We at The Data Monk hold the vision to make sure everyone in the IT industry has an equal stand to work in an open domain such as analytics. Analytics is one domain where there is no formal under-graduation degree and which is achievable to anyone and everyone in the World.

We are a team of 30+ mentors who have worked in various product-based companies in India and abroad, and we have come up with this idea to provide study materials directed to help you crack any analytics interview.

Every one of us has been interviewing for at least the last 6 to 8 years for different positions like Data Scientist, Data Analysts, Business Analysts, Product Analysts, Data Engineers, and other senior roles. We understand the gap between having good knowledge and converting an interview to a top product-based company.

Rest assured that if you follow our different mediums like our blog cum questions-answer portal www.TheDataMonk.com , our youtube channel - The Data Monk, and our e-books, then you will have a very strong candidature in whichever interview you participate in.

There are many blogs that provide free study materials or questions on different analytical tools and technologies, but we concentrate mostly on the questions which are asked in an interview. We have a set of 100+ books which are available both on Amazon and on The Data Monk e-shop page

We would recommend you to explore our website, youtube channel, and e-books to understand the type of questions covered in our articles. We went for the question-answer approach both on our website as well as our e-books just because we feel that the best way to go from beginner to advance level is by practicing a lot of questions on the topic.

We have launched a series of 50 e-books on our website on all the popular as well as niche topics. Our range of material ranges from SQL, Python, and Machine Learning algorithms to ANN, CNN, PCA, etc.

We are constantly working on our product and will keep on updating it. It is very necessary to go through all the questions present in this book.

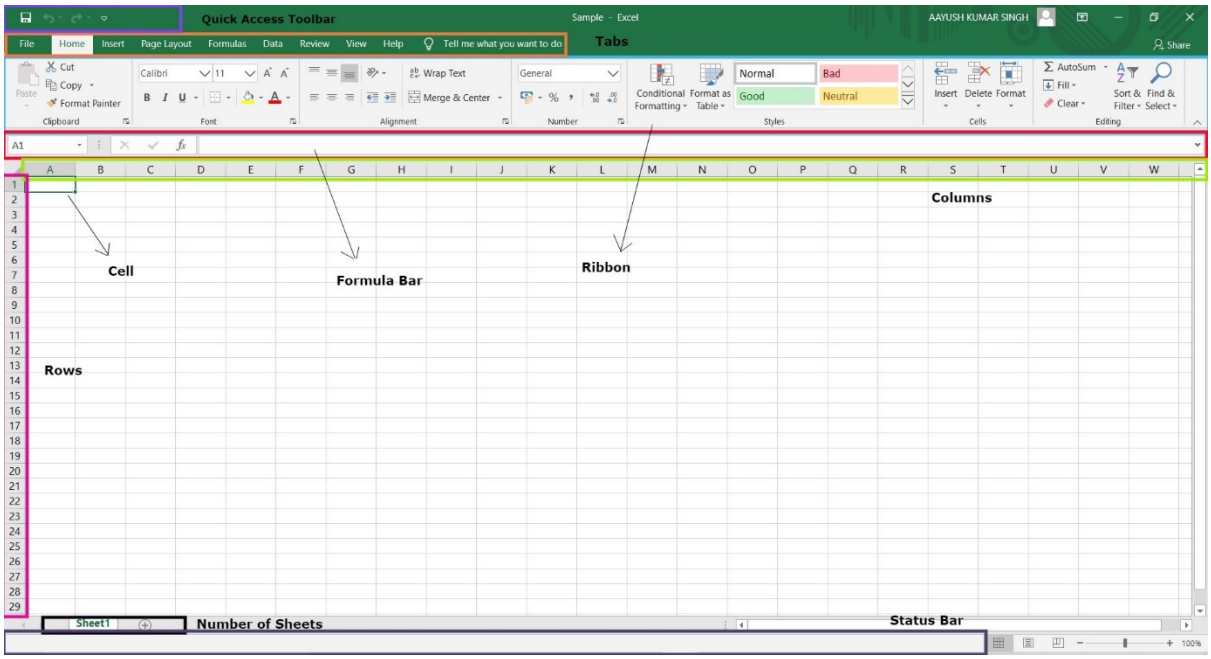
Give a rating to the book on Amazon, do provide your feedback and if you want to help us grow then please subscribe to our Youtube channel.

MS EXCEL TUTORIAL

The tutorial is performed on 2016 Version of MS Excel, there might be a few differences in terms of locations of icons and implementation methods for different versions or Mac users

MS Excel is one of the most widely used software for data entry, data retrieval and data analysis. It is used by most of the large MNCs for tackling with humongous datasets or to maintain important records. Excel has its uses ranging from simple data collection to complex data handling and analyzing. With growing Modernization, Excel continues to remain an important Tool for Big Business and firms and a requisite for any job seekers in the Data Analysis industry.

Getting Used to the Interface



Above shown is a picture of the interface of MS Excel, a few portions are marked in the above picture, that are as follows,

- **Ribbon** – It contains almost all the import icons related to the functionalities in MS Excel
- **Tabs** – Each tab contains its different set of icons for different functions, like as shown above, **'HOME'** is opened which contains the basic functionalities like cut, copy, bold, alignment, text options and many more.

One can close the ribbon by double clicking on the already opened tab, it creates more space for the spreadsheet

- **Quick Access Toolbar** – It contains some items that you may want to access immediately, it can be customized (with the help of small dropdown (triangle) present in the toolbar or through other options in the properties of an icon) as per one's needs.
- **Columns** – Columns in excel are represented by Alphabets (A, B, C … Z, AA, … AZ, …)
- **Rows** – Rows are represented by numbers (1, 2, 3, …)

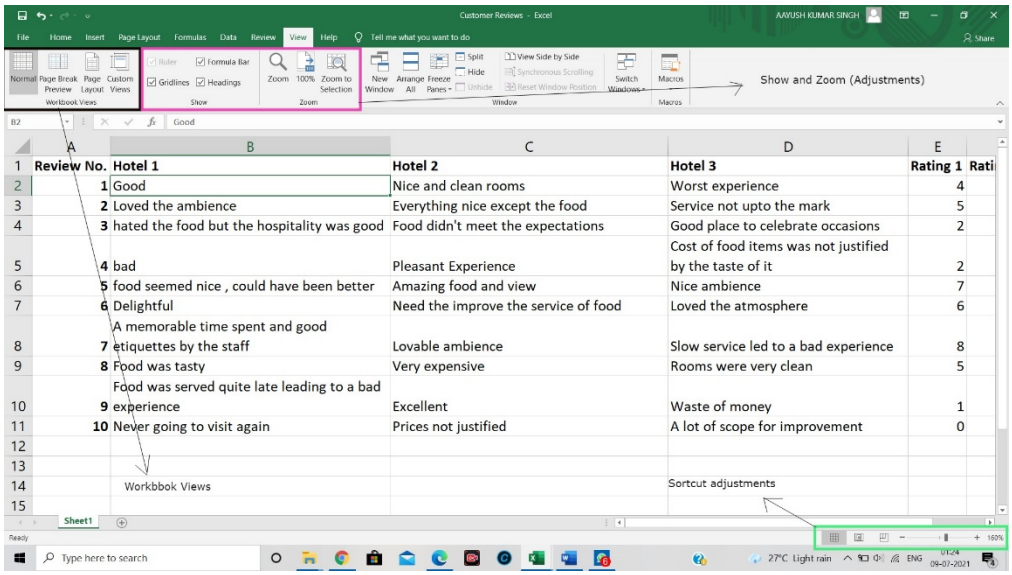
Excel has in total 10,48,576 rows and 16,384 columns

- **Cell** – A Cell is the location where one enters the required data. Its location is determined by the crossover of rows and columns, as in this case cell **'A1'** is selected.
- **Formula Bar** – It shows which cell is presently selected and what is the value entered in it.
- **Number of sheets** – We open an excel **'Workbook'** which contains **'Worksheets'**. Here, we can see the different worksheets in a Workbook and notice the one we are presently working on.

****One can rename, color, arrange the sheets in any required sequence as per the needs, just by right-clicking on the sheet and selecting the required options****

- **Status Bar** – Incase of any numerical entries over a range of cells in a row or column, selecting the entries will show basic mathematical calculations like **count, sum and average** in the Status Bar.

View Tab

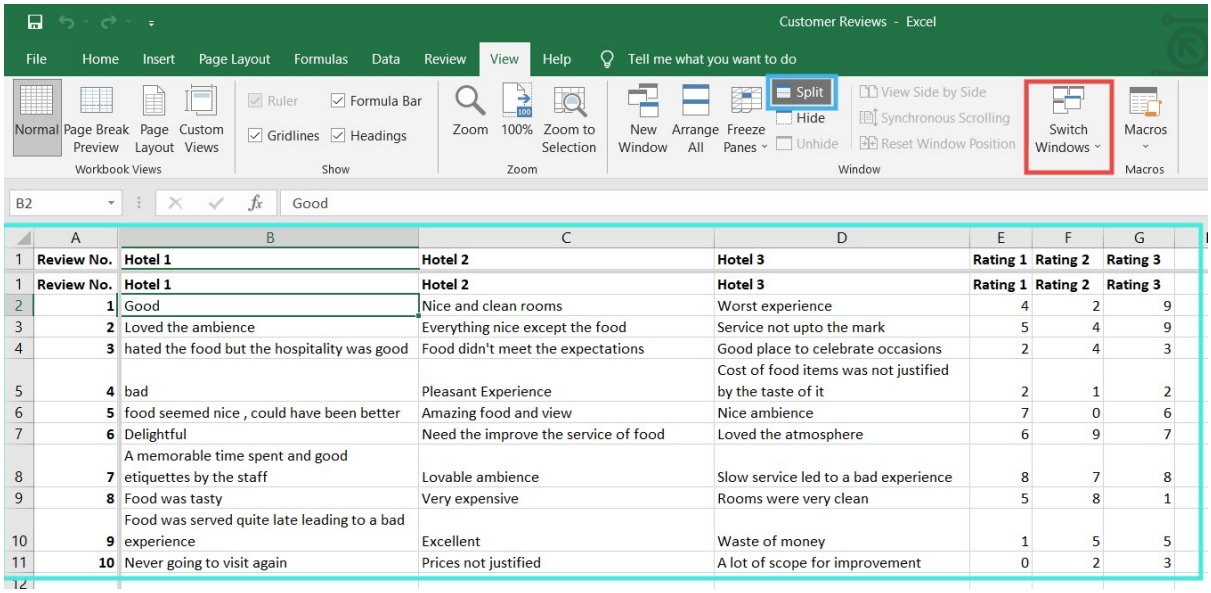


	A	B	C	D	E
1	Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1
2	1	Good	Nice and clean rooms	Worst experience	4
3	2	Loved the ambience	Everything nice except the food	Service not upto the mark	5
4	3	hated the food but the hospitality was good	Food didn't meet the expectations	Good place to celebrate occasions	2
5	4	bad	Pleasant Experience	Cost of food items was not justified by the taste of it	2
6	5	food seemed nice , could have been better	Amazing food and view	Nice ambience	7
7	6	Delightful	Need the improve the service of food	Loved the atmosphere	6
8	7	A memorable time spent and good etiquettes by the staff	Lovable ambience	Slow service led to a bad experience	8
9	8	Food was tasty	Very expensive	Rooms were very clean	5
10	9	Food was served quite late leading to a bad experience	Excellent	Waste of money	1
11	10	Never going to visit again	Prices not justified	A lot of scope for improvement	0

- **Workbook views** – This section of the view tab shows the various ways in which we can view our worksheet or workbook. It has the following options:
  a) Normal View
  b) **Page Break View** (Shows where the pages will break on printing)
  c) **Page Layout view** (Shows how printed document will look)
  d) Custom Views
- **Show and Zoom (Adjustments)** – Show section allows us to perform various miscellaneous functions on our worksheet with respect to presentation of **Gridlines, Headings** and **Formula Bar**. The Zoom sections as the name suggests caries out the magnification requirements of the user with the help of options present in it.
- Also, at the lower right-hand corner of our workbook you can notice that we have shortcuts of various commands like **Page Break Preview, Normal View, Page Layout View and a ruler for our magnifying purposes.**

This portion really comes in handy for fast operations

SPLIT SCREEN AND SWITCH WINDOWS



	A	B	C	D	E	F	G
1	Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3
1	Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3
2	1	Good	Nice and clean rooms	Worst experience	4	2	9
3	2	Loved the ambience	Everything nice except the food	Service not upto the mark	5	4	9
4	3	hated the food but the hospitality was good	Food didn't meet the expectations	Good place to celebrate occasions	2	4	3
5	4	bad	Pleasant Experience	Cost of food items was not justified by the taste of it	2	1	2
6	5	food seemed nice , could have been better	Amazing food and view	Nice ambience	7	0	6
7	6	Delightful	Need the improve the service of food	Loved the atmosphere	6	9	7
8	7	A memorable time spent and good etiquettes by the staff	Lovable ambience	Slow service led to a bad experience	8	7	8
9	8	Food was tasty	Very expensive	Rooms were very clean	5	8	1
10	9	Food was served quite late leading to a bad experience	Excellent	Waste of money	1	5	5
11	10	Never going to visit again	Prices not justified	A lot of scope for improvement	0	2	3

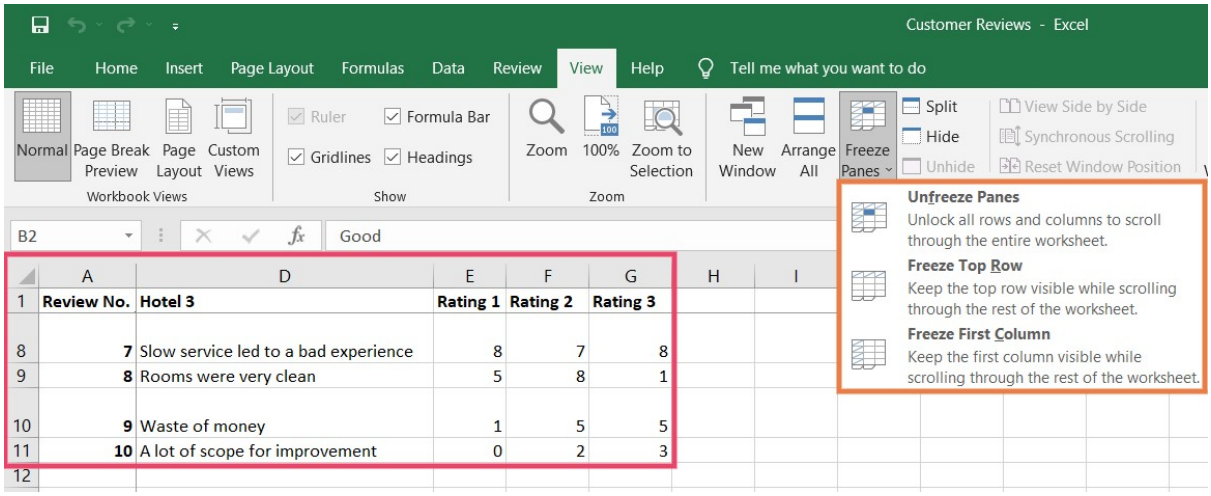
- In the picture above, we can observe **'Row1'** being **repeated**, it is actually because the split action has been performed keeping **cell 'B2' as the base**. The split screen command splits the screen into individual components and the components can be observed simultaneously.

You may observe a grey line where the screen split has been performed, to remove the split you can either re-click the icon to

remove it completely, or double click the grey line to remove the split one at a time (row-wise or column-wise) *

- **Switch Windows** – You can move to any other workbook opened in your device using this option. Basically, **ease of Navigation** within Workbooks is ensured by Switch Windows.

