contribute and keep track of who is contributing what. It's also possible to easily add graphics and other touches using Word that might be difficult when working with other report formats. Of course, the downside to using Word is that everyone must have a full copy in order to participate in the editing process.

If someone doesn't have Word, but still wants to review the report, they can use WordPad. The WordPad view of the information isn't as nice as the one provided in Word, but it does work.

10.15 FUNCTIONS

IDEA provides @Functions for performing operations such as Date, Arithmetic, Financial and Statistical calculations as well as Text searches. The @Functions are accessed through the Equation Editor. Quick help including the syntax, description, and an example of use for each of the @Functions is available when the @Function is highlighted. Additional and detailed help is provided in the Equation Editor Help system accessed through the Help button on the Equation Editor toolbar.

The list of functions is as follows:

@Function	Description
@Abs	Returns the absolute value of a numeric expression.
@Afternoon	Returns 0 if time is in the AM and 1 if time falls in the PM and -1 for an invalid time.
@Age	Calculates the number of days between two dates.
@AgeDateTime	Returns the number of seconds between two dates and times.
@AgeTime	Returns the number of seconds between two times.

@Function	Description
@AllTrim	Removes all leading and trailing spaces.
@Ascii	Provides the ASCII value of a character. (Not available in Unicode versions of IDEA.)
@Between	Determines if a numeric expression falls within a specific range.
@BetweenDate	Returns a number indicating whether a date value falls within a specified range (1) or not (0).
@BetweenTime	Returns a number indicating whether a time value falls within a specified range (1) or not (0).
@Bit	Identifies a bit value.
@BitAnd	Masks out unwanted bits.
@BitOr	Sets required bit.
@Chr	Provides the character equivalent of a specified ASCII code. (Not available in Unicode versions of IDEA.)
@CompareNoCase	Ignores uppercase letters when comparing expressions.
@CompIf	Determines if a record satisfies multiple criteria.
@Ctod	Converts character dates to IDEA Date format.
@Ctot	Converts a Character field containing time values stored as a string to a Time field with a Time format (HH:MM:SS).
@CurForm	Converts numeric value into a formatted text.
@CurVal	Converts formatted Character fields to Numeric fields.
@Date	Returns the present date.
@Day	Returns the day in a date expression.
@DaysToD	Converts a number of days since Jan. 1, 1900 to date format.
@Db	Calculates the fixed declining-balance depreciation for a specified period.
@Ddb	Calculates double declining-balance depreciation.
@Delete	Deletes a specified number of characters from a string.
@Dow	Returns the day of the week.
@Dtoc	Converts date expressions to character.
@DToDays	Reveals the number of days between Jan. 1, 1900 and a specified date.
@Dtoj	Converts dates to Julian format.

@Function	Description
@Exp	Calculates the exponent of a numeric expression.
@FieldStatistics	Returns the numeric value for a specified field statistic.
@FindOneOf	Finds the position of the first matching character in 2 strings.
@FinYear	Returns the financial year for a given date based on the year end.
@Format12HourClock	Returns a string representing time formatted as HH:MM:SS TT.
@Fv	Calculates the future value of an investment.
@GetAt	Returns the character that appears in a specified numeric position.
@GetNextValue	Returns the next value in the selected field.
@GetPreviousValue	Returns the previous value in the selected field.
@Hours	Returns the hours portion of a given time.
@If	Allows a choice of two results based on the evaluation of a condition.
@Insert	Inserts a string into an existing string.
@Int	Returns the integer portion of a numeric value.
@Ipmt	Calculates the interest payment for a given period.
@Irr	Calculates internal rate of return.
@IsBlank	Tests if a Character field is blank.
@IsFieldDataValid	Returns a 1 if the data in the field is valid, or a 0 if the data is invalid.
@Isin	Returns the starting position of a string within another string (case sensitive).
@Isini	Returns the starting position of a string within another string (NOT case sensitive).
@Jtod	Converts Julian dates to IDEA Date format.
@JustLetters	Returns a string with all the numeric characters removed.
@JustNumbers	Returns all the numbers (leading and trailing).
@JustNumbersLeading	Returns the leading numbers.
@JustNumbersTrailing	Returns the trailing numbers.
@LastDayofMonth	Returns the last day for any given month and year combination.
@Left	Returns the specified left-most characters in a string.
@Len	Returns the length of a string, including any trailing spaces.
@List	Determines which criteria in a list of values is met by an expression.

@Function	Description
@Log	Calculates natural logarithms.
@Log10	Calculates logarithm 10x.
@Lower	Converts all characters in a string to lowercase.
@Ltrim	Removes leading spaces from a string.
@Match	Determines which criteria in a list of values is met by an expression.
@Max	Returns the greater value of two numeric expressions.
@Mid	Extracts a portion of text from within a string.
@Min	Returns the smallest value of two numeric expressions.
@Minutes	Returns the minutes portion of a given time.
@Mirr	Calculates modified internal rate of return.
@Month	Returns the month in a date expression.
@NoMatch	Determines if an expression meets none of the criteria in a list of values.
@Npv	Calculates the net present value of an investment.
@Ntod	Converts a numeric expression into an IDEA Date format.
@Ntot	Converts a Numeric field containing time stored as seconds to a Time field with a Time format (HH:MM:SS).
@Pmt	Calculates a loan payment.
@Ppmt	Returns the principal amount of a loan payment.
@Precno	Returns the physical record number.
@Proper	Capitalizes the first letter of each word in a string.
@Pv	Returns the present value of an investment.
@Qtr	Returns 1-4 representing the quarter a given date falls in based on the specified year end.
@Random	Generates a random number.
@Rate	Calculates the interest rate of an investment/loan.
@Recno	Returns the logical record number (index sensitive).
@RegExpr	Matches character expressions using a complex set of rules. (Not available in Unicode versions of IDEA.)

@Function	Description
@Remove	Eliminates all instances of a specified character.
@Repeat	Repeats the first character of a string a specified number of times.
@Replace	Replaces a string or substring with another.
@Reverse	Reverses the order of characters in a string.
@Right	Isolates the specified right-most characters in a string.
@Round	Rounds to the nearest integer.
@Seconds	Returns the seconds portion of a given time.
@Seed	Sets the random number seed.
@SimilarPhrase	Measures the similarity between two specified phrases or Character fields.
@SimilarWord	Measures the similarity between two strings (either single words or character expressions) or Character fields.
@SimpleSplit	Extracts a segment of a character string that resides between the specified xth occurrence of a specified start character (or character sting) and the specified end character (or character string).
@Sln	Returns the straight-line depreciation of an asset.
@Soundex	Returns the sound code for a word. (Not available in Unicode versions of IDEA.)
@SoundsLike	Determines whether two words are phonetically alike. (Not available in Unicode versions of IDEA.)
@SpacesToOne	Strips spaces leaving only one space between words in a string.
@SpanExcluding	Returns the characters in a string that appear before any characters in a specified string.
@SpanIncluding	Returns the characters at the beginning of a string that match any character of a specified string.
@Split	Breaks a character string into segments separated by characters, such as spaces or commas, and returns a specified segment.
@Sqrt	Calculates a square root.
@Str	Converts numeric expressions to strings.
@Stratum	Groups records by interval.
@Strip	Removes all spaces, punctuation and control characters.
@StripAccent	Removes an accent from an accented character.

@Function	Description
@Syd	Returns the sum-of-years digit depreciation for an asset.
@Time	Returns the present time.
@Trim	Removes trailing spaces.
@Ttoc	Converts a time or number into a string with the HH:MM:SS format.
@Tton	Converts a Time field that has values stored in Time format (HH:MM:SS) to a Numeric field with the time value converted into seconds. This is the reverse of @Ntot.
@Upper	Converts all characters in a string to uppercase.
@Val	Converts a character expression to numeric.
@Workday	Returns 1 if a given date falls between Monday- Friday and 0 if the date falls on a Saturday or Sunday.
@Year	Returns the year in a date expression.

10.16 SUMMARY

In this chapter, we learned about extraction of data using Direct Extraction, Indexed Extraction, Top Records Extraction and Key Value Extraction. It is similar to filter option in Excel. Depending on nature of data, we need to use the extraction type. We learned about Summarization, which groups the records based on the fields specified. It's then possible to create a result with certain statistics, including sum, minimum, average, or maximum value for a specified field. We can do summarisation up to 8 levels.

