

Lesson:



Excel Numerical Functions



In the previous session, we learned

- Cell References
- Introduction to a large dataset
- Writing Excel functions
- Important Text Functions

In this session, we will be learning

- Key Numerical Functions in Excel
 - SUM and related functions
 - AVERAGE and related functions
 - COUNT and related functions
 - MAX and MIN functions

Key Excel Functions

We have already learnt the important arithmetic operators in Excel. Now let's get started with some important Excel Functions

- Addition

Basic addition is performed by the "+" operator and "SUM" function.

The syntax of writing a sum statement is

=SUM(value1,value2,...)

The benefit of the "SUM" function over "+" operator is that it is able to handle a large range of values without the user having to enter each cell reference individually. Let's see this in practice in our dataset.

When performing sum operations over a range, instead of entering each cell reference individually, this syntax become

=SUM(range start:range end)

Let's say we want to add the number of reviews for TCS and Accenture –

In the basic SUM function, this can be written as,

=SUM(47000, 31200)

This obviously does not have any benefit over the "+" operator

This can also be written as

=SUM(D2,D3)

This also does not have any benefit over the "+" operator

But let's say you want to add the number of reviews for the first 100 companies. With the "+" operator, this becomes

=D2+D3+D3+.....D101

Typing this entire formula out will be almost impossible. This is where the SUM function really shows its benefit.

This calculation with SUM is simply

=SUM(D2:D101)

We get 966300, which is the final value we are looking for.

But this is not all – Excel can also add multiple ranges in a single SUM function. Let's say you want to add the first 100 numbers, and to it, add numbers in rows 301 to 401. This is very simply

=SUM(D2:D101,D301:D401)

These ranges need not necessarily be in the same column. The following function will also work assuming all the entries in the range are numbers.

=SUM(D2:D101,F2:F101)

In our dataset, this will naturally give an error because entries in column F are not numbers.

Let's practice this

- Sort the dataset in decreasing order (highest to lowest) of number of reviews and calculate the sum of number of reviews of first 150 companies
- In this sorted dataset, calculate the sum of number of reviews of first 150 companies and last 150 companies

SUMIF

Let's say in our dataset we want to calculate the sum of the number of reviews of companies that are tagged as IT Services & Consulting. How do you do this?

One way to do this is to filter the companies which have the tag as "IT Services & Consulting", copy the data to another sheet, and use the SUM function.

The better way to do this is with SUMIF. As the name suggests, it will only SUM values IF a single specified condition is satisfied.

The syntax for SUMIF is

=SUMIF(criteria range, criterion, sum range)

In the case mentioned above, the criteria range will be the column where tags are recorded.

Criteria will be "IT Services & Consulting"

Criterion is the condition – each value in the criteria range is checked against this condition, and only those are selected which meet the condition.

And the sum range will be the column where the number of reviews are recorded. Here it becomes

=SUMIF(I2:I9960,"IT Services & Consulting",D2:D9960)

What SUMIF does is, it selects the criteria range, and the criterion, identifies the entries in the criteria range which matches the criterion, and only adds those corresponding entries.

Interestingly, sum range is optional in this formula – if sum range is not specified, the formula assumes criteria range as the sum range.

Criterion can also be numerical –

Let's calculate the total number of reviews for companies who have been rated 3.9. Here the formula becomes

=SUMIF(C2:C9960,3.9,D2:D9960)

As you can see, whenever a text string is used, we put it in double quotes, and whenever a number is used, we put it without quotes.

This is also a good point to learn about basic excel comparison operators –

| Operator | Use |
|----------|---|
| = | Checks if a value is equal to another value |
| <> | Checks if a value is not equal to another value |
| < | Checks if a value is less than another value |
| > | Checks if a value is greater than another value |
| >= | Checks if a value is greater than or equal to another value |
| <= | Checks if a value is lesser than or equal to another value |

Logical operators always return a BOOLEAN Datatype. This simply means it will return "TRUE" if the condition is true (in the formula =A1=B1, where A1 and B1 are the same) and return "FALSE" if the condition is false (in the formula =A1=B1, where A1 and B1 are not equal). We will learn more about Boolean Datatypes in our upcoming class on Logical functions.

In the case of SUMIF, we can define our criterion using logical operators also

We use double quotes also when we use a logical operator. For example, if we want to calculate the total number of reviews above 3.9 in the dataset,

=SUMIF(C2:C9960,">3.9",D2:D9960)

Exercise – Create a new sheet called “Sum of each tag” and calculate the sum of at least 5 tags in the following format –

| Tags | Total Number of Reviews |
|--------------------|-------------------------|
| BPO / Call Center | |
| Financial Services | |
| Internet | |
| Consulting | |
| BPO/KPO | |

SUMIFS

What about a case where you want to find the number of reviews for all Internet companies that are Private? Here, because you have multiple conditions – Tags = Internet, Company_type = Private, SUMIF will not help. This is where SUMIFS helps. This is a function that adds based on more than one criteria being matched. Syntax: =SUMIFS(sum range, criteria 1 range, criteria 1, criteria 2 range, criteria 2,...)