FREEZE PANES

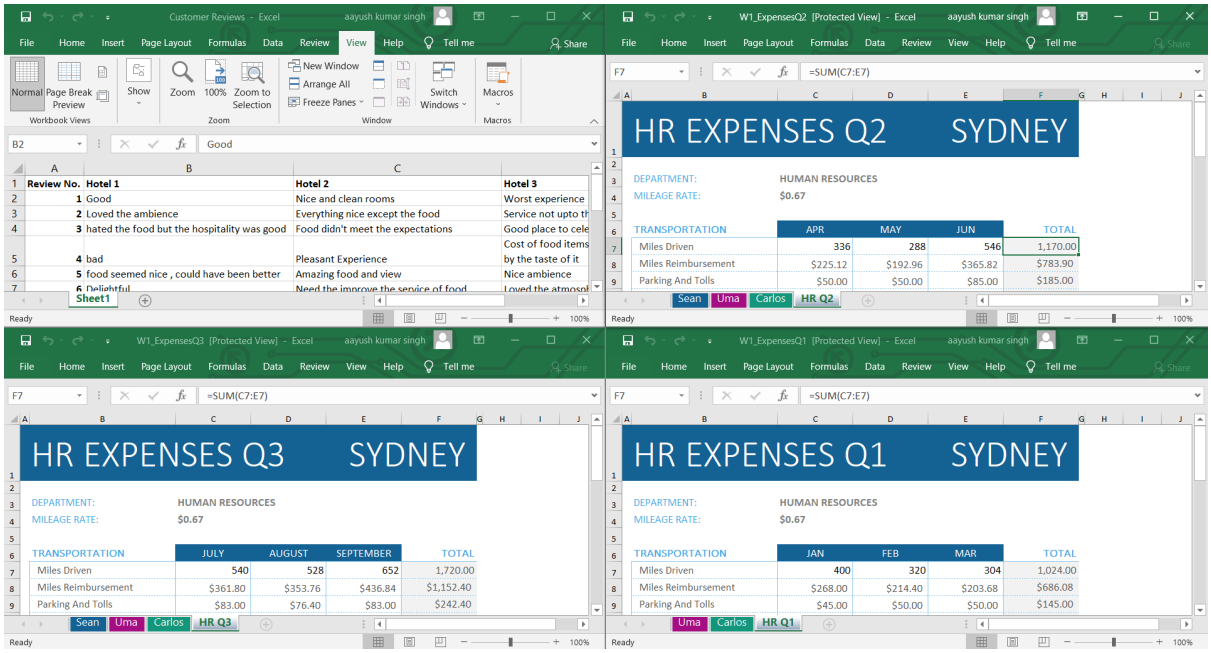


This option in the view tab enables us to Freeze our required Rows or Columns, it is mostly to used to freeze the headings of our data. As in to above picture you can observe that the first row and column are freeze and we can scroll and observe our data without losing the heading. One can easily click on the **Unfreeze icon** to remove freeze panes that can be observed as **thin grey lines**.

It is different from split as the freeze panes can't repeat itself in other panes, i.e., row 1 and column A can't be observed in the below portion of the sheet unlike in split screen*

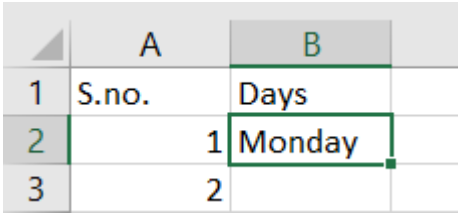
ARRANGE ALL

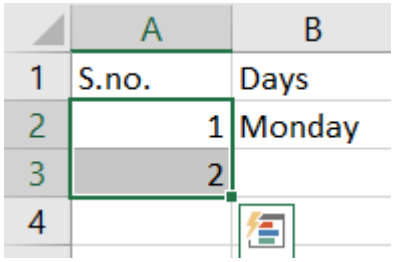
This function present in the view tab adjacent to the Freeze Panes icon is very useful when dealing with different workbooks. As shown below you can fit all your workbooks in a single frame and work simultaneously. You can import data from different workbooks, use **3-D formulae** (formulas including cells of different worksheets) and many more. Below picture shows the workbooks in **'tiled'** form, you may arrange the workbooks as per your fit.



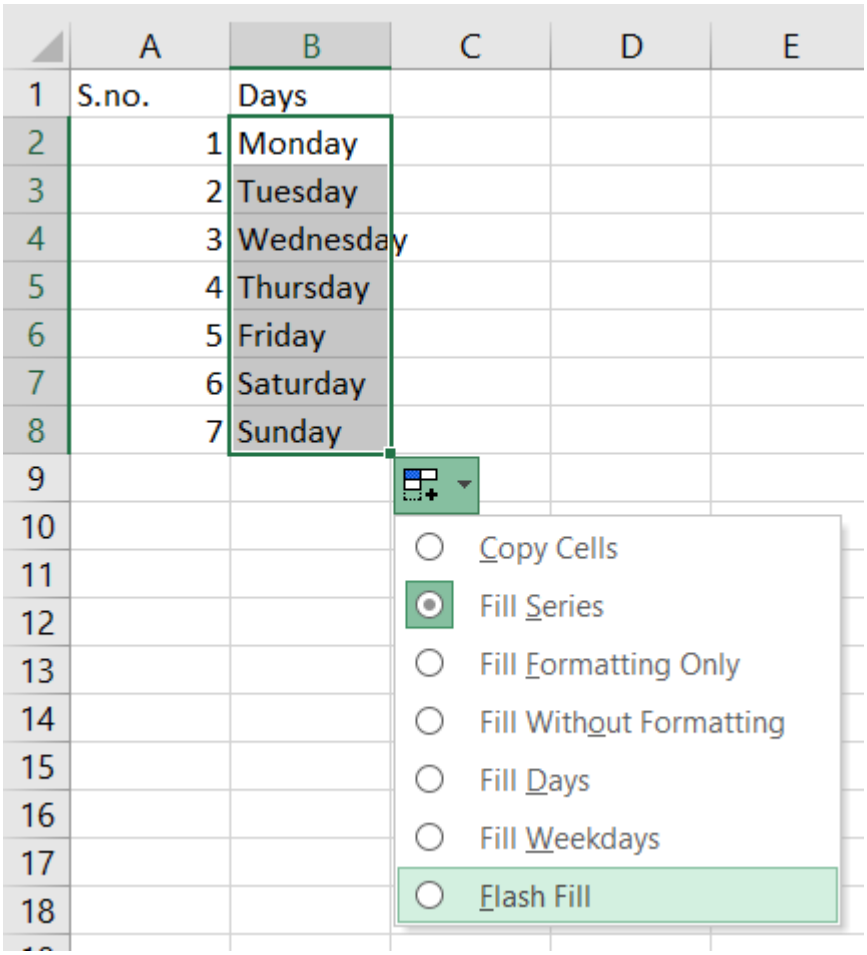
FILL HANDLE

- It is again one of the most efficient tools of MS Excel that is used for data entry and surely saves a lot of time.
- It is not present in any of the tabs but can be observed when the pointer is moved to the below right-hand corner of a cell. **(It is small dark '+' sign)**
- Let's see its use with the help of an example,





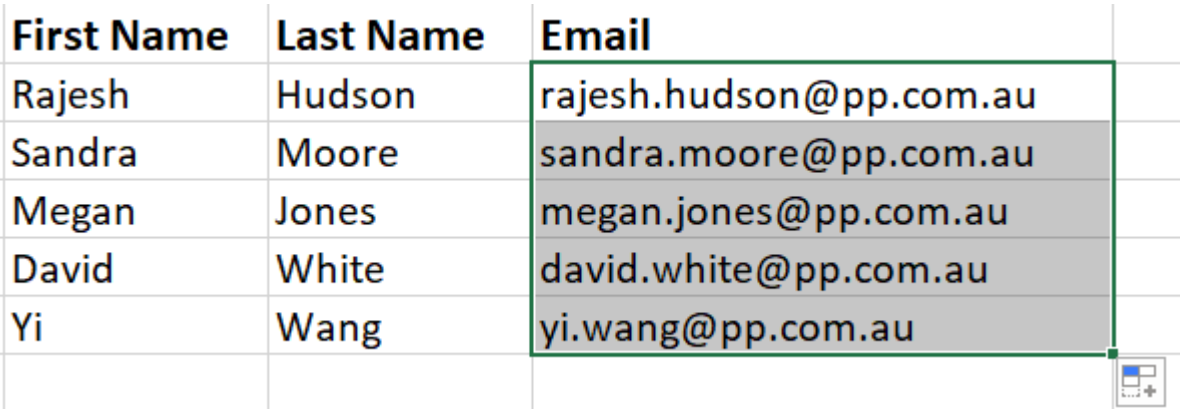
Here observe the two columns, Serial Number and Days, when applied fill handle, i.e., when the small green square at right hand corner is dragged along then we see the following results.



Observe that the days and serial number are filled in their respective order without actually entering them manually. This is how fill handle makes our work much easier. Observe the different options of fill handle as shown below and try implementing on your own.

Flash Fill is used to fill the cells in a particular manner, considering the pattern followed by a cell

Shown below is an example of flash fill,



CALCULATION

Calculations in Excel can be performed through **Formulas** or **Functions**,

1) **Formulas** – These are simple mathematical formulas that are entered in cells in order to perform calculations.
   Say you want to add up the values of cells **'A1'** and **'B1'** and display the result in **'C1'** then you can follow the following steps to do so,
   Step1) Click on cell **'C1'** and type **'='**
   Step2) Move your cursor to **'A1'** and press it.
   Step3) type **'+'**
   Step4) Move your cursor to **'B1'** and press it. The whole formula will look like **'=A1+B1'**
   Step5) Press **'Enter'** to see the sum/result.

   This however was a simple addition operation, you can perform various other calculations like subtraction, multiplication and division individually or altogether taking into consideration more than two cells

   Say you have several values in columns A1-A50 and B1-B50 and you have to give the sum of each pair of column A and B in respective cell of column C, here, rather than applying formula in each cell, one can simply apply formula in C1 and drag or double click the fill handle to calculate the values in subsequent cells of column C

   ****Fill handle applies to formulae as well****

2) Functions

What is the need of Functions when we have easy to implement formulas to take care of calculations?

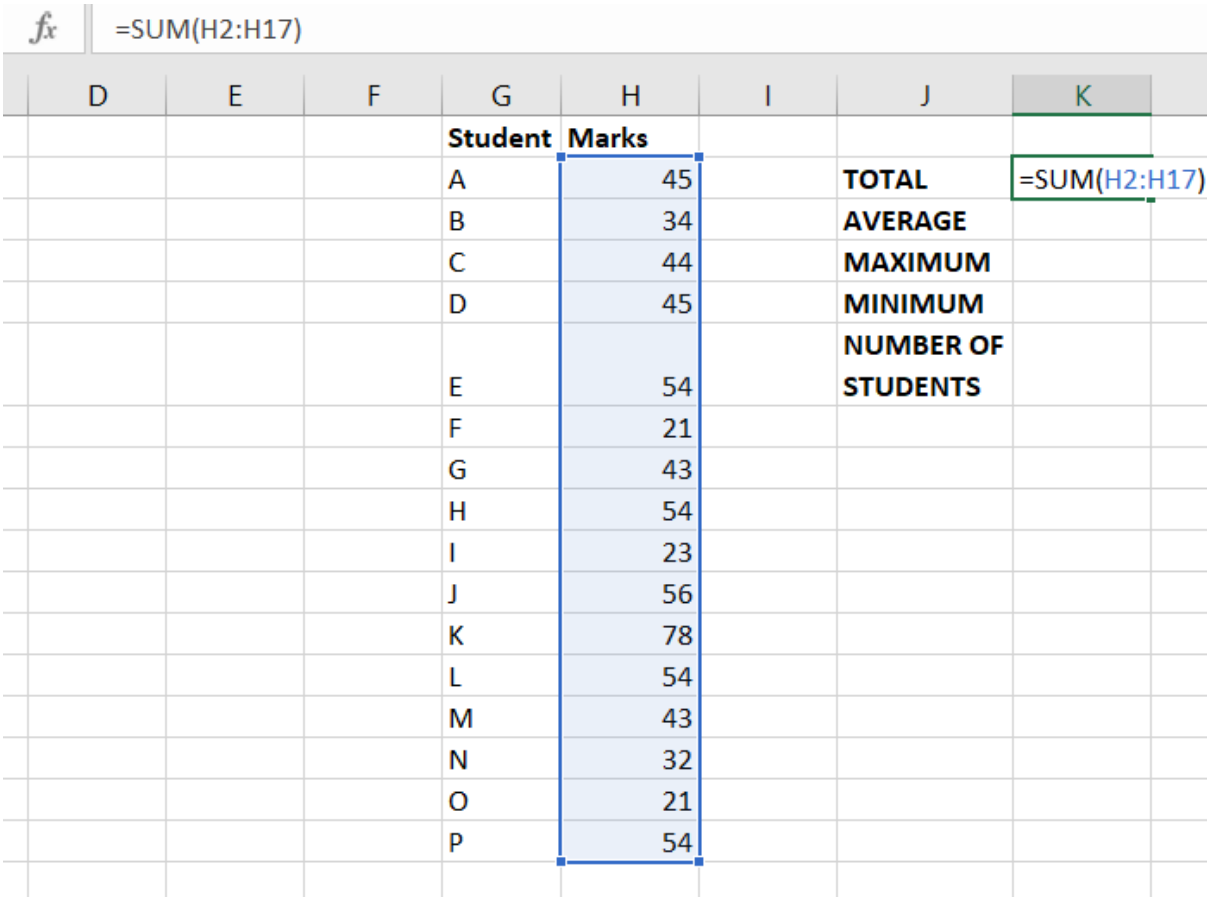
When dealing with large datasets Formulas might seem quite time consuming and this is where functions come into act and save the day for the users. Let's understand functions with the help of an example,
Here we have a list of students and their semester marks we are supposed to find the **'Total Marks', 'Average Marks', 'Maximum marks Obtained', 'Minimum Marks Obtained' and 'Total Number of Students'**.

- Using formula might be a bit time taking, as just in case of total marks, the formula used will be **=H2+H3+H4+…,** this will take a lot of time plus increase the probability of human error.

Here using functions is the ideal choice, observe the function written in column **K2** for adding up the marks of all the students **(=SUM (H2:H17))**, observe the formula bar as well!