The Stratification of data helps to create a view of data. Creating a view lets us build graphs, drill down into the data, and perform other analysis tasks. A database option allows to do more after the stratification. A stratified database contains a special STRATUM field that shows which value band holds a particular record. We use the Stratification task to place the data in layers or strata by a date, numeric, or character field. It's then possible to create a result with certain statistics, including sum, minimum, average, or maximum value for a specified field. It helps to look at the distribution of related data with greater ease.

The Aging feature helps to create the aging of the database from a specified date for up to 6 specified intervals. We can reduce the intervals as per our requirement. In order to use this analysis, the database must have at least one date field and one numeric field. We learned the usage of PivotTable. It is used to display specified data using a grid format. Pivot tables are helpful because it can help in summarizing the information or gain a new view of the data without becoming overwhelmed with detail. It creates a dynamic summary of the data in table format.

We saw that IDEA supports a number of chart and graph types to satisfy a variety of needs. The kind of chart or graph that we select determines the message that the output delivers. We can also use Visualisation and Discover feature for visual analysis of data.

We used the feature Relate in IDEA. The relate feature is like Lookup in Excel. There are 4 different types of Relate features in IDEA – Join, Visual Connector, Compare and Append. The Join Database is used to join two databases together based on a common field, or fields (called a key). We use the Visual Connector task to create a single database from two or more other databases. One of these databases must be the primary database. After we define the databases, we need to define the fields to include in the output database. The Compare feature helps to compare values from both the databases and gives the difference as an additional column by comparing Total field in both databases. The Append feature concatenate two or more IDEA databases into a new database.

We learned various options for sample selection in IDEA. We also learned about gap detection in data under various scenarios (Date and Invoice no.). This helps to identify the missing invoices and dates on which transaction not done. This helps in audit process. We learned about Duplicate Key option. It helps to identify duplicate items within a database. It has 3 options i.e. Detection, Exclusion and Fuzzy.

We learned about Benford's law, where this law analysis the digit occurrences. As per this law, the number consistently fall into a pattern with low digits occurring more frequently in the first position than larger digits. The mathematical tenet defining the frequency of digits is known as Benford's Law. In Benford's Law, number patterns are used to detect potential fraud, possible errors, manipulative biases, irregularities, etc. We also saw an example of usage of Macros in later part of the chapter.

REFERENCE

[1] https://idea.caseware.com

[2] Mastering IDEA Script – The Definitive Guide by John Paul Mueller, John Willey & Sons, Inc.

[3] IDEA Tutorial - IDEA 10

[4] Data analysis for auditors using CAAT by ICAI

[5] Fraud auditing using CAATT by Shaun Aghili, CRC Press

[6] Fraud and Fraud Detection – A data analytics approach by Sunder Gee, John Willey & Sons, Inc

PRACTICAL LAB EXERCISES AND CASE STUDIES

CHAPTER 2: EXCEL BASICS AND CUSTOMISATIONS

SMART COPY PASTE

Exercise 2.1

In an audit, your principal CA Chandiwala gives you a workbook containing four sheets for different quarters which he wants you to append into a single sheet so that that annual data can be analysed.

Note – For above use Exercise2.1.xlsx file from EXCEL folder.

For Result, please refer to Page No 94 of Chapter 2

Go to Special

Case Study 2.1

You are in a Forensic Audit Assisting your principal CA Chandiwala. Your client gives you a sheet which apparently looks right, you suspect some figures are fraudulent, also there are some error formulas in some cells, due to which unpivoting the data and analysis thereof is a challenge.

Your principal gives you a division wise quarterly sales report received from a client. as shown in Fig

You are to find inconsistencies in the sheet and answers to the following questions:

Are blank cells really blank?

Are any numbers stored as text?

Are there any formulas resulting in errors?

	A	B	C	D	E
1	**Dealer No.**	**Q1 19**	**Q2 19**	**Q3 20**	**Q4 20**
2	Dealer No. 1	228542	801452	824166	757331
3	Dealer No. 2	696929	686970	425651	242191
4	Dealer No. 3	885148		412479	263394
5	Dealer No. 4	279692	385973	#NAME?	461786
6	Dealer No. 5	285418	#DIV/0!	301274	240518
7	Dealer No. 6		295364	#DIV/0!	622029
8	Dealer No. 7	375196	232764	299778	621954
9	Dealer No. 8	318074	709511	695116	205276

Note – For above use Case Study 2.1.xlsx file from EXCEL folder.

For Result, please refer to Page No 88 of Chapter 2

CHAPTER 3: EXCEL- MAKING DATA CONSISTENT

TABLES

Exercise 3.1

In an Investigation your principal CA Chandiwala gives you an Excel sheet 3.1 where data Balance confirmations are being added periodically by various field assistants and the formatting is a mess, moreover he wants to maintain some consistency in formulas. He tells you to find a way in which formatting remains consistent and presentable even when data is added by different field assistants and formula could be added to consistently find the difference between Book balances and confirmations. You are a smart professional, is there a smart way to format this sheet?

Note – For above use Exercise 3.1.xlsx file from EXCEL folder.

For Result, please refer to Page No 128 of Chapter 3

DATA VALIDATION

Exercise 3.2

Your principal CA Chandiwala wants you to create two dropdown lists. The second list should be dependent on what is selected in the first one. He wants the first dropdown list with CA Partner-in-charge i.e. CA Chandiwala & CA Sonawala and independent List in 2nd cell, we want the dropdown list of Audit Managers from the respective lists.

Note – For above use Exercise 3.2.xlsx file from EXCEL folder.

For Result, please refer to Page No 143 of Chapter 3

Case Study

Your principal CA Chandiwala wants you to prepare a risk register shown in Figure. At the same times to maintain data consistency he wants you to ensure the following:

	A	B	C	D	E	F	G	H	I	J	K
1					RISK REGISTER					RISK RESPONSE	
2	Risk ID	Risk Title	Risk Statement	Date Risk Identified	Risk Originator	Risk Category	Probability 1-5	Impact 1-5	Total Score	Risk Response Strategy	Risk Response Plan Description
3											

Title	**Data Validations**
Risk ID	***Numeric***
Risk Title	***Text at least 20***
Date Risk Identified	***The day of entry***
Risk Category	***Only out of COMPLIANCE, INFOTECH, OPERATIONAL, REPUTATIONAL,***

	SECURITY
Probability	***Values 1 to 5***
Impact	***Values 1 to 5***
Risk Response Strategy	***Only out of ACCEPT, AVOID, MITIGATE, TRANSFER***

He asks you to find a way in which data should be restricted for entry as per above parameters You are a smart professional, is there a smart way to do it?

Note – For above use Case Study 3.1.xlsx file from EXCEL folder.

For Result, please refer to Page No 145 of Chapter 3

CHAPTER 4 : ORGANISE DATA USING EXCEL

SORTING CUSTOM LISTS

Exercise 4.1

In an audit, your principal CA Chandiwala gives you a Customer wise, month-wise list and he wants to do a Monthly Trend Analysis, but the list has the monthly columns arranged alphabetically. Your principal tells you to cut and paste columns and arrange them in the Calendar Year order. You are a smart auditor, is there a smart way to do it?

Note – For above use Exercise 4.1.xlsx file from EXCEL folder.

For Result, please refer to Page No 154 & 158 of Chapter 4

SUBTOTAL

Exercise 4.2

Your principal CA Chandiwala gives you an excel sheet containing date wise payment made to contractors as per Exercise 4.2 . He wants you to check for cases where total freight paid to any Contractor is more than Rs.1,00,000 so that it could check for TDS compliance under Income Tax.

Note – For above use Exercise 4.2.xlsx file from EXCEL folder.

For Result, please refer to Page No 177 of Chapter 4

CHAPTER 5 : EXCEL – FORMULA AND FUNCTIONS

Exercise 5.1 : Use of LEFT / MID / RIGHT Function

We have a list of dates, which are not compatible with our machine date format. We need to convert these dates in correct format using DATE, LEFT, MID and RIGHT functions.

Note: For above exercise, use File Chapter 5, sheet no. 5.1 and for answer, refer sheet no. 5.1 (A) in Excel Folder.

Exercise 5.2 & 5.3 : Use of CONCATENATE Function and "&" Operator

We have list of Indian cricketers. First Name and Full name in 2 separate columns. We need to connect them together in one cell. For this process, we need to use CONCATENATE and & operator.