Sum range is the set of values to be added

Criteria 1 range is the range of values from which criteria 1 needs to be checked

Criteria 1 is the condition which needs to be met in criteria 1 range for the value to be added

And same with criteria 2 and so on

Let us look at another use case for SUMIFS

Exercise - Calculate the total number of reviews for all IT Services and Consulting companies that have rating greater than 3.8

- Average

Average of a set of values = Sum of all values/Total number of values

In Excel, calculating the average of a set of values is done with the AVERAGE functions

Syntax: =AVERAGE(value1, value2,...)

Similar to SUM, we can also use a range here -

For example,

=AVERAGE(A2:A25)

Or

=AVERAGE(A2:A25, A30:A55)

Assuming column A contains some numerical values.

Exercise -

- Calculate the average rating for all the companies given in the dataset
- Arrange the companies in ascending order of name (A to Z) and calculate the average number of ratings in the first 200 companies

AVERAGEIF

Similar to SUMIF, AVERAGEIF calculates the average of only those values in a set which meet a certain criterion.

The syntax for this is very similar to SUMIF

Syntax: =AVERAGEIF(criteria range, criteria, average range)

Criteria range refers to the range of values from which the function will check whether the criteria is met

Criteria is the condition – each value in the criteria range is checked against this condition, and only those are selected which meet the condition. The average is calculated only for those values in the average range which correspond to the selected values in the criteria range.

Similar to SUMIF, if the average range is not specified, the formula assumes the criteria range as the average range.

Exercise -

- Calculate the average number of reviews received by an NGO/NPO (Company_type = NGO/NPO)
- Calculate the average rating of an Automobile company (Tags = Automobile)

AVERAGEIFS

This is similar to the SUMIFS function, where you can calculate the average of a set of values after having it meet multiple criteria.

Syntax: =AVERAGEIFS(average range, criteria 1 range, criteria 1, criteria 2 range, criteria 2,...)

Exercise -

- Calculate the average rating received by Public IT Services & Consulting companies

Count

COUNT function goes through the given range and counts the number of cells in the range that contain numeric data

Syntax: =COUNT(value1, value2, value3,...)

Similar to SUM and AVERAGE, the function can also take a range

For examples

=COUNT(A2:A35,A40:A200)

All numbers are counted, including negative numbers, percentages, dates, times, fractions, and formulas that return numbers. Empty cells and text values are ignored.

Exercise -

Count the number of ratings that has been populated in the dataset

COUNTA

COUNTA function counts the number of cells in a range that are not blank. Any kind of entry – numeric or text, is counted by this function. This means that if there is any entry in a cell – be it numbers or text, it will be counted by the function

Syntax: =COUNTA(value1, value2...)

Similar to the COUNT, this can also take a range as input.

The opposite of COUNTA is COUNTBLANK function, which counts the number of blank cells in a range. Syntax

=COUNTBLANK(range)

Exercise -

- Use COUNTA to calculate the number of missing entries in the No. of Employees column

COUNTIF

Similar to AVERAGEIF, COUNTIF calculates the number of only those cells in a set which meet a certain criterion.

Syntax: =COUNTIF(range, criteria)

Unlike SUMIF and AVERAGEIF, there is no provision to give a separate criteria range and count range – the criteria is checked in the range, and the counting is done in the same range.

Exercise -

- Use COUNTIF to calculate the number of missing entries in the No. of Employees column
- Use COUNTIF to calculate the number Manufacturing companies in the dataset (tags = Manufacturing)

COUNTIFS

COUNTIFS evaluates multiple criteria and calculates the number of cells that match all criteria.

Syntax: =COUNTIFS(criteria range1, criteria 1, criteria range 2, criteria 2,...)

Exercise -

- Calculate the number of FMCG companies that are private in the dataset
- Calculate the number of Financial Services companies which are rated higher than 4

| Function | Operation |
|------------|--------------------------------|
| COUNT | Count numerical values only |
| COUNTA | Count all non-numerical values |
| COUNTIF | Count with one condition |
| COUNTIFS | Count with multiple conditions |
| COUNTBLANK | Count blank cells |

MAX and MIN

MAX and MIN are two simple functions – they return the maximum and minimum numeric value in a dataset

Syntax:

=MAX(value1, value2,...)

=MIN(value1, value2,...)

Like the other functions discussed so far, these can also take range as input

Exercise -

- Find the maximum rating received by a company in the dataset
- Find the number of companies that received the maximum rating (use COUNTIF and MAX together)

Final Exercise

- In the dataset, calculate the number NBFC companies
- Calculate the number of Public NBFC companies
- Calculate the average rating received by an NBFC company
- Calculate the total number of reviews received by NBFC companies that have rating >4.0

Additional Tips: Shortcut

Shift + Space It is used to select the entire row of the current cell

Ctrl + Space It is used to select the entire column of the current cell

Ctrl + - Delete Row/Column

Alt + = Autosum values in the column upto current row

Ctrl + D Copy value from the cell above

Ctrl + R Copy value from the cell to the left