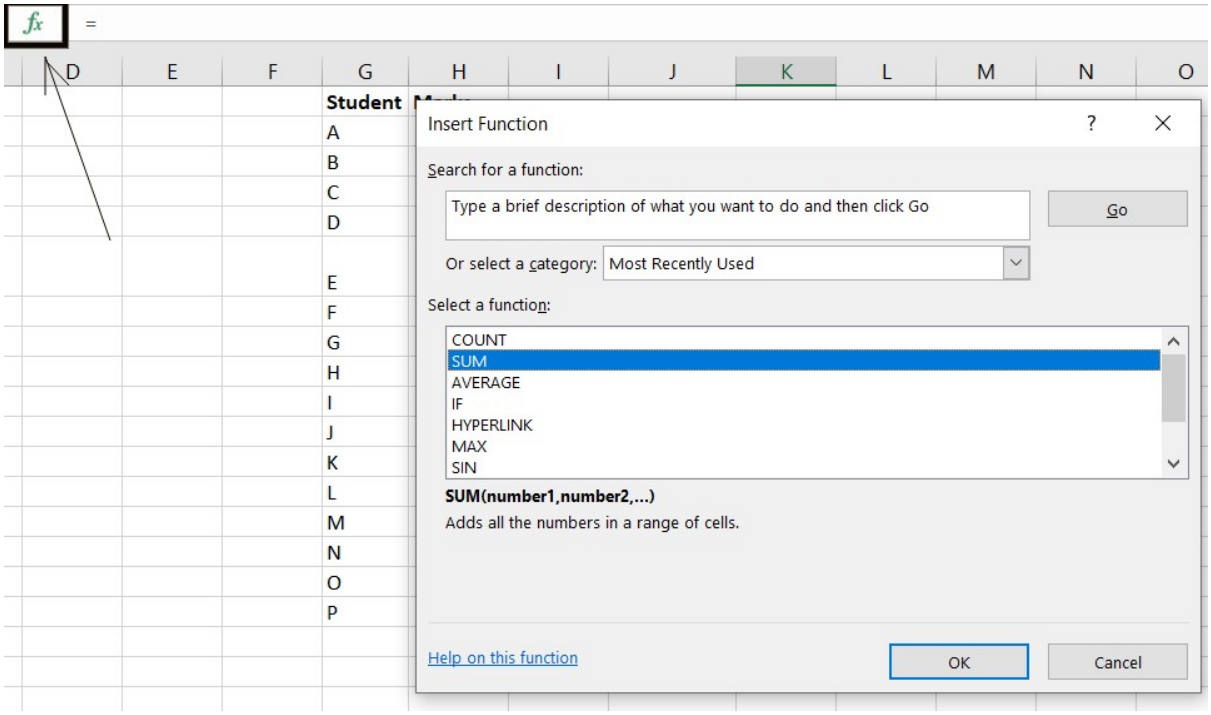
This is a simple and effective technique to perform calculations and we use the following functions in this example for our specific requirements,

- **SUM** – To get the total marks
- **AVERAGE** – To get the average marks
- **MAX** – To get the maximum marks
- **MIN** – To get the minimum marks
- **COUNT** – To get the number of students

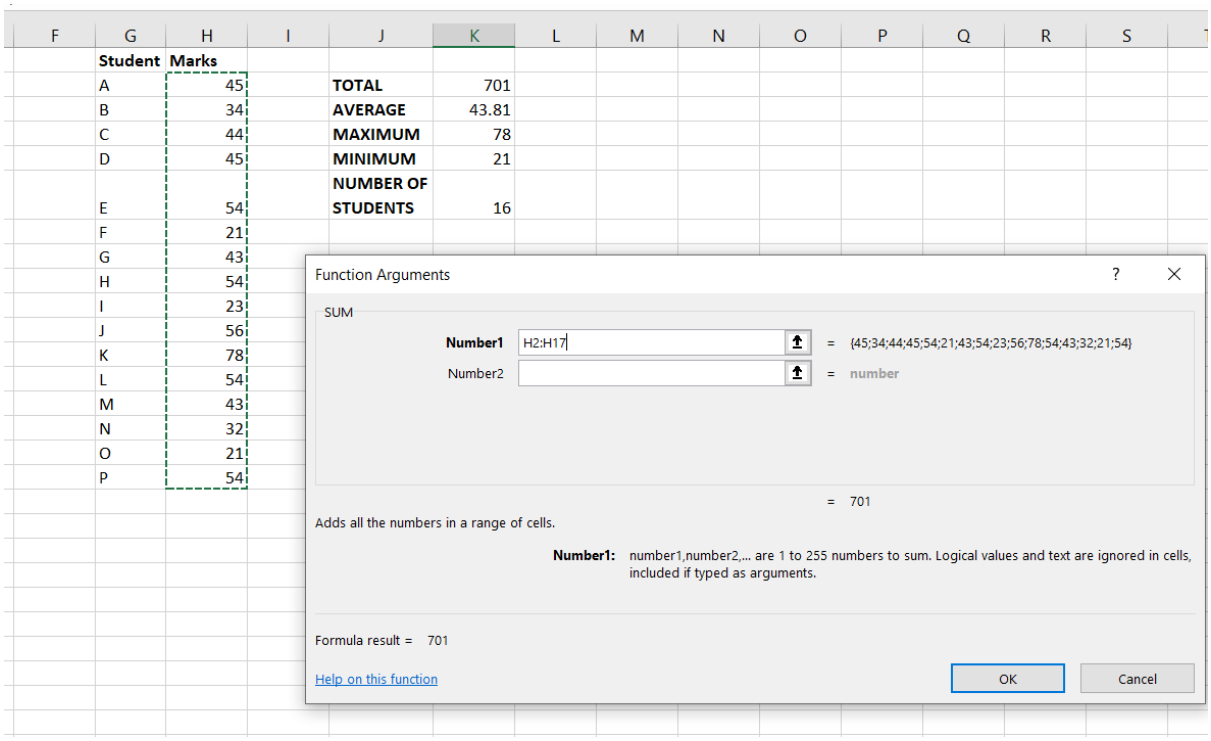


Another way we can implement functions is as follows,

Step 1) Click the highlighted **'f_x'** icon near the formula bar, the following dialog box appears on your screen, select the desired function and click **'OK'**.



Step2) Further select the range of number you want to apply the operation on and click **'OK'**.



This is a bit lengthy but useful method to implement functions efficiently.

**We can apply functions on the basis of conditions as well (SUMIF, COUNTIF) **

Function generates option while implementing, i.e., typing '=S' will show you ample of options like SUM, SEC, SECOND, SUMIF etc. along with its syntax, so this feature of excel makes it easier for the users to use functions

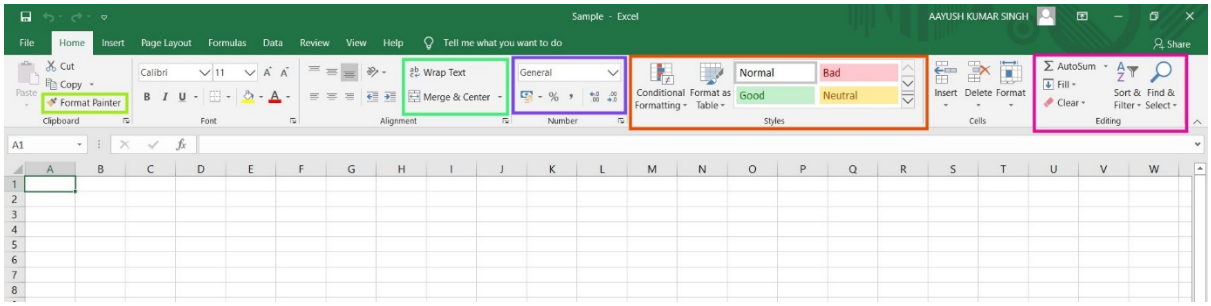
Numerical calculation formulae discussed above can be implemented within worksheets as well as across worksheets

Relative and Absolute Cell References

- **Relative Cell Reference –** By Default all the cell references are relative cell references, i.e., the **value changes with respect to different cells** (that is why we are able to use fill Handle; like in case of above example of list of student's marks)
- **Absolute Cell References –** We use absolute cell reference when we need to fix a cell's location when used in a formula, we can do so by adding a **dollar sign ($)** before the column name and row number **(A1)**. It is useful is various mathematical calculations regarding a dataset when a single value is involved multiple times, without repeating the value.

  $A1 indicates that the column A is fixed while row keeps on changing when applied elsewhere or on application of fill handle, while A$1 indicates that row 1 is fixed while the column keeps changing

HOME TAB

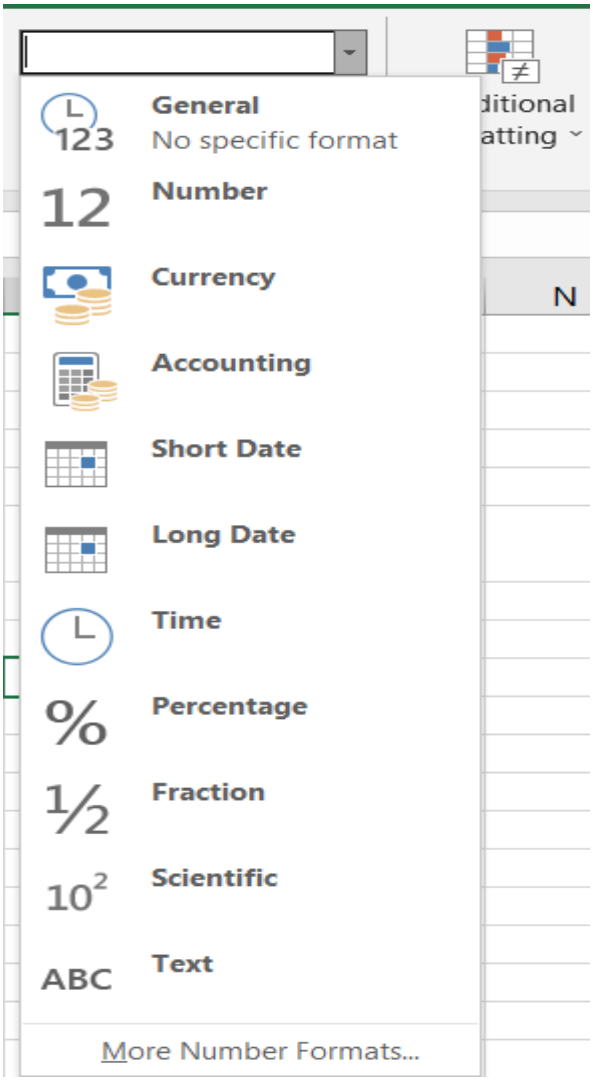


The **'Home'** tab contains several important features, icons and functions, to point out a few:

- **Format Painter** – It copies the formatting of one cell to another cell.
- **Warp Text** – It fits the contents or text in a single cell
- **Merge and Center** – Merges two or more cells and centers the content present in them. *Observe the drop down and try exploring other features of merge and center by yourself. You can also perform the same*

functions by selecting the cell/cells and right clicking the mouse, it will show the options and you can further select the required one.

- **Number** – It basically shows the different ways in which numbers can be interpreted in excel. One can select the option he/she wants to represents numerical values in. Let's look at the drop-down options of **'Number'** to look through the options provided by excel,



So, it is quite obvious that excel offers a number of ways in which the user can get the numerical values interpreted in. *The 'More Number Formats…' option present in the lowest position gives the liberty to implement numerical values in a better or detailed way. One should probably try out different number formats on their own and observe the presentation of the values.*

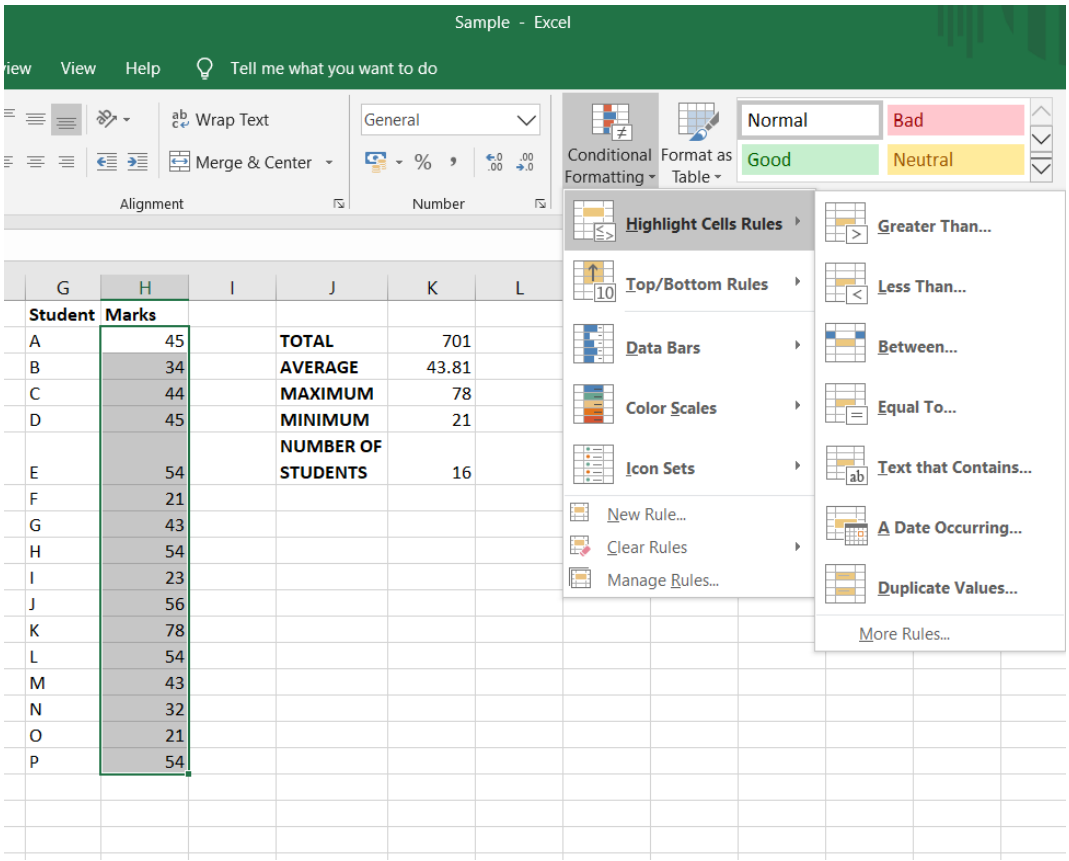
- **Styles** – Here one can color or make changes in a particular cell or a group of cells in order to make them unique or distinguishable from other cells in the worksheet. *One feature to notice in this section of*

***Home Tab** is **'Conditional Formatting'**, we will explore it in detail further in the article.*

- **Editing** – This section consists of a few important functions like **Sort and Filter** that are used to Sort the dataset in ascending or descending column or apply filters in the required rows or columns. We can also see the **find and replace icon** in this section which is used to find certain words in the worksheet and if required replace them with a certain word or phrase. *One of the important icons that can be seen in this section is the **AUTOSUM** icon, Autosum calculates the sum (or other values that you can observe in the dropdown of autosum) just by a single click, without the use of any formulas, it estimates the sum of all the numerical values above it or on the left side of the cell being applied to autosum when one or less numerical values are present above it.* ***There shouldn't be any empty cell in between the range of cells when applying Autosum****

CONDITIONAL FORMATTING

Conditional Formatting is basically used to **highlight cells** that follow a certain criterion. Let's see how we can implement it,



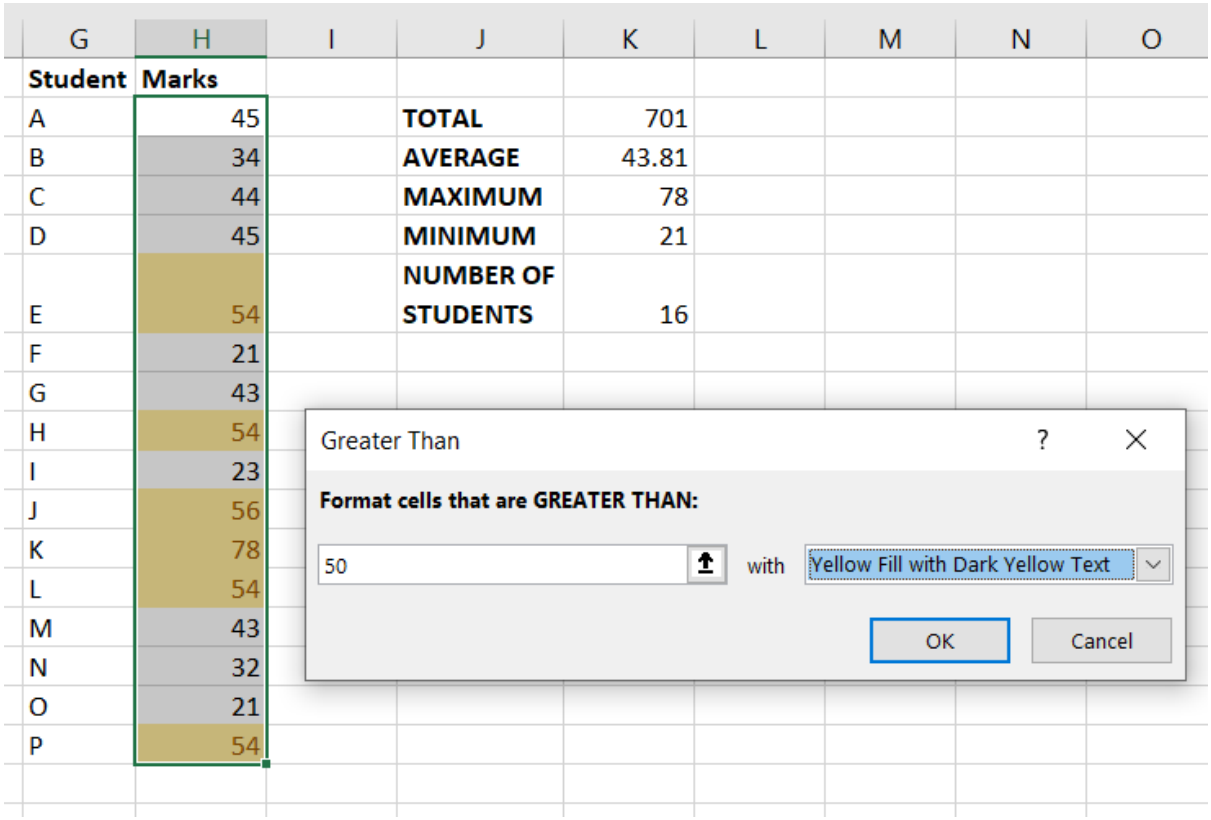
Here, say we want to highlight the marks of students that are greater than 50,