Note: For **CONCATENATE** exercise, use sheet no. 5.2 and for answer, refer sheet no. 5.2 (A) in Excel Folder.

For **&** operator exercise, use File Chapter 5, sheet no. 5.3 and for answer, refer sheet no. 5.3 (A) in Excel Folder.

Exercise 5.4 : Use of FLASH FILL

Example 1 - Name split: We have list of names, which want to separate based on many options – First Name, Last Name, First Name (Upper), Last Name (Proper) and Full Name (Proper).

Example 2 - Name extraction: We have list of bank narrations. From these narrations, we want to extract names.

Note: For above exercises, use File Chapter 5, sheet no. 5.4 and for answer, refer sheet no. 5.4 (A) in Excel Folder.

Exercise 5.5 : Use of TEXT Function

We have a Sales transaction dump for FY2019-20. We want to identify on which day sales have been made. As per the company policy, there is Holiday on Sunday.

Note: For above exercises, use File Chapter 5, sheet no. 5.5 and for answer, refer sheet no. 5.5 (A) in Excel Folder.

Exercise 5.6 : Use of IF Function

We have a list of Individual persons and their income. We need to compute the tax for each person as per Flat Tax Rate and as per Incremental Tax rate assumption.

The Tax Rate slab is as follows:

0 to 2,50,000	NIL
2,50,001 to 5,00,000	5%
5,00,001 to 10,00,000	20%
> 10,00,000	30%

Note: For above exercises, use File Chapter 5, sheet no. 5.6 and for answer, refer sheet no. 5.6 (A) in Excel Folder.

Exercise 5.7 : Use of SUMIF, SUMIFS, COUNTIF and COUNTIFS Function

We have Sales information for 6 months. We need to find the following information:

1) Sum of Quantity sold to Sagar Bakshi
2) Sum of Quantity sold to Jayesh Thakur
3) Sum of Qunaitty sold if Quantity > 75
4) Sum of Qunaitty sold if Order < 3
5) Sum of Quantity sold to Sagar Bakshi when order >7
6) No. of Orders placed by Sagar Bakshi
7) No. of Orders placed by Sagar Bakshi when Quantity > 50

Note: For above exercises, use File Chapter 5, sheet no. 5.7 and for answer, refer sheet no. 5.7 (A) in Excel Folder.

Exercise 5.8 : Use of VLOOKUP, HLOOKUP, INDEX and MATCH Functions

We have Sales value in a tabular form (Customer-wise in rows and Zone wise – in column). We need to find sales value in the following combination using VLOOKUP, HLOOKUP, INDEX, MATCH and CHOOSE function:

		VLOOKUP	HLOOKUP	INDEX	CHOOSE
Product	**Zone**	**Sales**	**Sales**	**Sales**	**Sales**
A	West				
B	South				
A	North				
C	South				
D	West				
B	East				

Note: For above exercises, use File Chapter 5, sheet no. 5.8 and for answer, refer sheet no. 5.8 (A) in Excel Folder.

Exercise 5.9 : Use of DCOUNT Function

We have Employee Register, containing list of 10 employees, their rate, hours and Total Income.

We need to calculate following:

1) Give the count when Employee Name is “Ram” – exact match
2) Give the count when Employee Name is “Ram” – approximate match

3) Give the count when Employee Name is “Ram” – exact match and Rate > 60

4) Give the count when Employee Name is “Ram” – exact match and “Vijay” – approximate match

Note: For above exercises, use File Chapter 5, sheet no. 5.9 and for answer, refer sheet no. 5.9 (A) in Excel Folder.

Exercise 5.10 : Use of DSUM Function

We have Employee Register, containing list of 10 employees, their rate, hours and Total Income.

We need to calculate following:

1) Calculate Total Income when Employee Name is “Ram” – exact match

2) Calculate Total Income when Employee Name is “Ram” – approximate match

3) Calculate Total Income when Employee Name is “Ram” – exact match and Rate > 60 and Hours <> 340

4) Calculate Sum of Total Income when Employee Name is “Ram” – exact match and Rate > 60 and Hours <> 340 and for Employee Name is “Vijay” – approximate match and Rate > 75

Note: For above exercises, use File Chapter 5, sheet no. 5.10 and for answer, refer sheet no. 5.10 (A) in Excel Folder.

Exercise 5.11 : Use of LARGE and SMALL Function

We have Employee Register, containing list of 10 employees, their rate, hours and Total Income.

We need to calculate following:

1) Sum of Total Income for Top 3 Employees (based on Total Income)

2) Sum of Total Income for Last 3 Employees (based on Total Income)

Note: For above exercises, use File Chapter 5, sheet no. 5.11 and for answer, refer sheet no. 5.11 (A) in Excel Folder.

Exercise 5.12 : Use of AND & OR operator as an ARRAY function

We have Employee Register, containing list of 10 employees, their rate, hours and Total Income.

We need to calculate following:

1) Sum of Quantity sold to Dipak Bhasin when Order < 3

2) Sum of Quantity sold to Sagar Bakshi when order >7

3) Sum of Quantity sold to Sagar Bakshi OR Order >7

4) Count Unique no. of Customers

5) Unique no. of Orders (using FREQUENCY Function)

Note: For above exercises, use File Chapter 5, sheet no. 5.12 and for answer, refer sheet no. 5.12 (A) in Excel Folder.

Case Study 5.13 : Use of PivotTable for data analysis

The company has provided Sales Database in sheet 5.13, which contains 4 columns i.e. Date, Customer Name, No. of Orders and Quantity Sold. We need to create a PivotTable to compile the following information:

a) The number of Sales transactions, No. of Orders and Sales Amount for each Sales Person.

b) For each salesperson, the Amount in crore using Calculated Field Option

c) Total revenue by salesperson by year and month.

d) Zone-wise sales for each Sales Person and sort the data by Amount in crore in descending order.

Note: For above exercises, use File Chapter 5, sheet no. 5.13 and for answer, refer sheet no. 5.13 (A) in Excel Folder.

Exercise 5.14 Tax Calculations

Example 1: Compute Tax Payable by Mr. Shri Ram using VLOOKUP for the AY 2012-13 using Tax Rates given in the following table:

M/s. Being Computerized Co.				
Tax Computation Sheet				
Assessment Year 2012-13		**For**	**Male Less < 60 Yrs**	
		GTI	**Rate**	**Amt+**
	Shri Ram	1	0%	0
		180,000	10%	0
	1,000,00			
	0	500,000	20%	
SHE Cess		800,000	30%	
Total Tax		**Slab**		
		Tax Rate		
		Amt+		

Note – For above use Exercise5.14_Income Tax_CS.xlsx file from NEWCS_EXCEL folder

CHAPTER 6 : ANALYSING AND VISUALISING DATA IN EXCEL

CONDITIONAL FORMATTING

Exercise 6.1

In a forensic investigation your principal CA Chandiwala gives you a sheet Exercise 6.1.xlsx which contains some amounts relating to some customers it is suspected that some of these figures are in fact text that is why not reflected in the total amount. You are required to find which of these amounts is not a number. You are a smart auditor, is there a smart way to locate Numbers stored as text so that they are not included in the total?

Note – For above use Exercise 6.1.xlsx file from EXCEL folder.

For Result, please refer to Page No 295 of Chapter 6

EXCEL CHARTS

Exercise 6.2

Your principal CA Chandiwala has given you data relating to Unitwise Expenditure and Budget. It is suspected that in some Units expenditure is much beyond the Budget allocated, He wants to investigate those units. You have to make a presentation on Expenditure vs Budget so that defaulting units could be investigated.

Note – For above use Exercise 6.2.xlsx file from EXCEL folder.

For Result, please refer to Page No 301 of Chapter 6

PIVOT TABLES

CASE STUDY 6.3

As a part of the investigation your principal CA Chandiwala has given you data relating to power consumption in different units of a company for various dates as per Case Study 6.1.xlsx He wants you to analyse the power consumption on Unit wise & Weekday wise.

After you submit your results to your principal CA Chandiwala showing Unitwise and Daywise Power consumption, your principal wants to know the monthly unit wise power consumption.

Your principal CA Chandiwala, on receiving the report in the previous case study sees that consumption of power for April for Brake is less compared to other months & Units, He wants you to find details of the underlying data so that he can look for the reason if any.

After preparing the report CA Chandiwala wants you to do a Weekday wise, Month wise, and also unit wise Power Consumption and He also wants the report of each unit on a separate Excel sheet.

You are a smart auditor how will you proceed smartly?

Note – For above use Case Study 6.1.xlsx file from EXCEL folder.

For Result, please refer to Page No 306 onwards of Chapter 6

Exercise 6.4: Case Study of Payroll Frauds

You are performing internal audit of ABC Company and have decided to perform the following tests to confirm whether fraud has been committed in payroll. The tests with sample data, results and additional test to be performed are given here.

Example 1:-

Identify whether salary has been paid only to employees in employee master file.

Files to use:

1. Employee_6.4 file.
2. Payroll_6.4 file

Tests to Perform:

Compare Employee ID's in Employee file with Employee ID's on Payroll.

Steps to perform Test:

Step 1: In the Employee file, use the following Excel formula:

=VLOOKUP (B2, Payroll! $B$2:$B$17, 1)

Emp_Id	emp_first	emp_lastr	DOB	emp_fath	emp_mot	Address	City	State	Phone_Nc	Mobile_N	Email_Id	Pincode	country
1	Vikas	Patel	########	Ram Kuma	Maya	12, Laxmi	Indore	Madhya P	1234	9578481	vik@gmai	135001	India
2	Ankur	Mehra	########	Deepak	Nilima	23, Krishn	Delhi	Delhi	5678	8976755	ank@yahc	135002	India
3	Anoj	Kumar	########	Lalit	Anjali	22, Saket	Delhi	Delhi	3333	5463284	ano@redi	135003	India
4	Priya	Singh	########	Nirmal	Shilpa	A-29, Shiv	Surat	Gujrat	5555	7895468	priy@gma	171025	India
5	Surekha	Gupta	########	Manoj	Shanti	Sector 7, N	Noida	Uttar Prad	8888	7895462	sur@redif	201101	India
6	Ashwini	Jain	########	Pradeep	Pramila	7, Kailas N	Pune	Maharash	6954	7845127	ashw@yal	411001	India
7	Mayur	Kumar	########	Dev	Shubha	M - 25, Se	Mumbai	Maharash	7777	1234567	may@red	400003	India

Step 2: Use Excel's Auto-fill function to apply the formula to all the records.

Inference:

- The results show all rows where ID's do not match with the payroll file.
- **(#N/A)** are records that exist in the Employee Master file but not in the Payroll file indicating that some employees have not been paid salary.
- Employees who have been terminated (Employee ID: **JD** & **AD**) should not have appeared in the Payroll file but they do and have possibly been paid salary even though they are no longer with the company.

Step 3: In the Payroll file, use the following Excel formula:

=VLOOKUP (B2, Employee_6.4! $B$2:$B$16, 1, FALSE) Step 4: Use Excel's Auto-fill function to apply the formula to all the records.

Inference:

- The results show all rows where ID's do not match.
- **(#N/A)** are records that exist in the Payroll file but not in the Employee Master file indicating that salary has been paid to employees that are no longer with the company.

Example 2:-

Identify whether any employees are vendors.

Files to use:

1. Employee_6.4 file.
2. Vendor_6.4 file

Tests to Perform:

Compare Employee Addresses to Vendor Addresses

Steps to perform Test:

Step 1: In the **Employee** file, use the following Excel formula:

=VLOOKUP (F2, Vendor_6.4!$C$2:$C$12,1,FALSE)

Step 2: Use Excel's **Auto-fill** function to apply the formula to all the records.

Inference:

The results show employees whose address matches the address of vendors indicating that these employees may have created fictitious companies to defraud their employer.

Example 3:-

Identify whether there are Ghost employees to whom salary has been paid.

File to use:

1. Payroll_6.4 file.

Tests to Perform:

Compare Employee Addresses to Vendor Addresses

Steps to perform Test:

Identify employees who have paid salary but don't have an address.

Step 1: In the **Payroll** file, click on the **Data** menu and select the **Filter** function.

Emp_Id	emp_first	emp_lastr	DOB	Emp Addr	emp_fath	emp_mot	Address	City	State	Phone_Nc	Mobile_N	Email_Id	Pincode	country
1	Vikas	Patel	########		Ram Kum	Maya	12, Laxmi	Indore	Madhya P	1234	9578481	vik@gmai	135001	India
2	Ankur	Mehra	########		Deepak	Nilima	23, Krishn	Delhi	Delhi	5678	8976755	ank@yahc	135002	India
3	Anoj	Kumar	########		Lalit	Anjali	22, Saket	Delhi	Delhi	3333	5463284	ano@redi	135003	India
4	Priya	Singh	########		Nirmal	Shilpa	A-29, Shiv	Surat	Gujrat	5555	7895468	priy@gma	171025	India
5	Surekha	Gupta	########		Manoj	Shanti	Sector 7, N	Noida	Uttar Prad	8888	7895462	sur@redif	201101	India
6	Ashwini	Jain	########		Pradeep	Pramila	7, Kailas N	Pune	Maharash	6954	7845127	ashw@yal	411001	India
7	Mayur	Kumar	########		Dev	Shubha	M - 25, Se	Mumbai	Maharash	7777	1234567	may@red	400003	India

Step 2: Click on the drop down box on the bottom-right of the "**Employee Address**", "**Employee ID**", **etc.** cell. Click on the "**Select All**" check box to de-select all the options. Then, select only the "**(Blanks)"** option. Click OK.

Inference:

The results show all records where employee address or the employee ID is blank.

Additional tests:

- Please check whether employees in employee master file are without address.
- Please identify employees whose address is different in employee master and payroll file.
- Re-compute net salary in payroll to identify errors in payroll computation.

Example 4:-

Identify whether there are any duplicate payment to same employee.

File to use:

1. Payroll_6.4 file.

Tests to Perform:

Identify duplicate employee Id.

Steps to perform Test:

Step 1: In the **Payroll_6.4** file, select the column to perform duplicate check (**Employee ID**) click on the **Home** tab and select the **Conditional Formatting** function. Select **Highlight Cell Rules** and select the **Duplicate Values**... function.

Emp_Id	emp_first	emp_lastr	DOB	Emp Addr	emp_fath	emp_mot	Address	City
1	Vikas	Patel	########		Ram Kum	Maya	12, Laxmi	Indore
2	Ankur	Mehra	########		Deepak	Nilima	23, Krishn	Delhi
3	Anoj	Kumar	########		Lalit	Anjali	22, Saket	Delhi
4	Priya	Singh	########		Nirmal	Shilpa	A-29, Shiv	Surat
5	Surekha	Gupta	########		Manoj	Shanti	Sector 7, N	Noida
6	Ashwini	Jain	########		Pradeep	Pramila	7, Kailas N	Pune
7	Mayur	Kumar	########		Dev	Shubha	M - 25, Se	Mumbai
2	Aman	Kumar	########		Parveen	Pramila	7, Kailas N	Pune

Step 2: In the dialog box that pops up, click OK.

Inference:

The results show duplicate payments where same employee is been paid twice for the same month.

Additional Tests:

- Verify whether there are two employees are having same phone no.
- Generally, limit is related to seniority of scale. Hence, higher the salary, higher would be limit. Compare limit as per department with salary and identify whether there is any inconsistency.
- Summarise salary of payroll as per department.

Exercise 6.5: Case Study of Purchase frauds

New

Purchases are a critical area which is prone to fraud and hence auditor has to perform tests to confirm whether data is correct and complete and to test for potential areas of fraud. The tests with sample data, results and additional test to be performed are given here.

Example 1:-

Identify whether there are any employees who are vendors by verifying common address.

Files to use:

1. Employee_6.4 file.
2. Vendor_6.4 file

Test to Perform:

Compare Employee address in Employee file with Vendor address in vendor master.

Steps to perform Test:

Step 1: In the **Vendor_6.4** file, use the following Excel formula:

=VLOOKUP (C2, Employee_6.4!$F$2:$F$16,1,FALSE)

Step 2: Use Excel's **Auto-fill** function to apply the formula to all the records.