Step 1) Select the target cells

Step 2) Go to Conditional Formatting

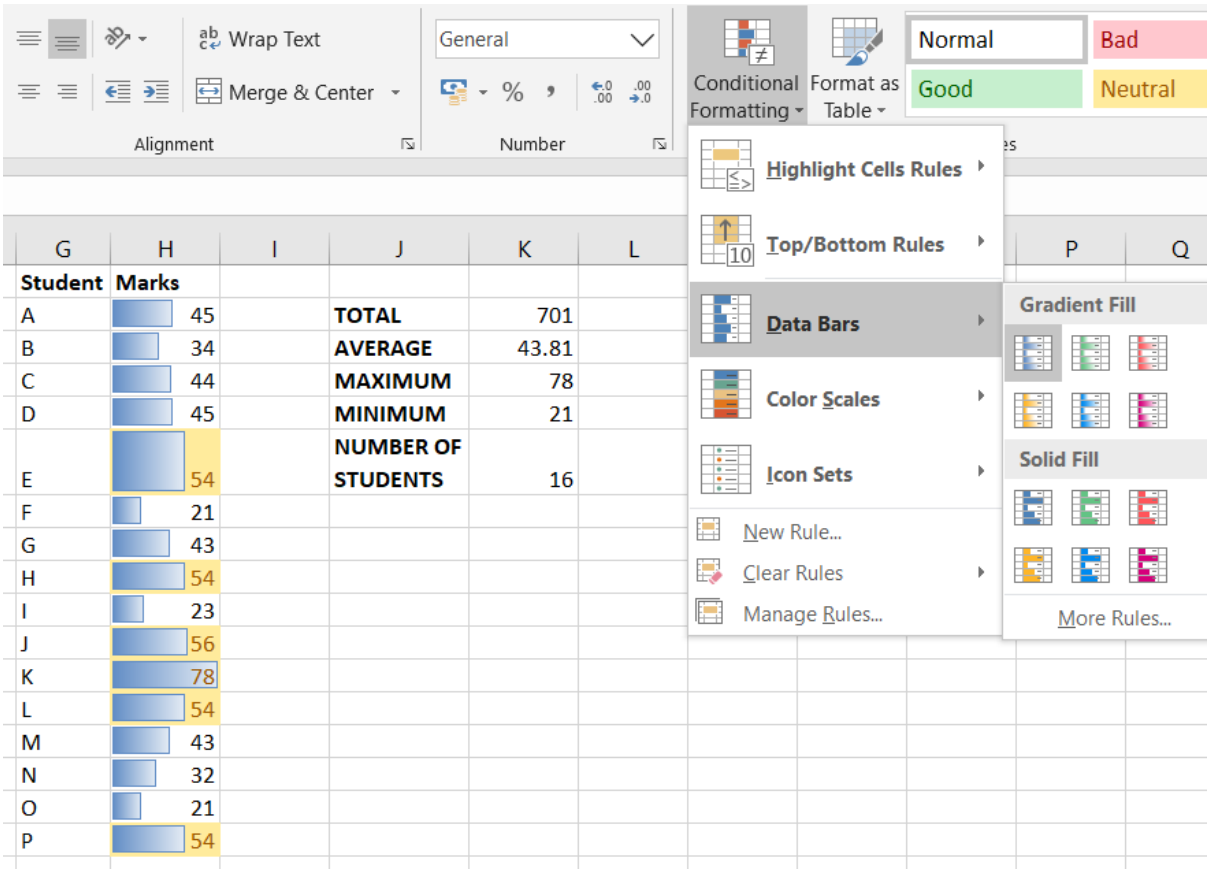
Step 3) In the highlight cell rules select **'Greater than'** option

Step 4) Enter the required values in the dialog box. You can change the highlighting color as per your needs from the side mentioned drop down in the dialog box.



Some attributes or options provided by the conditional formatting section are:

1) **Highlight Cell Rules** –Try to explore other options within this section just like we did the 'Greater than' option in the above example.
2) **Top/Bottom Rules** – It highlights the cells that are in the top certain percentage are bottom certain percentage of the selected data.
3) **Data Bars, Color scales and Icon sets** – It gives a relative representation of the values. Let's see what it means by implements **Data Bars** on our above-mentioned Students Marks dataset,

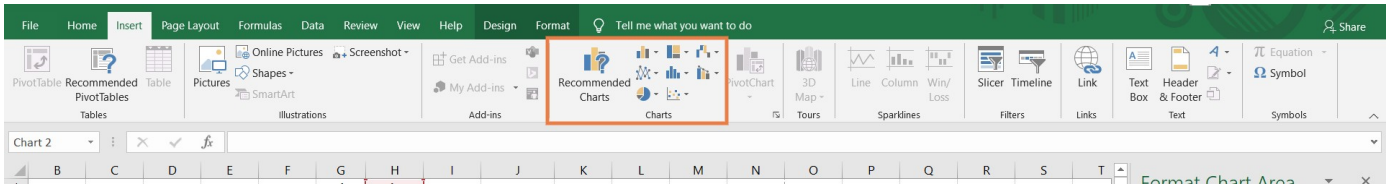


Observe that bars are created within the cells giving a **visual representation of the magnitude of the values**. *Similarly, Color scales and Icon sets are visual representation techniques.*

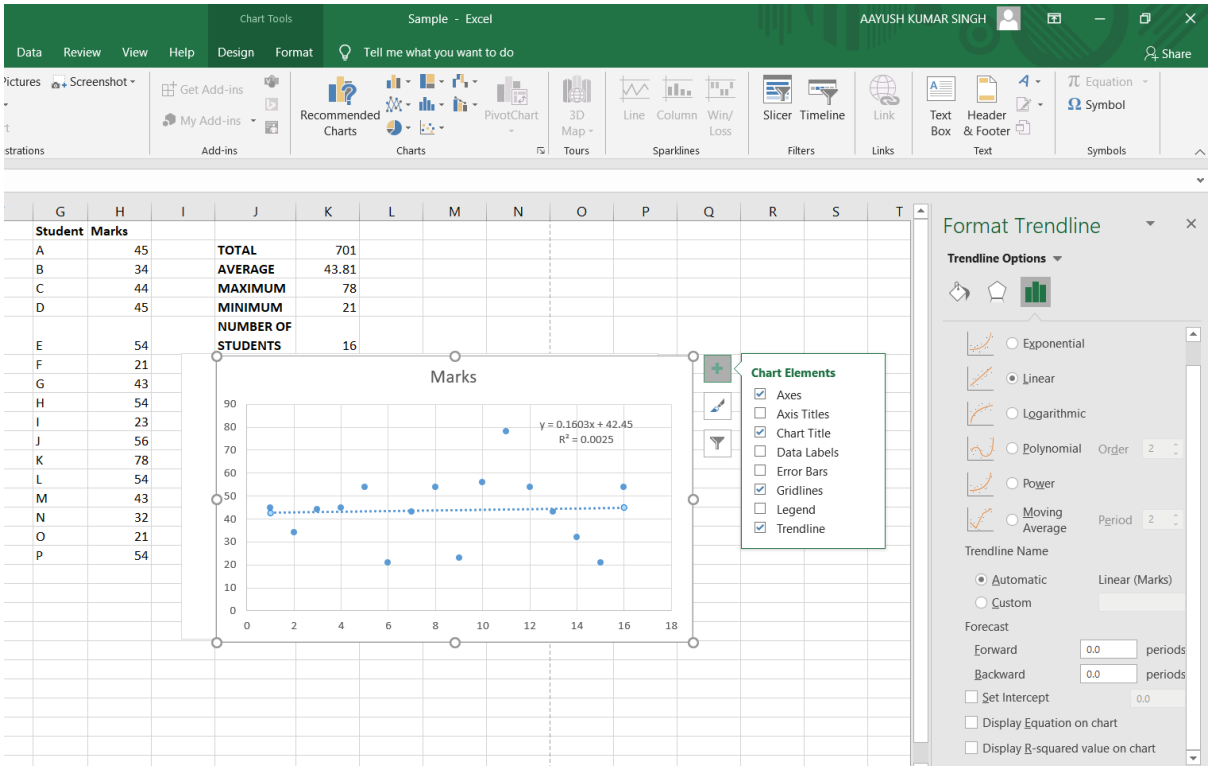
4) **Clear Rules** – It is used to clear conditional formatting from the entire worksheet or the selected cells.
5) **New Rule …** - Conditional Formatting gives us the liberty to create rules by ourselves and then implement them in the required dataset, this can be done with the help of this feature, it provides a number of ways in which we can create a rule and also the way in which we wish to highlight the cell following the rule. Reader might want to go through this feature of excel and try creating a random rule on a random dataset.
6) **Manage Rules** – This feature consists the list of all the rules applied in the spreadsheet and the user can easily edit, delete or create a new rule from here.

DATA VISUALIZATION IN EXCEL

Excel provides us with a range of functions to visualize our data and make conclusions out of it. Let's see the various options offered for data visualization.



In the 'Insert' tab we can find options to create charts of various types like Pie Chart, Column Chart, Bar Chart, Scatter chart and many more. One can simply create a chart by selecting the rows and columns taking in consideration and selecting the desired chart type. Let's try out the scatter chart on our students marks list,



From the above picture we can observe the following points,

- **Chart Elements** – It provides us with tools like **Axes, Title, Labels, Legends etc.** giving us the liberty of make our chart the way we wish to.
- **Chart types** – We can customize our chart in terms of presentation with the help of this tool.
- **Filter** – There is an inbuilt filter feature attached with the graph.
  ***Trendline -** *Observe the last chart element 'Trendline', it observes the trend in the selected dataset from which the chart is made up of and sketches a best fit line to suit the data, we however can format the trendline from the 'Format Trendline' bar as shown in the right side of the picture. We can also figure out the equation of the line or cure and*

*the R^2 value. This is one of the major tools used for Data Analysis in Excel for Data Visualization**

Pivot Charts are also one of the most important data visualization tools present in excel which we will cover along with Pivot Tables further in this article

TEXT RELATED FUNCTIONS

We have already discussed a few numerical functions provided by excel for calculations. However, excel also provides us with a variety of a Text functions that help us to deal with Texts within the dataset and perform certain operations required. Let's discuss a few text related functions and the operations performed by them:

1) **CONCAT or CONCATENATE**– It is used to join the texts of different cells. Excel provides us with the syntax of all the functions while executing them but still the expression for concatenating is **'=CONCATENATE (A1," ",B1)'**. This expression shows the addition of contents of cell A1 and B1 with space in between. ***We can also use the '&' symbol or concatenating text instead of the CONCAT function***
2) Functions to deal with the Case of the Text
    - **UPPER –** Changes the selected text in upper case.
    - **LOWER –** Changes the selected text in lower text.
    - **PROPER –** First letter of each word is in the upper case while others are in lower case.
3) Functions related to extracting the text
    - **LEFT –** Extracts characters from the left side of a word or phrase.
    - **Right –** Extracts characters from the right side of a word or phrase.
    - **MID –** Extracts characters from in between the word or phrase.
    - **FIND –** Finds the required characters from a phrase.

DATE RELATED FUNCTIONS

We discussed function related to numerical values and categorical texts, however, excel makes it easy for us to deal with dates as well by including some very important date related functions. To point out a few functions, we have:

1) **NOW()** – It returns the present date as well as time.
2) **TODAY()** – It returns the date of the present day.

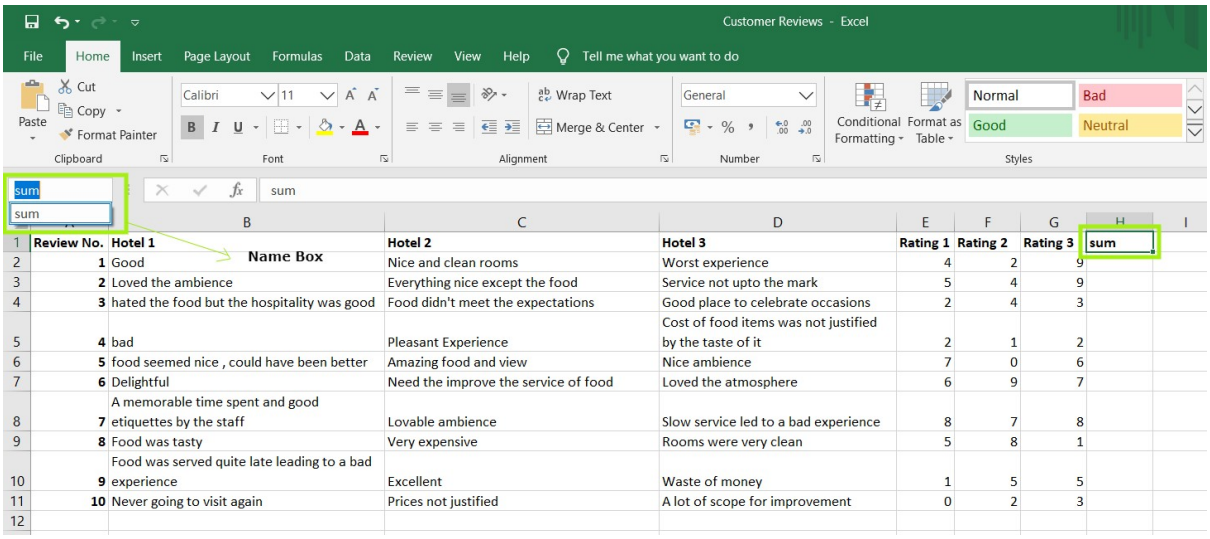
3) **YEARFRAC** – It returns the fraction of years when the time is represented in terms of days or months to make it a clearer presentation of the time elapsed.
Normal subtraction of days gives the days elapsed in the time period, here is where YEARFRAC really comes in handy to give a better time period presentation

NAMED RANGES

Named Ranges are names assigned to a particular cell or a range of cells. It basically helps to get work done faster and one can use named ranges in formulae as well making the formula look very convincing and understandable.