	A	B	C	D	E
1	**Vendor ID**	**Name**	**Vendor Address**	**PAN**	**Address Check**
2	PD/11240	10 FRONT	22956 GROEN WOHLD	AVEAF2242N	#N/A
3	PD/10288	COMBINED AGENCIES & SUPPLIES	16 OPP LAL BUNGLOW	AGBEA5235H	#N/A
4	PD/10102	ARPITH COMPUTERS SERVICES PVT LTD	63 SHANKAR LANE	ABWRG525N	#N/A
5	PD/10825	BD SOLUTIONS PVT. LTD	734 JATASHANKAR DOSA ROAD	AQTVZ726GF	#N/A
6	PD/10555	GENIE CONTROL SYSTEMS LTD.	176B VALJI LADHA ROAD	AUGWS7934B	176B VALJI LADHA ROAD
7	PD/11289	MURTHY COLLEGE OF NURSING	25/31, III FLOOR, DR A M RD	AJDWE6113Y	#N/A
8	PD/10293	TOUCH SONIC	793, GOWSHALA ROAD	AOPGRX6513Z	793, GOWSHALA ROAD
9	PD/10287	VP POWERS SOLUTIONS	FLAT NO.303-A A WING ROAD	APPAM6512Q	#N/A
10	PD/11224	CELLTRON ELECTRONICS	53 A EVEREST VIHAR	AOWOQ7934W	#N/A
11	PD/10654	STAG ENTERPRISES	RUTU TOWERS, HIRANANDANI ESTATE	AONVV9835F	#N/A
12	PD/11147	M. K KHAN	SCHEME NO 6, NANDA PATKAR ROAD	AIOER4965Z	#N/A

Inference:

The results show vendors whose address matches the address of employees indicating that these may be fictitious vendors created by employees.

Example 2:-

Identify whether there are any employees who are vendors by verifying common address.

Files to use:

1. Purchase_6.4 file.

Test to Perform:

Find duplicate invoice numbers in purchases.

Step 1: In the **Purchases** file, select the column to perform duplicate check (**Invoice Number**) click on the **Home** menu and select the **Conditional Formatting** function. Select **Highlight Cell Rules** and select the **Duplicate Values…** function.

Step 2: In the dialog box that pops up, click OK.

Step 3: To view only the duplicates, click on the **Data** menu and select the **Filter** function.

Step 4: Click on the drop down box on the bottom-right of the "**Invoice Number**" cell, select the **Filter by Colour** option and click on the colour under either **Filter by Cell Colour** or **Filter by Font Colour.**

	A	B	C	D	E	F	G	H	I	J
1	ac	Invoice Number	Vendor	Name	Vendor Address	Product ID	Unit Pri	Quanti	Amou	Employee ID
21	29-04-2014	PINV/1020	PD/10287	VP POWERS SOLUTIONS	FLAT NO.303-A A WING ROAD	G/40	500	30	15000	VR
22	30-04-2014	PINV/1020	PD/10288	COMBINED AGENCIES & SUPPLIES		G/28	700	20	14000	DA
23	30-04-2014	PINV/1020	PD/10287	VP POWERS SOLUTIONS	FLAT NO.303-A A WING ROAD	G/40	500	40	20000	AB
94	02-08-2014	PINV/1093	PD/10825	BD SOLUTIONS PVT. LTD	734 JATASHANKAR DOSA ROAD	G/35	250	50	12500	LN
95	05-08-2014	PINV/1093	PD/10654	STAG ENTERPRISES	RUTU TOWERS, HIRANANDANI ESTATE	G/35	250	30	7500	SS
96	05-08-2014	PINV/1093	PD/10654	STAG ENTERPRISES	RUTU TOWERS, HIRANANDANI ESTATE	G/28	700	40	28000	GK
97	06-08-2014	PINV/1093	PD/10555	GENIE CONTROL SYSTEMS LTD.	176B VALJI LADHA ROAD	G/31	1200	20	24000	SS
98	07-08-2014	PINV/1093	PD/11289	MURTHY COLLEGE OF NURSING	25/31, III FLOOR, DR A M RD	G/30	400	50	20000	DA
151	12-10-2014	PINV/1150	PD/10293	TOUCH SONIC	793, GOWSHALA ROAD	G/42	1000	20	20000	GK
152	12-10-2014	PINV/1150	PD/10102	ARPITH COMPUTERS SERVICES PVT LTD	63 SHANKAR LANE	G/30	400	40	16000	LN
153	12-10-2014	PINV/1150	PD/10654	STAG ENTERPRISES	RUTU TOWERS, HIRANANDANI ESTATE	G/35	250	30	7500	GK

Inference:

The results highlights duplicate invoice numbers in the Purchase file with the same invoice number for purchases made from different vendors on different dates which is highly unlikely. This could also indicate records where the same invoice has been split into multiple invoices and entered to overcome purchase limits put on employees.

Example 3:-

Identify Favourable treatment of Vendors and classifying vendors by value

Files to use:

1. Purchase_6.4 file.

Test to Perform:

Step 1: In the **Purchases** file, select all the data in the worksheet (**Ctrl + A**), then click on the **Insert** tab and select

Pivot Table. In the resulting dialog box, click OK.

Step 2: Drag "**Name**" from the PivotTable Fields dialog box and drop it under **Rows**. Drag "**Amount**" and drop it under

Values

4	Row Labels	Sum of Amount
5	10 FRONT	199000
6	ARPITH COMPUTERS SERVICES PVT LTD	446000
7	BD SOLUTIONS PVT. LTD	400500
8	CELLTRON ELECTRONICS	603500
9	COMBINED AGENCIES & SUPPLIES	337000
10	GENIE CONTROL SYSTEMS LTD.	259500
11	M. K KHAN	344000
12	MURTHY COLLEGE OF NURSING	275000
13	STAG ENTERPRISES	1987000
14	TOUCH SONIC	572500
15	VP POWERS SOLUTIONS	349000
16	(blank)	74000
17	**Grand Total**	**5847000**

Inference:

- The results show the total amount purchased from each vendor indicating favourable treatment given to one vendor (**STAG ENTERPRISES**) over all others.
- Purchases with no vendor name are also displayed under the **(blank)** category. These are all purchases where the name of the vendor has not been entered into the Purchase register. To view the purchases made by each vendor, double-click the name of that vendor.

Exercise 6.6 – Pivot Tables

Refer to the file 'Exercise 6.6 - Pivot Tables'

Consider the data in the above file. Insert a Pivot Table for this data and answer the following questions:

1. How many material codes belong to class 0? And how many belong to class 3?
2. How many material codes have either PAIL or DRUM as their Labeling Unit?
3. Apply Report filter on 'Inventory Type' feld. Select filtering value as Cool but fgd. How many material codes are shown?

CHAPTER 7: PROTECTION AND DATA TOOLS

PROTECTION FEATURES

Exercise 7.1

Your principal CA Chandiwala wants you to protect the cells where you have put Formulas so that nobody can accidentally change formulas, and no one can see the formulas.

Note – For above use Exercise 7.1.xlsx file from EXCEL folder.

For Result, please refer to Page No 372 of Chapter 7

DATA TOOLS-CONSOLIDATE

Exercise 7.2

Your principal CA Chandiwala has the expenditure of 4 quarters in 4 worksheets Q1, Q2, Q3 & Q4 and they are having the same layout but in each of the sheets order of expense head is different, consolidating them manually is a challenge and he wants to consolidate them into annual expenditure.

Note – For above use Exercise 7.2.xlsx file from EXCEL folder.

For Result, please refer to Page No 391 of Chapter 7

WHAT-IF-ANALYSIS

Exercise 7.3

As a Mortgage Analyst in a Bank you propose to give a Housing Loan to a prospective client on the following terms

- *Loan amount Rs. 40,00,000.*
- *ROI Annual 12%*
- *Term 240 Months*

You calculate the EMI but your client says he cannot spare more than 42000 every month. You want to reset the term to achieve a targeted EMI of 42000 pm.

Note – For above use Exercise 7.3.xlsx file from EXCEL folder.

Exercise 7.4: Importing Data from MS Access in MS Excel

Note – Here use Exercise7.4_XYZ_CORP.dbf file from EXCEL_AUDIT_TOOL folder.

Text File format. Output

Book1 - Notepad

File Edit Format View Help

```
COUNTRY CUSTOMER_N      INVOICE_NU      DATE    REP     PROD_CODE       UNIT_PRICE
U.S.A.  "SIMPSON, MIROSLAW"     1000867 10-08-01        102     05      5.99    700
FRANCE  "BRANDSTACK, PEKKA"     1000018 26-07-01        111     03      35.1    561
CHINA   "PROYNOV, DENISLAV"     1000079 12-04-01        112     05      5.99    4703
COSTA RICA      "KWONG, KWAI HEUNG"     1000865 05-11-01        122     05      5.99
U.S.A.  "SIMPSON, MIROSLAW"     1000498 26-10-01        105     05      5.99    600
MEXICO  "TREUTEN, KATHERIN"     1000211 07-05-01        120     04      105.69  97
COSTA RICA      "KWONG, KWAI HEUNG"     1000248 16-05-01        122     05      5.99
MEXICO  "TREUTEN, KATHERIN"     1000368 19-02-01        119     05      5.99    95
BELGIUM "LAMMERANT, BENOIT"     1000718 09-11-01        103     05      5.99    107
MEXICO  "TREUTEN, KATHERIN"     1000227 16-03-01        119     05      5.99    111
BARBADOS        "SANCHEZ, LEONARDO"     1000298 06-03-01        125     05      5.99
FINLAND "BRANDSTACK, PEKKA"     1000769 05-11-01        108     05      5.99    122
CANADA  "PROYNOV, DENISLAV"     1000327 24-04-01        123     05      5.99    1582
CHINA   "KWONG, KWAI HEUNG"     1000879 14-09-01        113     05      5.99    1175
DENMARK "CHRISTENSEN, VITA"     1000329 10-04-01        107     05      5.99    30
U.S.A.  "SIMPSON, MIROSLAW"     1000413 30-03-01        102     04      105.69  400
U.S.A.  "VILAT, PHETSAMONE"     1000535 15-10-01        104     05      5.99    1431
FINLAND "BRANDSTACK, PEKKA"     1000221 06-04-01        108     05      5.99    34
U.S.A.  "SIMPSON, MIROSLAW"     1000362 10-04-01        102     05      5.99    25
```

Exercise 7.5: Consolidation of Data Case Study Output

Here students are requested to find the sum of expenses of April, May and June as per head of account to prepare total expenses of the quarter.

Step 1: Create a blank worksheet, click on the Data tab and select the Consolidate function.

Step 2: In the dialog box, click on the formula box beside Browse (), select the Exp_April file and select all the records in column B and column C (including the column headings). Now click on the formula box again to return to the dialog box (you should have the formula Exp_April! $BA$1:$C$30) in the box under Reference: Now, click Add in the dialog box to add this reference to the box under All references.

Repeat this process for the **Exp_May** file (**Exp_May! $B$1:$C$30**) and **Exp_June** file (**Exp_June! $B$1:$C$30**) Step 3: Tick the check box beside **Left column**. Click OK.

	A	B
1	**Head**	
2	Washing allowance/Charges	1895
3	Conveyance allowance to Award Staff	19236
4	Rent Paid for Office Premises	171858
5	Rent Paid for Other Premises	49276
6	Other Taxes	17183
7	Lighting	56536
8	Generator Expenses (Diesel etc)	127931
9	Printing stationery/ Sale Proceeds of old records	43023
10	Postage	1297
11	Courier Charges	888
12	Telephones	2532
13	Amount Paid for leased lines/ISDN conections	3159
14	Repairs to Furniture fixtures Other than Cars jeeps	42647
15	Travelling Expenses	31502
16	Travelling Expenses - Clerks	9836

Inference:

The results show a consolidated, quarterly view of all expenses as per head of account.

Exercise 7.6: Sensitivity Analysis Case Study with MS-Excel What-If-Analysis

M/s. ABC Ltd. is considering a project with an Initial Cost of Rs.10,00,000/- with Selling Price per Unit of Rs.60 and Cost per unit of Rs.40. The expected Sales Volume in Year 1, 2 and 3 are 20,000, 30,000 and 30,000 units respectively. The Cost of Capital for the company is 10%. You are required to:

1. Compute NPV and IRR of the Project and
2. Undertake What-If-Analysis as follows:
    a) Pivot Table Report of Scenarios of NPV for Cost of Capital of 10%, 15%, 20%, 25% and 30%
    b) Compute IRR of the Project through Goal Seek
    c) Generate a Data Table Report of NPV of the Project using Cost of Capital as 10%, 15%, 20%, 25% and 30%

Note – For above use Exercise7.6_Sensitivity Analysis_CS.xlsx file.

Exercise 7.7: Scenario Analysis Case Study with MS-Excel What-If-Analysis

M/s. ABC Co. Ltd. Currently sells 1000 units of a product @ Rs.10/- each that costs Rs.6/- each. The Company projections estimate following growth rates per annum for Sales @ 15%, Costs @ 10% and Selling Price @ 2%. You are first required to compute the NPV of the project and then undertake Sensitivity Analysis using "Scenario Manager" in "What-If-Analysis" of MS-Excel for following Scenarios:

Particulars	Best	Worst	Average
Tax Rate	50%	30%	70%
Sales Growth	15%	20%	15%
Cost Growth Rate	10%	2%	10%
Selling Price Growth Rate	2%	8%	2%

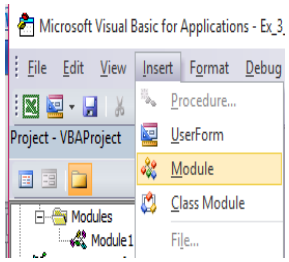
Note – For above use Exercise7.7a_Scenario Analysis_CS.xlsx file.

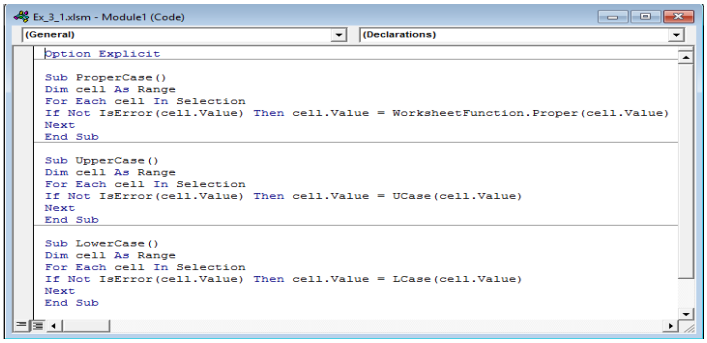
CHAPTER 8

GOOD TO KNOW FEATURES

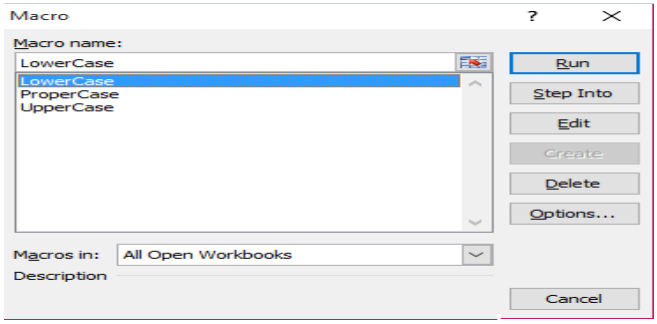
Exercise 8.1 Output

Step-1: Open Exercise8.1.xlsx file from EXCEL folder and press ALT + F11. Step-2: Insert new Module and type the below code as shown in below figure.





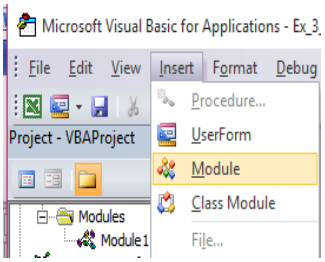
Step-3: Press ALT + F8 and run the macro.