Methods to apply names ranges:

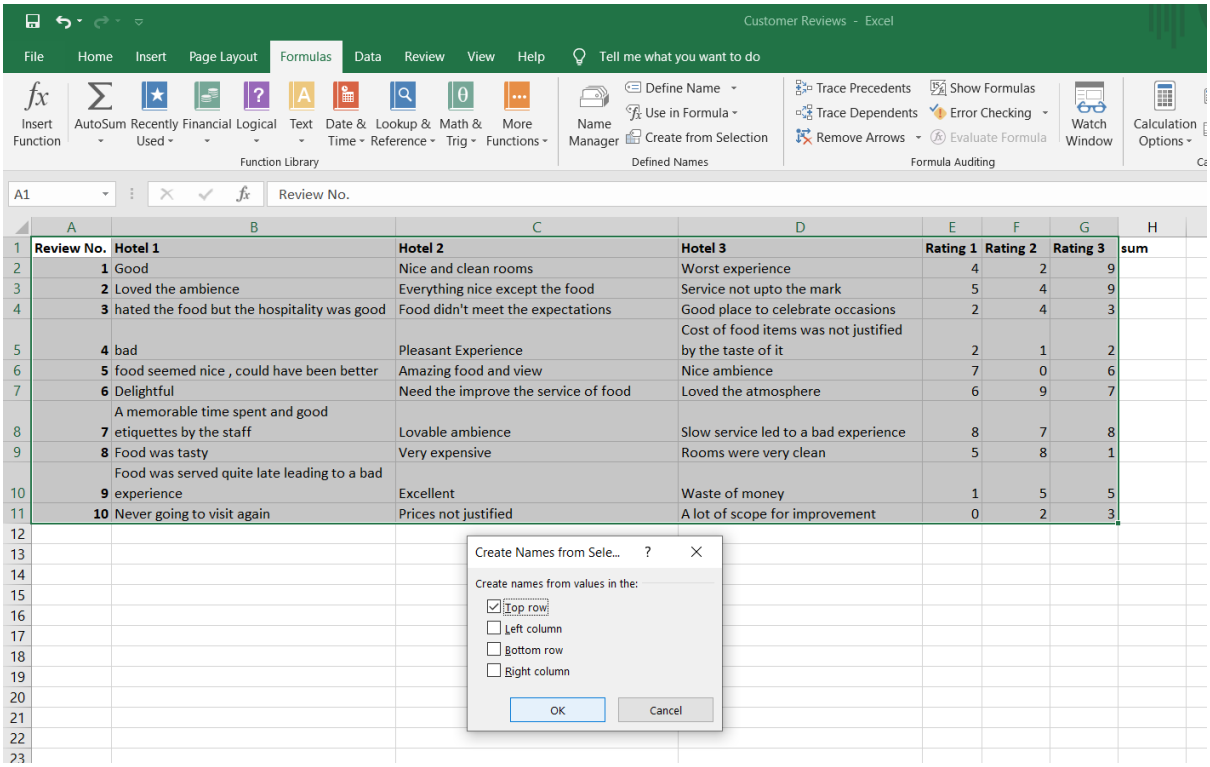
1) Directly select the cell or the ranges of cells and change the name in the **name box.**



	Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3	sum
2	1	Good	Nice and clean rooms	Worst experience	4	2	9	
3	2	Loved the ambience	Everything nice except the food	Service not upto the mark	5	4	9	
4	3	hated the food but the hospitality was good	Food didn't meet the expectations	Good place to celebrate occasions	2	4	3	
5	4	bad	Pleasant Experience	Cost of food items was not justified by the taste of it	2	1	2	
6	5	food seemed nice , could have been better	Amazing food and view	Nice ambience	7	0	6	
7	6	Delightful	Need the improve the service of food	Loved the atmosphere	6	9	7	
8	7	A memorable time spent and good etiquettes by the staff	Lovable ambience	Slow service led to a bad experience	8	7	8	
9	8	Food was tasty	Very expensive	Rooms were very clean	5	8	1	
10	9	Food was served quite late leading to a bad experience	Excellent	Waste of money	1	5	5	
11	10	Never going to visit again	Prices not justified	A lot of scope for improvement	0	2	3	
12								

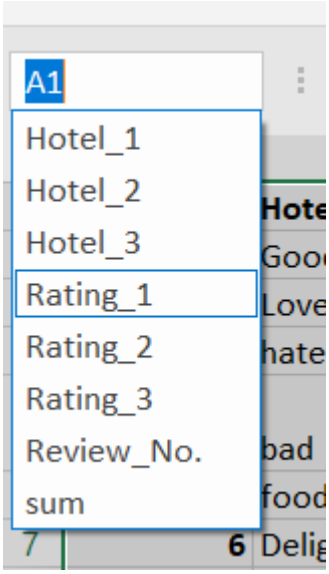
Here **H1 is named as 'sum'**

2) Select the cell or cells and click on the **'define names'** icon in the **formula tab** and name the cell or cells in the appearing dialog box.
3) **Create from Selection** – This is the most common and most useful way of generation of named ranges. It basically names the range of cells directly as the heading of the rows or column selected. Let's see this with the help of an example. Follow the following steps to generate named ranges from Create from Selection,
Step 1) Select the cells **(including the headings)** and click on **'Create from Selection'** in the **formulas tab.**



Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3	sum
1	Good	Nice and clean rooms	Worst experience	4	2	9	
2	Loved the ambience	Everything nice except the food	Service not upto the mark	5	4	9	
3	hated the food but the hospitality was good	Food didn't meet the expectations	Good place to celebrate occasions	2	4	3	
4	bad	Pleasant Experience	Cost of food items was not justified by the taste of it	2	1	2	
5	food seemed nice , could have been better	Amazing food and view	Nice ambience	7	0	6	
6	Delightful	Need the improve the service of food	Loved the atmosphere	6	9	7	
7	A memorable time spent and good etiquettes by the staff	Lovable ambience	Slow service led to a bad experience	8	7	8	
8	Food was tasty	Very expensive	Rooms were very clean	5	8	1	
9	Food was served quite late leading to a bad experience	Excellent	Waste of money	1	5	5	
10	Never going to visit again	Prices not justified	A lot of scope for improvement	0	2	3	

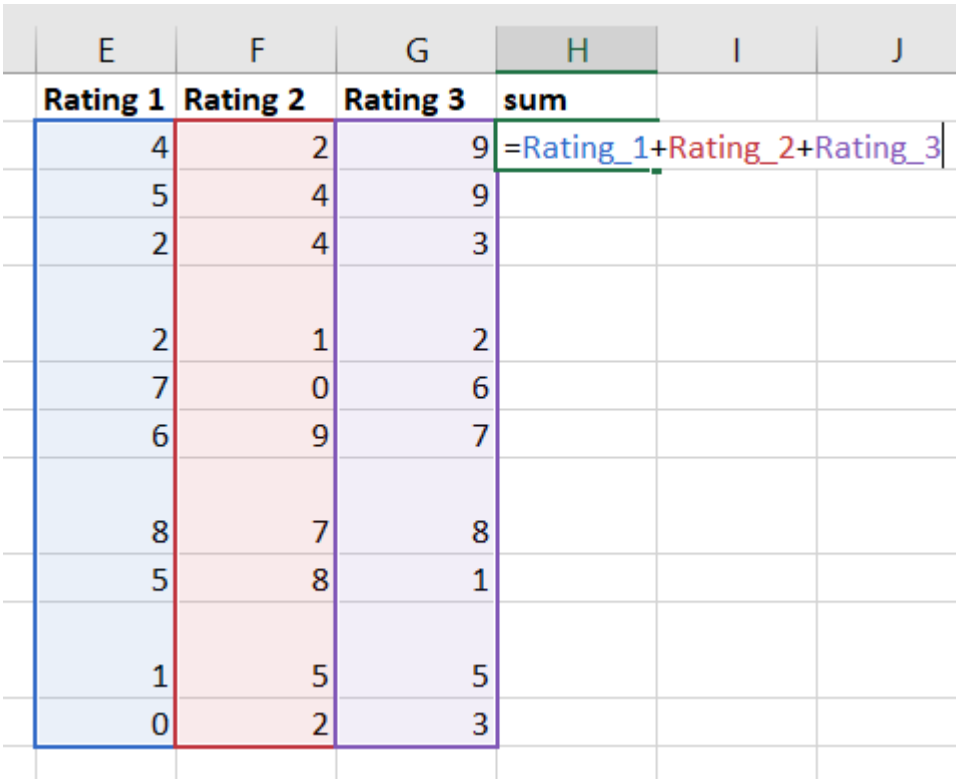
Step 2) Select the name you want to assign with respect to the dataset. Like here, naming the cells according to the top row heading will make most sense. Press **'OK'** and you have your named ranges created.



Observe the contents in the name box and you may want the select a particular named range from here and verify the data underneath it.

Calculation using named ranges

Named ranges make calculations really effective, easy to understand and implement. A sample calculation using named ranges is shown below,



E	F	G	H	I	J
Rating 1	**Rating 2**	**Rating 3**	**sum**		
4	2	9	=Rating_1+Rating_2+Rating_3		
5	4	9			
2	4	3			
2	1	2			
7	0	6			
6	9	7			
8	7	8			
5	8	1			
1	5	5			
0	2	3			

So, the above picture makes it pretty obvious why named ranges are so suitable to perform calculations.

Name Manager

The Name Manager is the user's one stop destination for all the stuff related to named ranges. The user can create new named ranges, delete, edit and apply filters to already created named ranges.

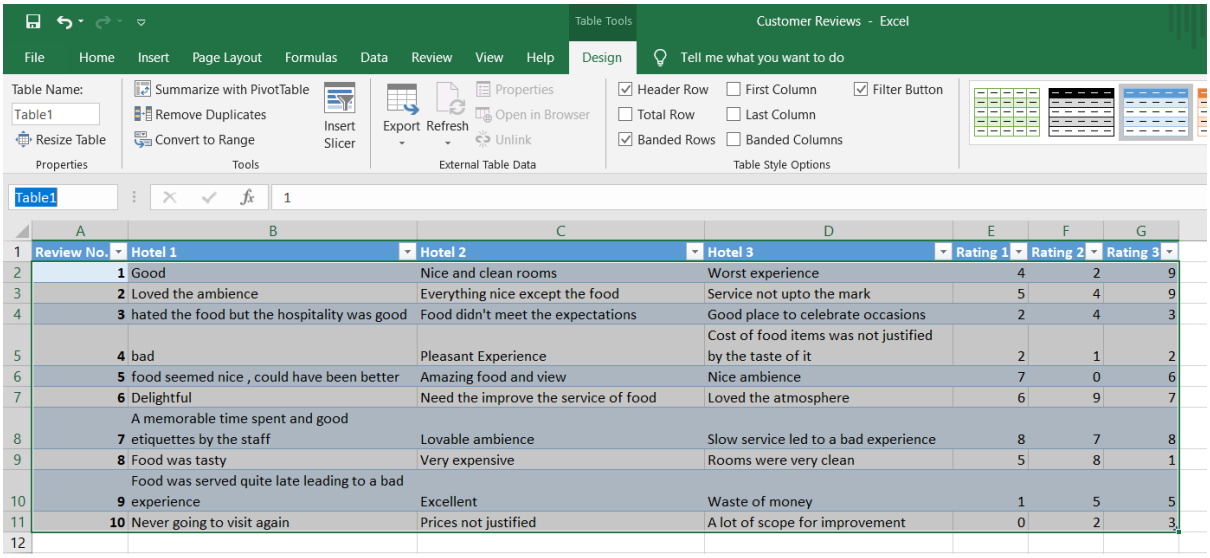
Use in Formula

This feature gives you the list of named ranges and the cells under those named ranges. *However, the most important feature in the **'use in formula'** drop down is **'paste names'** with its help one can copy and paste all the named ranges at a particular place or in a different worksheet, to get a list of all the named ranges along with their selected range. This feature really comes in handy when dealing with large datasets.*

TABLES

Tables are a very important and reliable feature in excel. Mostly analysts and researchers try to deal with data in excel in table format only due to the following features.

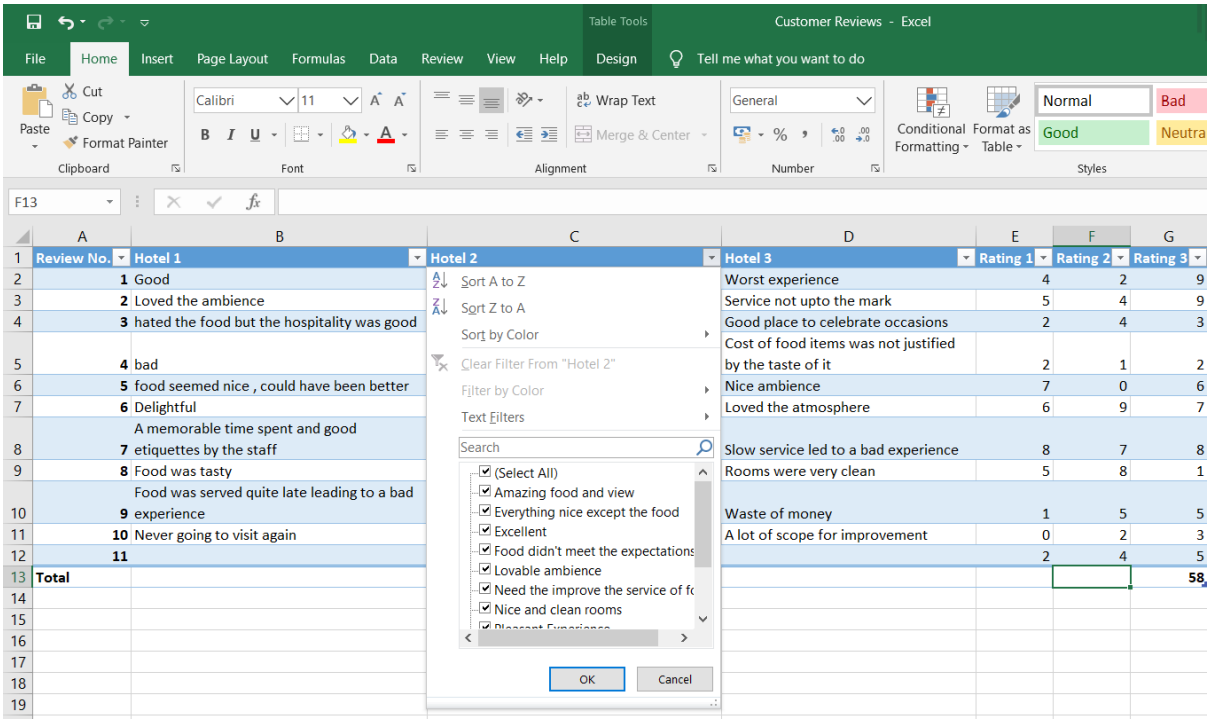
1) Easy to apply and name the whole dataset as a single unit.



Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3
1	Good	Nice and clean rooms	Worst experience	4	2	9
2	Loved the ambience	Everything nice except the food	Service not upto the mark	5	4	9
3	hated the food but the hospitality was good	Food didn't meet the expectations	Good place to celebrate occasions	2	4	3
4	bad	Pleasant Experience	Cost of food items was not justified by the taste of it	2	1	2
5	food seemed nice , could have been better	Amazing food and view	Nice ambience	7	0	6
6	Delightful	Need the improve the service of food	Loved the atmosphere	6	9	7
7	A memorable time spent and good etiquettes by the staff	Lovable ambience	Slow service led to a bad experience	8	7	8
8	Food was tasty	Very expensive	Rooms were very clean	5	8	1
9	Food was served quite late leading to a bad experience	Excellent	Waste of money	1	5	5
10	Never going to visit again	Prices not justified	A lot of scope for improvement	0	2	3

- To apply table one can easily select the required dataset and click on **table icon** in the **insert tab.**
- Notice the **'Name Box'** the whole data set is represented as **'Table1'** and the name can be varied if required.
- A new tab **Table Tools – Design** appears in the ribbon that contains functionalities related to the table.