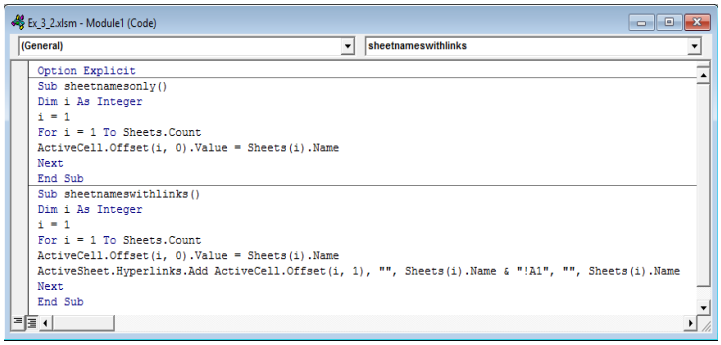


	A	B
1		Output
2	Lower Case	this is lower case
3	Upper Case	THIS IS LOWER CASE
4	Proper Case	This Is Lower Case

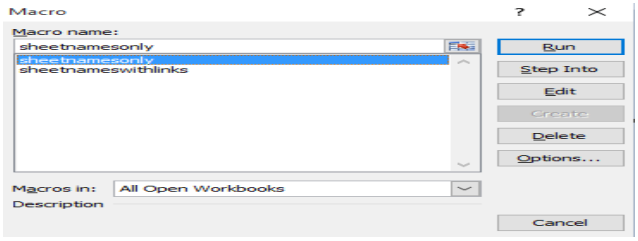
Exercise 8.2 Output

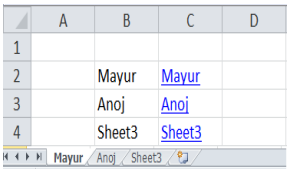
Step-1: Open Exercise 8.2.xlsx file from EXCEL folder and press ALT + F11. Step-2: Insert new Module and type the below code as shown in below figure.





Step-3: Press ALT + F8 and run the macro.





Exercise 8.3: Cleaning dates from ERP downloaded "DD.MM.YYYY" format in to Excel acceptable "MM/DD/YYYY" format

Output

Pre-requisite step before applying date-based Sorting, Filtering and applying Date formulas such as =EDATE (), EOMONTH (), TEXT () etc.

Applying date based formulas on dates stored in DD.MM.YYYY format will result in errors as Excel reads date in MM/DD/YYYY format. Sort, Filter, Pivot Table and other such techniques will not work correctly unless the date is corrected as per Excel Standards i.e. MM/DD/YYYY format.

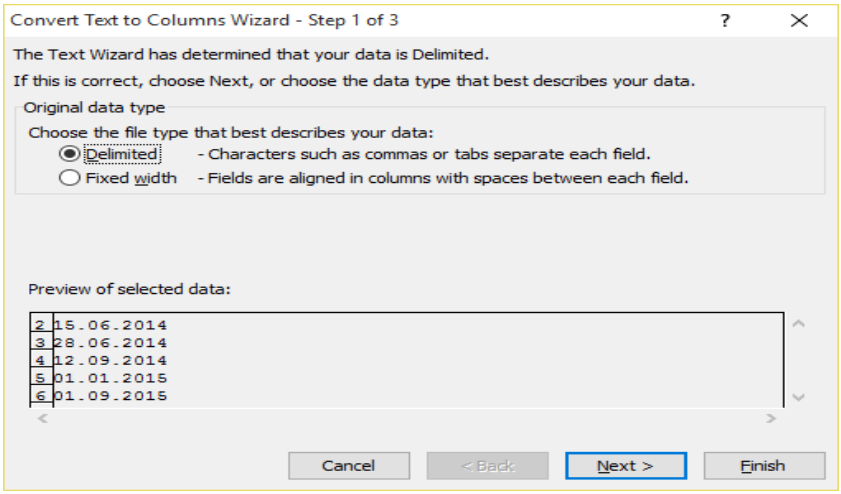
	A	B
1	**Invoice Date**	**Date Formula - NA**
2	15.06.2014	#VALUE!
3	28.06.2014	#VALUE!
4	12.09.2014	#VALUE!
5	01.01.2015	#VALUE!
6	01.09.2015	=EOMONTH(A6,0)

Step 1: Select affected dates

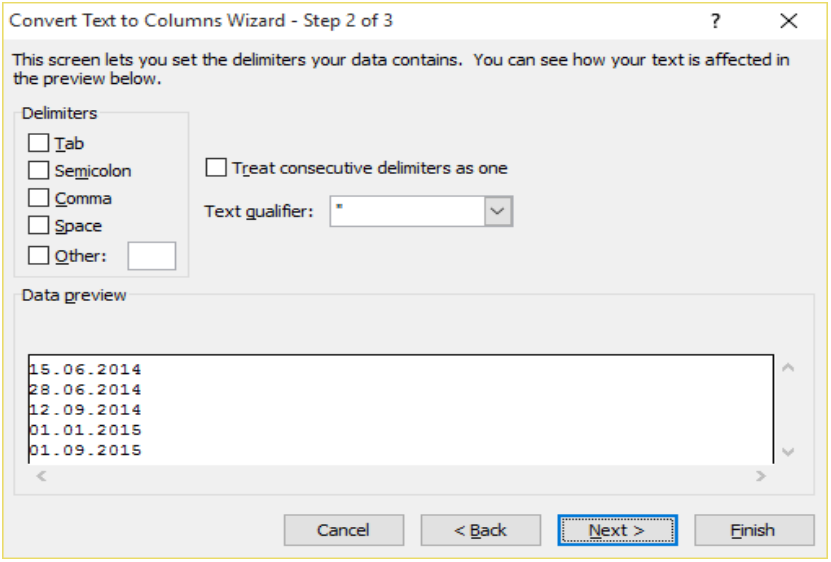
	A	B
1	**Invoice Date**	**Date Formula - NA**
2	15.06.2014	#VALUE!
3	28.06.2014	#VALUE!
4	12.09.2014	#VALUE!
5	01.01.2015	#VALUE!
6	01.09.2015	#VALUE!

Step 2: Select Data Tab->Text to Columns

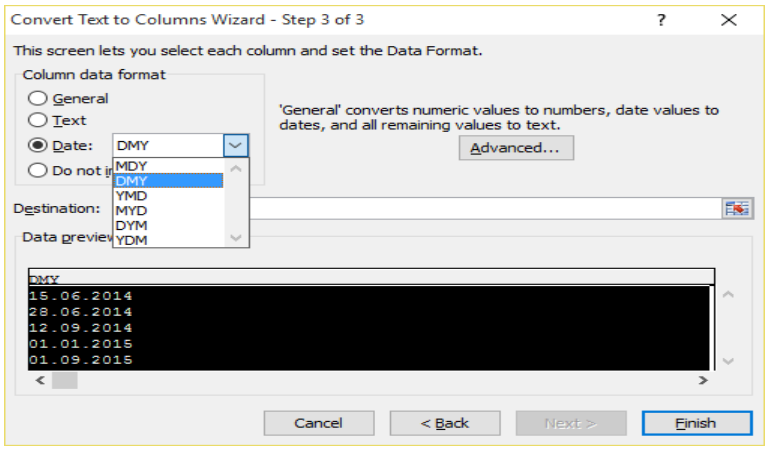
Step 3: Text to Columns (Step 1 of 3) Delimited



Step 4: TEXT TO COLUMNS (Step 2 of 3) -> Turn-off all checkboxes [e.g. Tab, Comma etc.



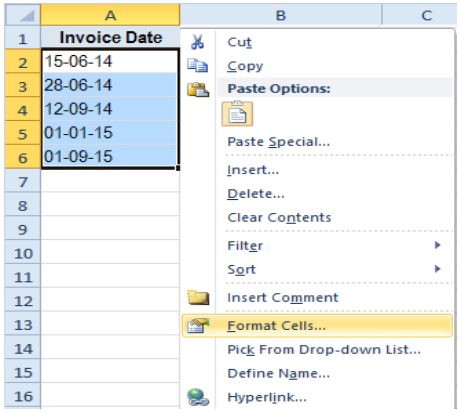
Step 5: TEXT TO COLUMNS (Step 3 of 3) -> “Date” drop-down list -> DMY -> “Finish” button

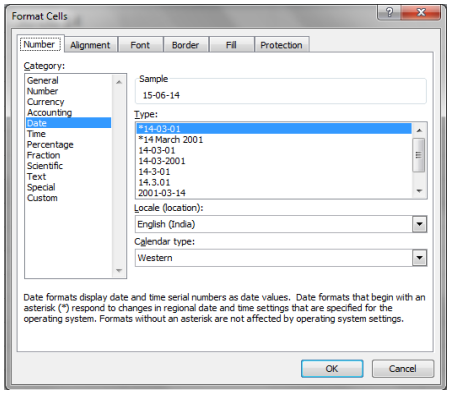


Step 6: Result: Internally, all dates turn into MM/DD/YYYY [refer cell A6]. Although, the skin (presentation) can be modified to DD-MMM-YY as shown in the subsequent step.

	A	B
1	Invoice Date	Date Formula - NA
2	15-06-14	30-06-14
3	28-06-14	30-06-14
4	12-09-14	30-09-14
5	01-01-15	31-01-15
6	01-09-15	30-09-15

Step 7: Select cells -> Right-click -> "Format Cells" -> "Date" option -> Choose appropriate format for display- presentation of dates





CHAPTER 9 & 10

COMPUTER ASSISTED AUDIT TECHNIQUES (CAATS)

DATA ANALYSIS IN CAATs (IDEA)

Exercise 10.1 : Use of various Extraction options in CAAT (IDEA)

The Statutory Auditor of the XYZ Ltd. has received the Sales Outstanding file, which has various columns i.e. Date, Customer Name, Address, City, State, Pin Code, Country, Mob_No, Cust_ID And Amount. The Auditor planned to use CAAT (IDEA) for extracting various samples/significant data from the file for the purpose of further analysis. After loading the file in IDEA, they want to perform following analysis on it:

1) Using Direct Extraction option, extract the list of Customers having Amount Outstanding > 3,00,000. The output file should contain only 3 columns i.e. Customer Name, City and Amount

2) Re-run the same query and extract the Amount Outstanding > 4,50,000. The output file should contain only 4 columns i.e. Customer Name, City, Cust ID and Amount

3) Using Indexed Extraction option, extract list of Customer having Amount Outstanding >3,00,000 and <=5,00,000 in Mumbai city. The output file should contain only 3 columns i.e. Customer Name, City and Amount

4) Using Top Records Extraction option, extract list of Top 3 Customer based on Amount Outstanding. The output file should contain only 3 columns i.e. Customer Name, Country and Amount.

5) Re-run the Top 3 Customer query and extract Top 3 Customers based on Amount Outstanding in each Country. The output file should contain only 3 columns i.e. Customer Name, Country and Amount.

6) Re-run the Top 3 Customer Country-wise query and extract Top 3 Customers based on Amount Outstanding > 10 Lakhs in each Country. The output file should contain only 3 columns i.e. Customer Name, Country and Amount.

7) Using Key Value Extraction option, extract the list of customer based on State and City as Key Value. Then from the list of combinations, select any 2 combinations only for output. The output file should contain only 4 columns i.e. Customer Name, City, State and Amount.

8) Using Append option, append both the files extracted using Key Value Extraction option one below the other.

9) Extract 25 random samples from the Sales Outstanding file. The output file should contain only 3 columns i.e. Customer Name, City and Amount.

Note: For above exercise, use file name CAAT-1_Sales Outstanding. For answer, please refer module.

Exercise 10.2 : Use of Summarisation option in CAAT (IDEA)

The Statutory Auditor also wants to summarise the Sales Outstanding file. After loading the file in IDEA, they want to perform following analysis on it:

1) Summarise the Sales Outstanding file based on Customer Name.

2) Summarise the Sales Outstanding file based on Country and then by State.

Note: For above exercise, use file name CAAT-1_Sales Outstanding. For answer, please refer module.

Exercise 10.3 : Use of Stratification option in CAAT (IDEA)

We have a Sales Outstanding file, having various columns i.e. Date, Customer Name, Address, City, State, Pin Code, Country, Mob_No, Cust_ID And Amount. After loading the file in IDEA, we need to perform following analysis on it:

1) Stratify the Sales Outstanding file based on Amount column in the following range:

Lower Limit	Upper Limit
0	2,00,000
2,00,000	4,00,000
4,00,000	6,00,000
6,00,000	11,00,000
11,00,000	16,00,000
16,00,000	24,00,000
24,00,000	26,00,000

Extract list of instances, which are below lower limit of Rs. 0.

2) Re-Run the above Stratification query and now group this stratification based on Country.

3) Stratify the Sales Outstanding file based on Date column and Total for Amount column in the following range: Month (first 3 interval by 3 month, the next 2 interval by 6 month, last interval by 3 month).

Note: For above exercise, use file name CAAT_1B_Sales Outstanding. For answer, please refer module.

Exercise 10.4 : Use of Ageing option in CAAT (IDEA)

We have a Sales Outstanding file, having various columns i.e. Date, Customer Name, Address, City, State, Pin Code, Country, Mob_No, Cust_ID And Amount. After loading the file in IDEA, we need to perform following analysis on it:

1) Prepare aging based on Date column as on June 30, 2016. Amount field to be used for Total. The proposed Aging interval is : Days (Default) with the default intervals (50, 100, 150, 200, 300).

2) Re-Run the above query and change the interval to Months (3, 6, 9, 12 and 15)

Note: For above exercise, use file name CAAT_1B_Sales Outstanding. For answer, please refer module.

Exercise 10.5 : Use of PivotTable and Chart option in CAAT (IDEA)

We have a Sales Outstanding file, having various columns i.e. Date, Customer Name, Address, City, State, Pin Code, Country, Mob_No, Cust_ID And Amount. After loading the file in IDEA, we need to perform following analysis on it:

1) Prepare a PivotTable based on Country and State for summarising the data based on Amount column. The output file should have following columns: Sum of Amount, Average of Amount, Count of Amount, Minimum Amount and Maximum Amount. Arrange all these columns vertically.

2) Prepare a State-wise summary based on Amount column and make a Bar chart based on number of records

3) Also visualise the data based on Visualisation option and Discover option as explained in Module.

Note: For above exercise, use file name CAAT_1B_Sales Outstanding. For answer, please refer module.

Exercise 10.6 : Use of Relate the Table option in CAAT (IDEA)

We have 2 files. First file is Sales Outstanding file, having various columns i.e. Date, Customer Name, Address, City, State, Pin Code, Country, Mob_No, Cust_ID And Amount. Second file is Customer Credit Limit file, having 2 columns i.e. Customer Name and Credit Limit.

After loading the file in IDEA, we need to perform following analysis on it:

1) Prepare a summary of both files based on Customer Name. We will use Sales Outstanding Summary as Primary file and Customer Credit Limit summary as Secondary file. Here Customer Name field will be used as Match Key Field.

2) Using Join option:

   a. Extract list of records which are Matching in both summary files

   b. Extract list of all records in Primary file

   c. Extract list of records which have no Secondary match

   d. Extract list of records which do not have any Primary match

   e. Extract list of records in both files

3) Using Visual Connector option:

   a. Extract list of records which are Matching in both summary files

   b. Extract list of all records in Primary database

4) Using Compare option, compare both summary files based on Customer Name. The Total filed will be Amount column in Primary file and Credit Limit in Secondary file for output.

Note: For above exercise, use file name CAAT_1B_Sales Outstanding and CAAT – 2_Credit Limit file. For answer, please refer module.

Exercise 10.7 : Use of Gap Detection and Duplicate Key option in CAAT (IDEA)

The Statutory Auditor has received 2 files from XYZ Ltd i.e. Sales Outstanding File and Sales Database file. The Auditor wants to find the following:

- Whether the company has deleted/removed invoices from the entire Sales Database file.
- Whether the transactions done on holidays as well
- Is there any duplicate invoice, which can be done to inflate sales
- Whether there is any customer manipulation i.e. same customer has been issued multiple customer Id. This can be done to manipulate the individual threshold limit for outstanding amount.

Hence, using the CAAT (IDEA), they want to identify the following from Sales Database file:

a) List of missing invoices

b) List of dates on which no transaction took place

c) List of Duplicate invoices

The Auditor also want to find the list of Customers having same mobile number but their Cust ID is different in Sales Outstanding file.

Note: For above exercise, use file name CAAT_1B_Sales Outstanding and CAAT – 3_Sales Database file. For answer, please refer module.

Exercise 10.8 : Use of Benford's Law option in CAAT (IDEA)

We have a Sales Database file, having various columns i.e. INV_NO., INV_DATE, SALESREP_NO., CUSTNO., PROD_CODE, UNIT_PRICE, QTY, SALES_BEF_TAX, TAX and SALES_PLUS_TAX.

After loading the file in IDEA, we need to perform following analysis on it:

1) Perform 1st Digit Test and 2nd Digit Test on Sales Data file. The field to be analysed is SALES_BEF_TAX. Evaluate the output in Graphical form as well as in Tabular form.

2) Extract the list of Suspicious and Highly Suspicious Transactions. Try to evaluate the reason of marking the transactions as Suspicious and Highly Suspicious by the IDEA.

Note: For above exercise, use file name CAAT-4_Benford Law - Sales Data file. For answer, please refer module.