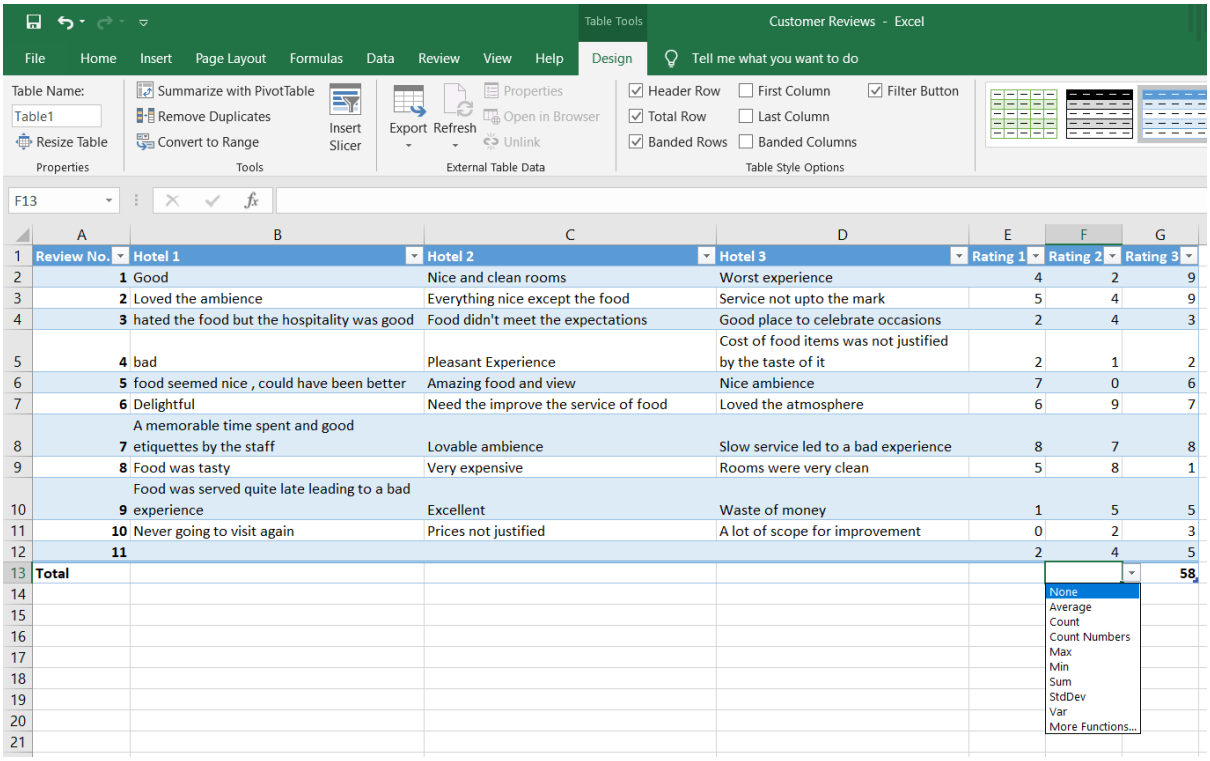
2) Auto filter and sort option



Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3
1	Good		Worst experience	4	2	9
2	Loved the ambience		Service not upto the mark	5	4	9
3	hated the food but the hospitality was good		Good place to celebrate occasions	2	4	3
4	bad		Cost of food items was not justified by the taste of it	2	1	2
5	food seemed nice , could have been better		Nice ambience	7	0	6
6	Delightful		Loved the atmosphere	6	9	7
7	A memorable time spent and good etiquettes by the staff		Slow service led to a bad experience	8	7	8
8	Food was tasty		Rooms were very clean	5	8	1
9	Food was served quite late leading to a bad experience		Waste of money	1	5	5
10	Never going to visit again		A lot of scope for improvement	0	2	3
11				2	4	5
Total						58

Notice the drop down in the right-hand side of each cell of the heading row. That is actually the filter and sort option, so that the user can filter and sort the data instantly as per his/her needs.

3) Automation and Total Row feature of Tables

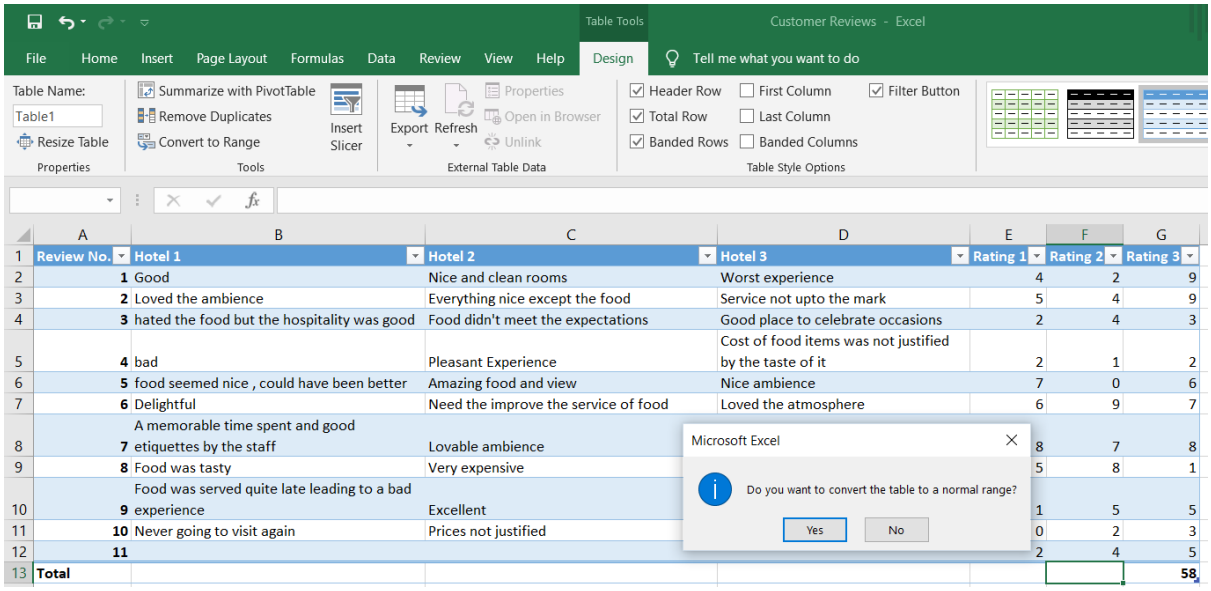


Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3
1	Good	Nice and clean rooms	Worst experience	4	2	9
2	Loved the ambience	Everything nice except the food	Service not upto the mark	5	4	9
3	hated the food but the hospitality was good	Food didn't meet the expectations	Good place to celebrate occasions	2	4	3
4	bad	Pleasant Experience	Cost of food items was not justified by the taste of it	2	1	2
5	food seemed nice , could have been better	Amazing food and view	Nice ambience	7	0	6
6	Delightful	Need the improve the service of food	Loved the atmosphere	6	9	7
7	A memorable time spent and good etiquettes by the staff	Lovable ambience	Slow service led to a bad experience	8	7	8
8	Food was tasty	Very expensive	Rooms were very clean	5	8	1
9	Food was served quite late leading to a bad experience	Excellent	Waste of money	1	5	5
10	Never going to visit again	Prices not justified	A lot of scope for improvement	0	2	3
11				2	4	5
Total						58

- Observe that a new column number 11 has been added to the dataset, and it has been automatically made a part of the Table. *This feature of table to automatically add the new rows and columns in the table is called* ***'Automation'***
- Also observe that the **'Total Row'** check box has been clicked in the **Table Tools – Design tab**, and this is because we can see a newly created row number 13 with displaying the total sum of Rating 3 column. Also, one can observe the drop-down feature besides every cell in row number 13 giving various options like **Sum, Average, Count** etc. This is a really efficient tool provided by tables in excel to make our working really fast and efficient.

Convert to Range and Subtotal Feature

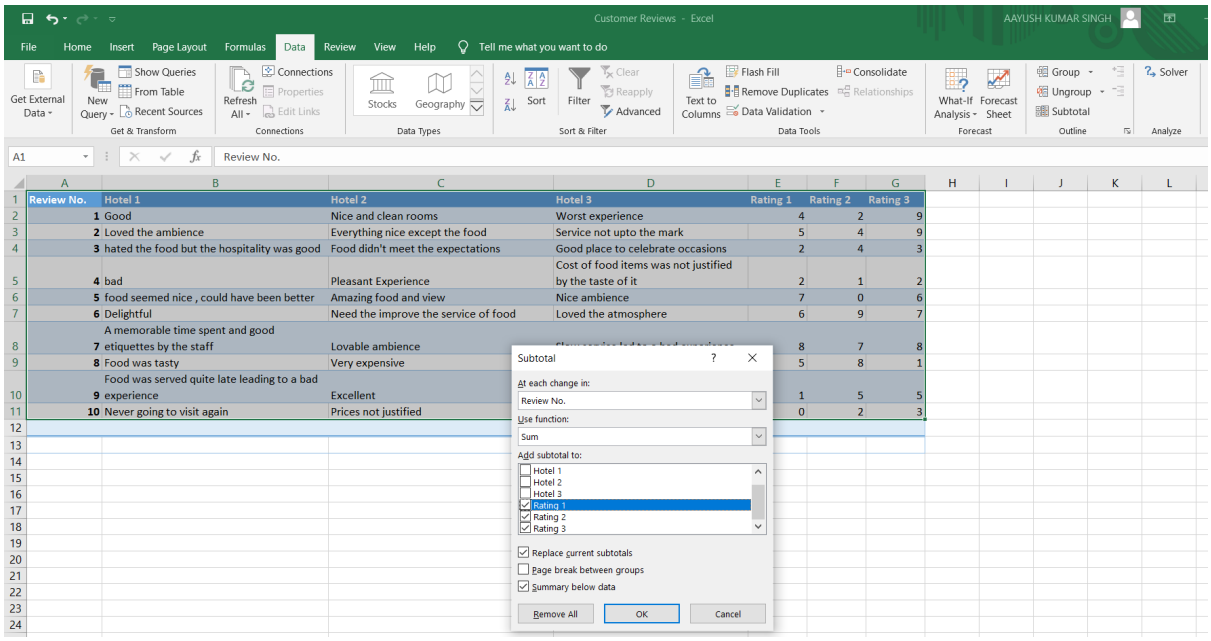
Subtotal feature helps us to get the total of the subsets within a dataset. However, this can't be obtained when the data is in form of Table. So, first we need to convert the table in range and then apply the subtotal feature. Let's see how to proceed,



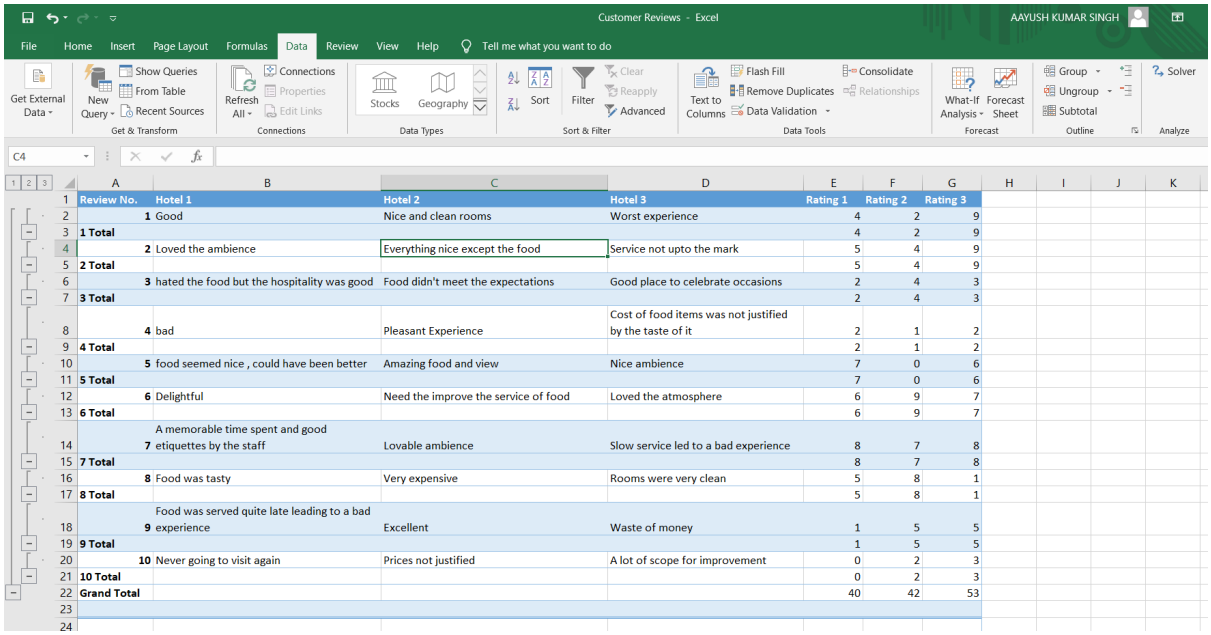
Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3
1	Good	Nice and clean rooms	Worst experience	4	2	9
2	Loved the ambience	Everything nice except the food	Service not upto the mark	5	4	9
3	hated the food but the hospitality was good	Food didn't meet the expectations	Good place to celebrate occasions	2	4	3
4	bad	Pleasant Experience	Cost of food items was not justified by the taste of it	2	1	2
5	food seemed nice , could have been better	Amazing food and view	Nice ambience	7	0	6
6	Delightful	Need the improve the service of food	Loved the atmosphere	6	9	7
7	A memorable time spent and good etiquettes by the staff	Lovable ambience		8	7	8
8	Food was tasty	Very expensive		5	8	1
9	Food was served quite late leading to a bad experience	Excellent		1	5	5
10	Never going to visit again	Prices not justified		0	2	3
11				2	4	5
Total						58

Notice the **'Convert to Range'** icon present in the **'Tools'** section of the **Table Tools – Design tab**. Click the icon and confirm your actions in the dialog box that follows.

However, notice that converting a table back to range won't alter or remove the rows, columns or elements added in the table



Now, notice the **'Subtotal'** feature in the **'Outline'** section of the **'Data' tab**. Click anywhere in the range and then on the subtotal icon. Select the options you want the subtotal of from the further appearing dialog box, click on the checkboxes according to your requirement and then click **'OK'**



Review No.	Hotel 1	Hotel 2	Hotel 3	Rating 1	Rating 2	Rating 3
1	Good	Nice and clean rooms	Worst experience	4	2	9
1 Total				4	2	9
2	Loved the ambience	Everything nice except the food	Service not upto the mark	5	4	9
2 Total				5	4	9
3	hated the food but the hospitality was good	Food didn't meet the expectations	Good place to celebrate occasions	2	4	3
3 Total				2	4	3
4	bad	Pleasant Experience	Cost of food items was not justified by the taste of it	2	1	2
4 Total				2	1	2
5	food seemed nice , could have been better	Amazing food and view	Nice ambience	7	0	6
5 Total				7	0	6
6	Delightful	Need the improve the service of food	Loved the atmosphere	6	9	7
6 Total				6	9	7
7	A memorable time spent and good etiquettes by the staff	Lovable ambience	Slow service led to a bad experience	8	7	8
7 Total				8	7	8
8	Food was tasty	Very expensive	Rooms were very clean	5	8	1
8 Total				5	8	1
9	Food was served quite late leading to a bad experience	Excellent	Waste of money	1	5	5
9 Total				1	5	5
10	Never going to visit again	Prices not justified	A lot of scope for improvement	0	2	3
10 Total				0	2	3
Grand Total				40	42	53

Now observe the output or the sum of the subsets. Here, each entry is considered as a different subset and results are shown accordingly. However, you can adjust your subsets from the left-hand side portion of the subtotaled worksheet.

A major advantage of dealing with Tables rather than ranges is the relation of Data Visualization in both cases. Due to the automation feature of Tables, the graphs change when a new row or column is added to the Table, which is not the scene in case of Ranges*

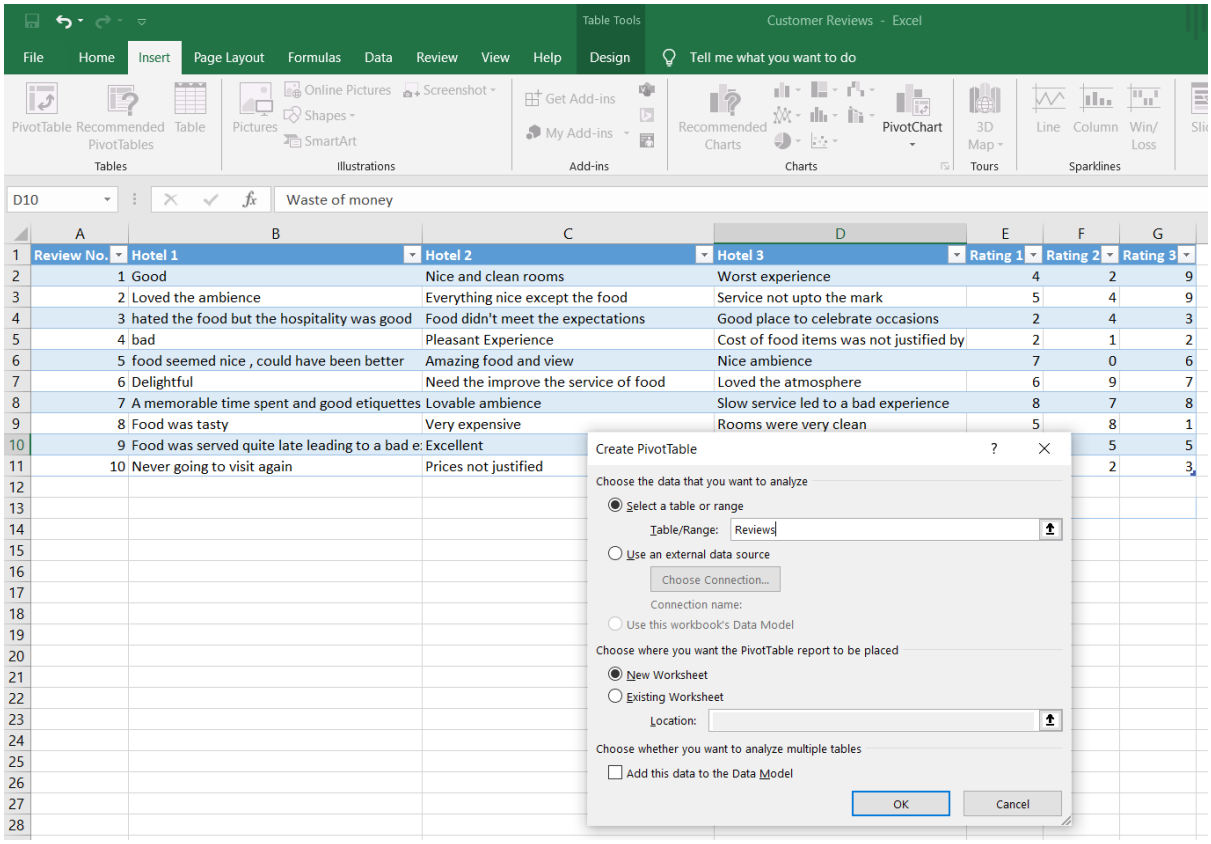
PIVOT TABLES

Pivot tables is a kind of **summarizing tool** in excel. Large datasets can be compressed in any desired form with the help of pivot tables. We can extract specific work related or relevant information from huge datasets with the help of Pivot Tables.

Let us see how to apply pivot table along with its various functionalities,

Follow the following steps in order to create Pivot table,

Step1) Select a cell from the range or table of which you want to create pivot table of and select the **'Pivot Table'** icon in the **'Insert' tab**. A dialog box will appear, select the range or table and the location you want to place your Pivot Table in and press **'OK'**

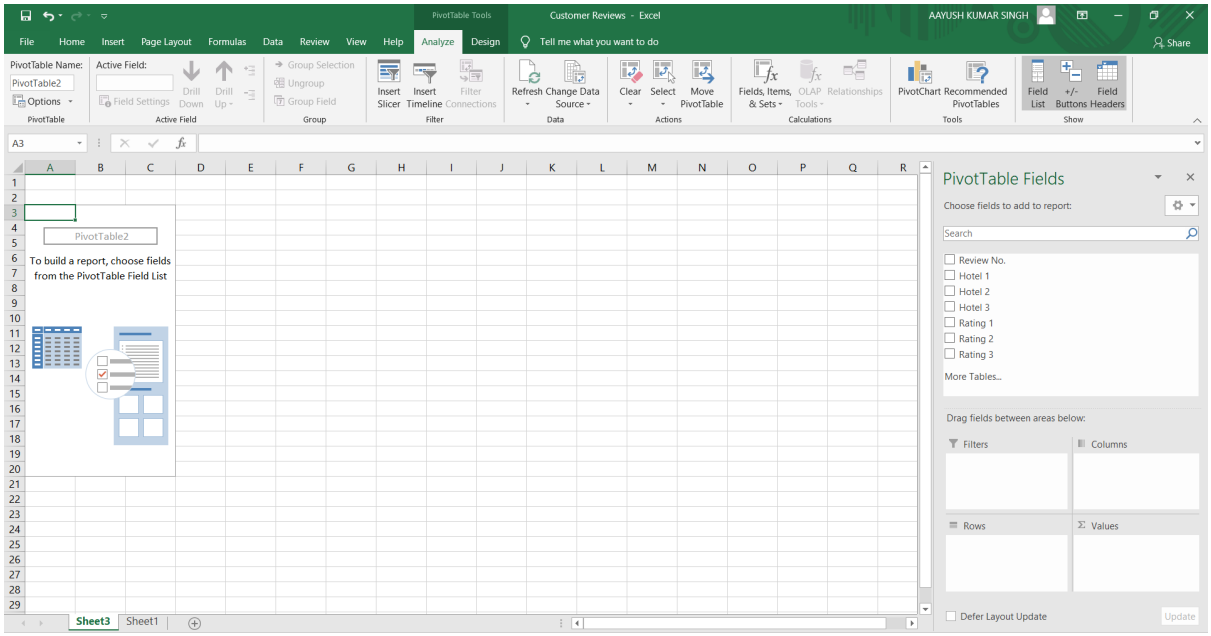


Here, I have selected the location of the pivot table as new worksheet.

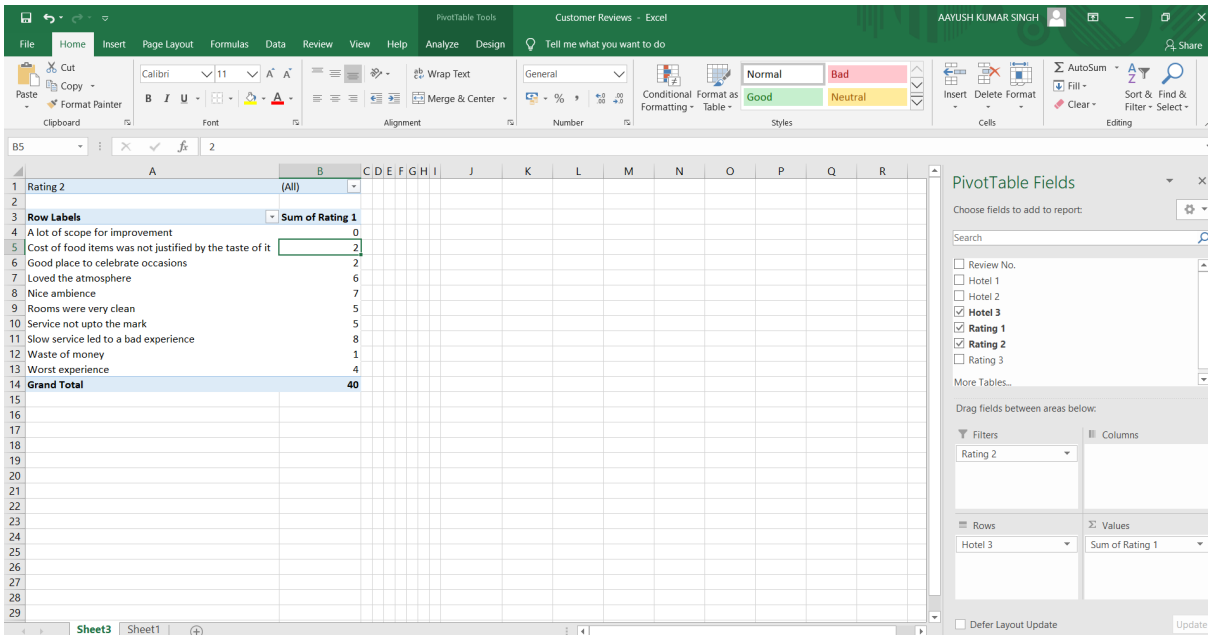
Step 2) Now we have a new sheet where we can observe the following

- Two new tabs **"Pivot Table Tools – 'Analysis' and 'Design'"** come into play; these tabs contain almost all the functionalities required by the user to deal with Pivot Tables. **Design tab** mostly contains **themes** and **visual presentation tools** while **Analysis tab** contains more **technical aspects of the pivot table**.
- A section **'Pivot Table Fields'** is introduced at the right-hand side of the screen. This is the portion where we will customize our pivot table according to our requirements. It contains four sections,
  1) **Filters** – Include variable that you might want to filter out of.
  2) **Rows**
  3) **Column** – It presents values under different conditions. *(Say you want to observe records for quarterly sales of a company, or yearly or per day; similarly, column deals with different conditions in different datasets)*

4) **Values** – It represents the numerical portion of the Pivot table and gives the overall sum, count or other required attributes of the data as selected.



Step 3) Drag and drop the variables in Filters, Column, Rows or Values section of the **'Pivot Table Fields'** as per requirement.



Here, you can observe the outcome when Rating 2, Hotel 3 and rating 1 are put in Filters, Rows and Values section respectively.

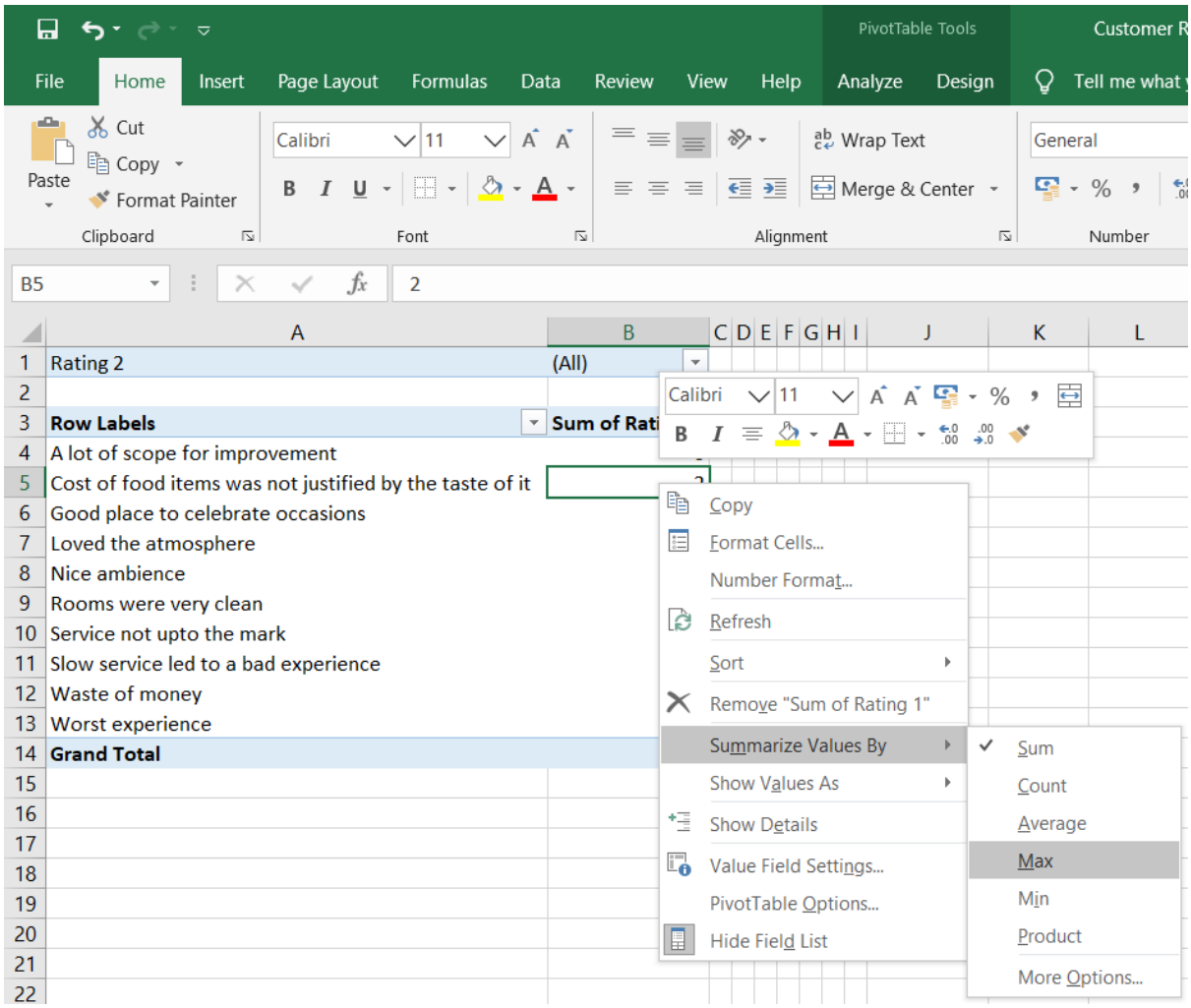
This might not make a lot of sense, but try to get the essence of the sections under Pivot Table Fields

So, this is how one can summarize their data with the help of Pivot Tables.

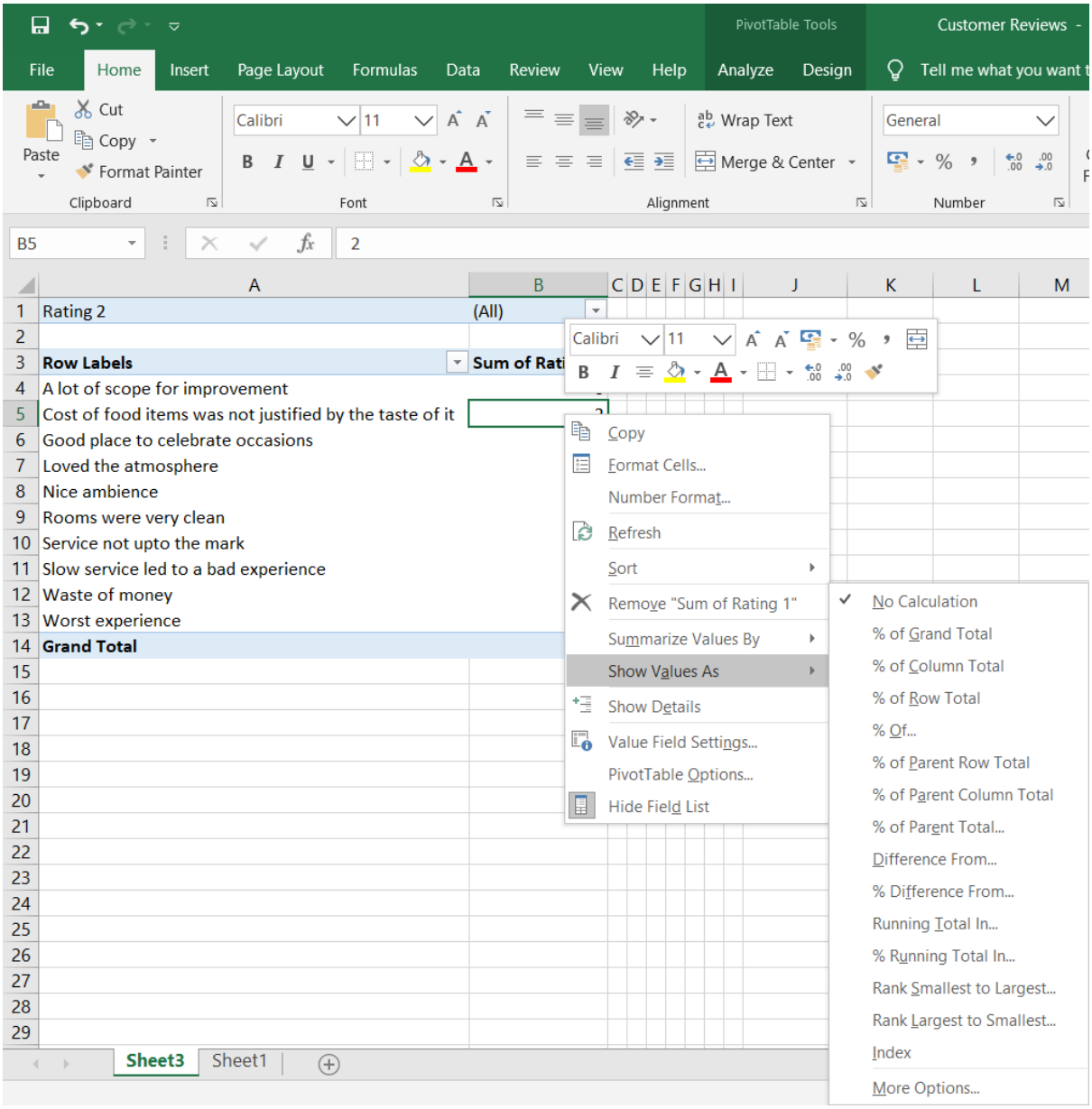
Various functionalities of Pivot Tables

Pivot Tables give us with a load of functions and features to get our data summarized in the best way possible. Let's look at few functionalities provided by the Pivot Table,

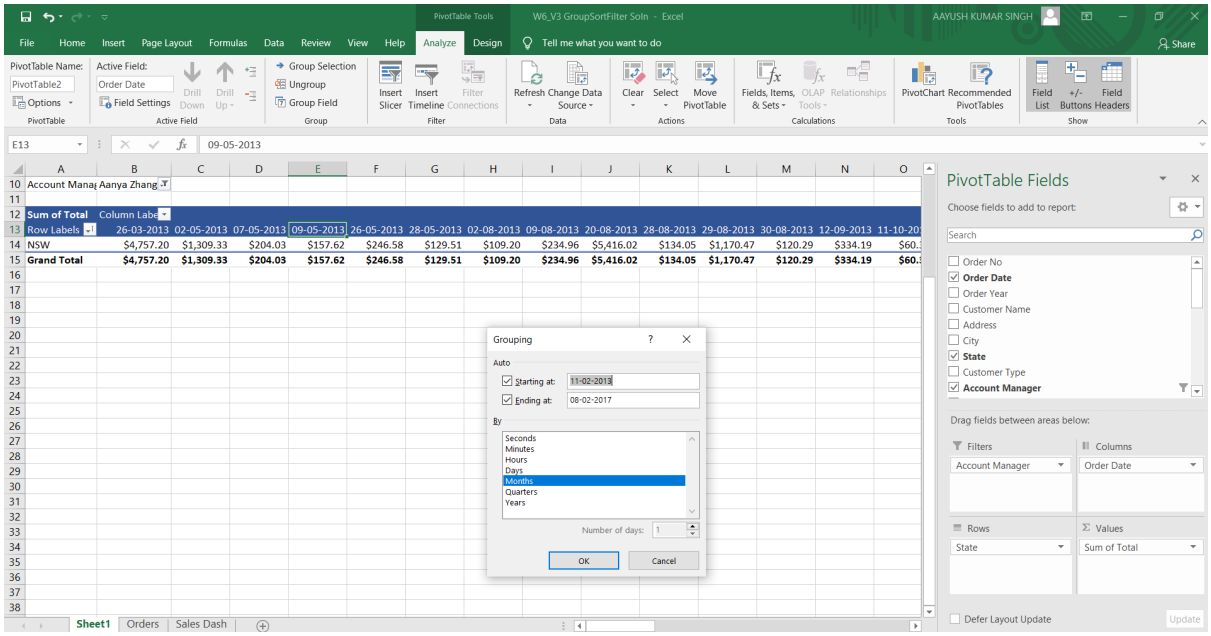
- **Summarize Values by** – It summarizes the contents of the variables under value section as per requirement. Just right click of any of the cells under value section of the pivot table and select the desired option.



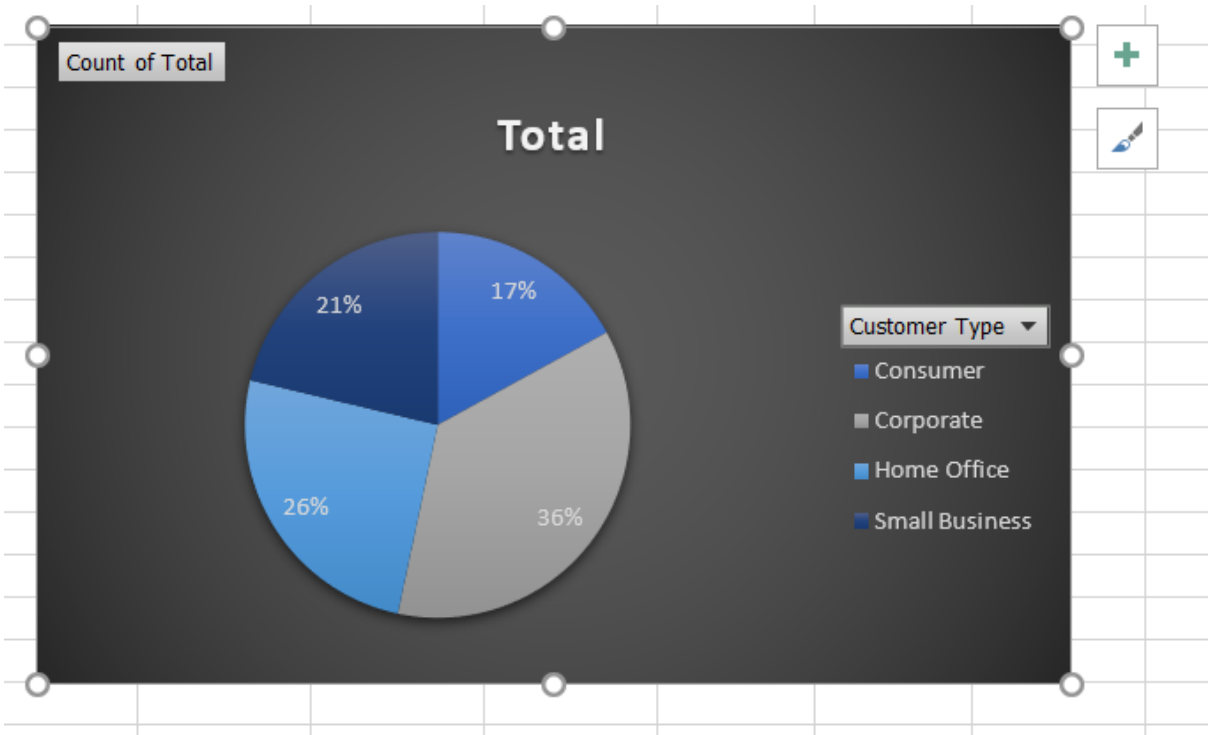
- **Show Values As** – It shows relative values, i.e., percentage of column total or rows total. Simply right click on any of the values cells from the pivot table and select the desired option.



- **Grouping** – Observe the group section in the **Analyze tab** within **Pivot Table tools**. This is where we can group our data and get the summary as per the requirements in the Column section of Pivot Table. Simply select a cell from the Column section variable of the Pivot Table, click on grouping and group as per requirement.
  Observe the below picture (it is from another dataset for better understanding), the column contains dates and by grouping we can get the data as per quarters, or years, days or any other way we wish to. Grouping is a vital function of Pivot Tables that allow dynamic analysis of the data as per requirement, one of the examples as shown below.



- **Pivot Charts** – Pivot Charts give a **Visual Presentation of the Pivot Table**. It changes with change in values of the original Pivot table from which it is derived from. A very unusual feature of the Pivot Charts is that they can filter data and represent the derived representation in terms of graphs alongside filtering the Pivot Table as well.



Observe the above shown Pivot Pie Chart (it is derived from a different dataset for better understanding), the drop down **'Customer Type'** basically allows us to sort or filter data and the same changes are represented in the Pivot table as well.

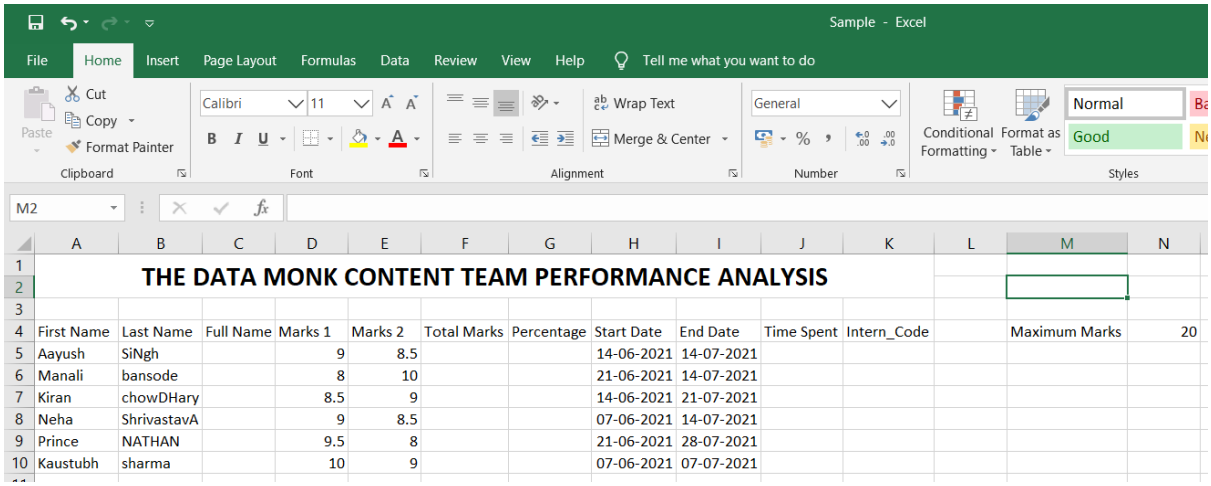
Pivot has a lot more to it than these above-mentioned functionalities. Try to go through the **Pivot Tables Tools – Design and Analyze tabs** and have a click around, maybe you'll find something interesting that may help you organize or present your Pivot table in a better way.

CASE STUDY

Now we are pretty much done with the learning part. Let's try to implement our learnings in the form of a Case Study.

The Data Monk hires 6 interns to prepare two reports. The input sheet given is as follows and we are asked to analyze it, make it presentable and easy to get conclusions out of or a summarized version of it.

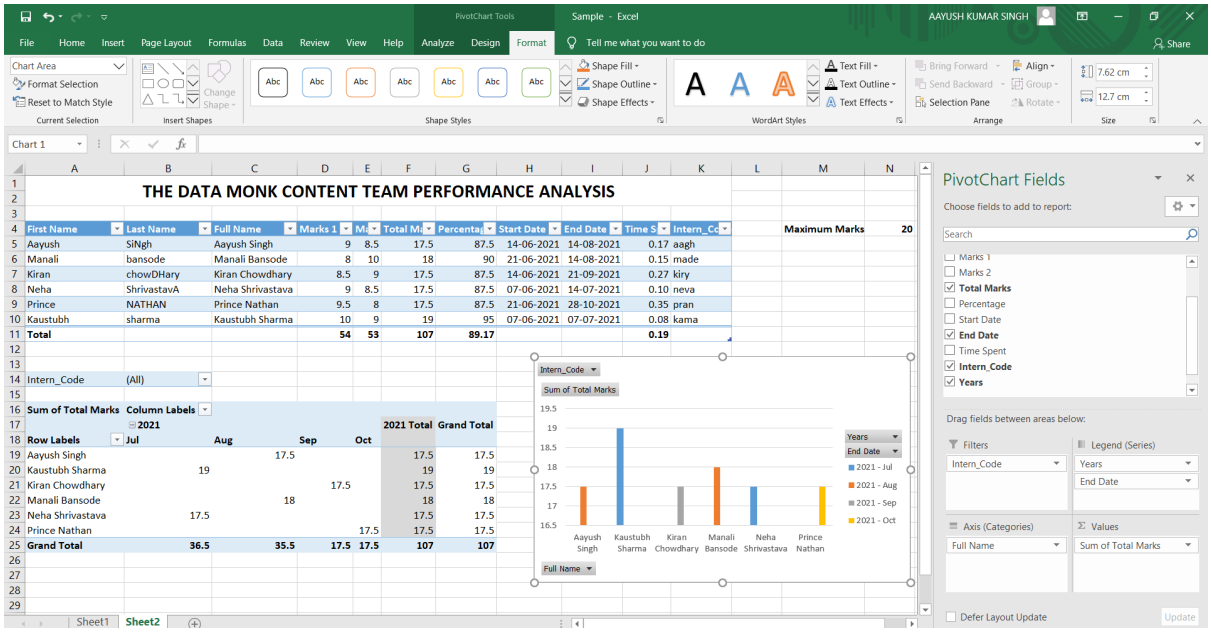
So, the input sheet,



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2	THE DATA MONK CONTENT TEAM PERFORMANCE ANALYSIS													
3														
4	First Name	Last Name	Full Name	Marks 1	Marks 2	Total Marks	Percentage	Start Date	End Date	Time Spent	Intern_Code		Maximum Marks	20
5	Aayush	SiNgh		9	8.5			14-06-2021	14-07-2021					
6	Manali	bansode		8	10			21-06-2021	14-07-2021					
7	Kiran	chowDHary		8.5	9			14-06-2021	21-07-2021					
8	Neha	ShrivastavA		9	8.5			07-06-2021	14-07-2021					
9	Prince	NATHAN		9.5	8			21-06-2021	28-07-2021					
10	Kaustubh	sharma		10	9			07-06-2021	07-07-2021					

- **Full Name** is the combination of **First name** and **Last name** in **Proper Case** with space in between.
- **'Marks 1'** and **'Marks 2'** are the marks assigned to each intern in the first and second report respectively.
- **Maximum Marks** as shown is 20.
- **Time spent** should be the presented in terms of **fraction of the whole year spent.**
- **Intern Code** is the combination of first two characters of the **'First Name'** and last two characters of the **'Last Name'** all in **lower case**.

After doing a set of operations, this is how the sheet looks like



The operations performed are as follows,

1) **Full Name** row is filled with the following Formula **=PROPER(CONCATENATE(A5," ",B5))** and then **fill handle** is used to fill the same for rest of the interns.
2) **Total Marks** are Calculated with the help of following Formula **=SUM(D5:E5)** and **fill handle** is applied to fill the same for rest of the interns.
3) **Percentage** row is filled with the help of following formula **=(F5/N4)*100.**

   Notice Absolute Cell Referencing
4) **Time Spent** is calculated with the help of following formula **=YEARFRAC(H5,I5).**
5) **Intern Code** is estimated with the help of following formula **=LOWER(LEFT(C5,2)&RIGHT(C5,2))**
6) **Named Ranges** is applied in the whole dataset.
7) The range is converted in a **Table** and the **'Total Row'** checkbox is clicked to determine the Total of required rows.
8) A **Pivot Table** and **Pivot Chart** is created out of the Table with
   - The **Column labels** being **'Years'** and **'End Date'** and grouped in **'Years'** and **'Months'**.
   - **Intern Code** being the **filter**
   - **Full Name** being the **Row**
   - **Total Marks** being in the **values** section.

SUMMARY

MS Excel is one of the most widely used software because of its ample number of features and functions that help in Data Analysis, Data Collection and Data Representation. In this article we try to cover major functionalities of excel, their role, the way of implementation and its applications in the real-world scenario.

We start off with a basic description of the interface of Excel Workbook and move on to the icons present in the View Tab and Home Tab of our Excel Worksheet. We study various methods of dealing with multiple worksheets and workbooks and along side explore various icons and features like format painter and many others in the Home Tab.

Somewhere in between we set our eyes on the calculations performed in excel with the help of formulae and functions.

Further we discuss Conditional Formatting in detail, followed by Data Visualization Methods. After this we jump to the important concepts of Date and Text functions where we see the versatile nature of Excel in providing us with function other than numerical related.

Further on, we are introduced to the concept of Named Ranges and we understand the need of the same. After this, we get to learn about Tables in Excel and differentiate them with range. We also learn about pivot table, its method of implementation and various functionalities related to it.

We end our article with a case study where we try to apply our learnings from the whole article.