Excel

As a **data analyst**, Excel is one of the foundational tools you’ll use for a wide range of tasks. Here’s what you typically do with Excel in a data analytics role:

**1. Data Cleaning**

**2. Data Analysis**

**3. Data Visualization**

**4. Data Importing & Exporting**

**5. Data Transformation with Power Query**

**6. Forecasting & Modeling**

**7. Automation with VBA (Optional)**

**8. Reporting**

**Microsoft Excel** is a **spreadsheet software** developed by Microsoft. It’s part of the **Microsoft Office suite** and is used for **Organizing data** **Performing calculations** **Analyzing information** **Creating charts and reports.** It’s widely used across industries like finance, marketing, operations, research, and data analytics

A **workbook** is the entire Excel file.

A **worksheet** (also called a "sheet") is a **single page/tab** within the workbook

Ctrl+right arrow = for counting maximum rows and columns

A **range of cells** is simply **two or more cells** selected together in Excel. These cells can be next to each other (contiguous) or separated (non-contiguous). A1:A5 → Cells A1, A2, A3, A4, A5 (a vertical range)

### ****Contiguous Range****

A group of **adjacent** cells.

### ****Non-Contiguous Range****

Multiple **separated** cells or ranges selected at the same time using **Ctrl**

Head of cell is that cell from which range of cell is selected.

Page up or page down + ctrl = moving from one page to another in excel

Formula box = box where formulas of cell contents are there

Name box = box where address of cells are given

Functions are the pre-defined formulas

Ctrl+O = open existing workbook

Ctrl+N = open new workbook

Ctrl+W= close current workbook

We can password protect our workbook from tools option in save and worksheet from right click on sheet tab.

Status bar shows us the status of the cell which is pointed

Ctrl + y = redo

Drag drop cells to copy or write serialize numbers

We can add columns and rows in the worksheet

Sum function = for summation of numbers SUM (A3:B5).

Randbetween = generating random numbers Randbetween(2,99)

### Referencing in Excel refers to the way you refer to cells or ranges of cells in formulas. There are three main types of cell referencing: ****Relative Reference****

* **Example:** =A1
* **Behavior:** Changes when copied to another cell.
* **Use case:** When you want the formula to adapt based on its new location.

**Mixed Reference**

* **Example:** =A$1 or =$A1
* **Behavior:** Part of the reference is fixed:
  + A$1 → Row is fixed.
  + $A1 → Column is fixed. **Use case:** When copying across rows or columns, but you only want part of the reference to adjust.

**Fixed referencing in Excel** refers to a way of keeping a specific cell reference **constant** when copying a formula to other cells. This is done using the **dollar sign ($)** to lock either the column, the row, or both

Conditional Formatting

Change currency format using accounting options , number using number format

Advanced conditional formatting using “New Rules “in which we use formulas to apply formatting on rows by selecting whole sheet like $C1 = “Sales “.

MAX function, MIN function, AVERAGE COUNT

Use clear rules option to remove conditional formatting

Syllabus of excel for data analyst :

Data cleaning & Management = removing duplicates , text to columns , data visualization , flash fill

Formula Mastery = sum , count,average,sumifs,countifs,averageifs,vlookup,hlookup,xlookup,index,match,index & match , if, iferror, and ,or , not , nested functions , array formulas , let , sumproduct , indirect , choose , offset , left , right

Data analysis & Reporting = pivot tables & pivot charts , data sorting & filtering , subtotals , data tables , scenarios ( what-if analysis ) , goal seek and solver

Visualization Expertise = Conditional Formatting , Basic to Advanced Charting , Creating Dynamic Dashboard

Efficiency Enhancers = keyboard shortcuts , data consolidation techniques , error checking

Advanced Excel Capabilities = Advanced filters , Slicers and Timelines in Pivot Tables

Formulas:

If = it is used for conditions If (logical test, value if condition is true, value if condition is false)

Ifs = it is used for multiple conditions =IFS(logical\_test1, value\_if\_true1, [logical\_test2, value\_if\_true2], ...)

The LEN function returns the **number of characters** in a cell, including spaces. LEN( cell reference or string ).

Left = extract characters from the left side of a text string

Syntax: =LEFT (text, [num\_chars])

Right = extract characters from the right side of a text string

Syntax: =RIGHT(text, [num\_chars])

you want to **convert a date to text**, you can use the **TEXT function** =TEXT (value, format\_text)

dd day with zero 01-30

dddd full weekday Monday

mm month with zero 01-12

mmmm full month January

yyyy year 2025

In Excel, the TRIM function is used to remove **extra spaces** from text. It keeps only **single spaces between words** and removes all leading and trailing spaces. =TRIM (text)

If cell A1 contains: Hello World

=TRIM(A1)

Will return: Hello World.

The SUBSTITUTE function in Excel replaces specific text within a string.

=SUBSTITUTE (text, old\_text, new\_text, [instance\_num])

**Ampersand (&) =** is used to **combine text from multiple cells or strings** into one.

=A1 & " " & B1 Does the same thing as CONCATENATE.

The SUMIF function in Excel adds up values **based on a single condition**

=SUMIF (range, criteria, [sum\_range])

* =SUMIF (A2:A4, A2, B2:B4) =SUMIF (B2:B10, ">100")

**sum\_range** (optional): The cells to sum if the condition is met. If omitted, Excel sums the range.

The SUMIFS function in Excel allows you to sum values **based on multiple conditions** — much more powerful than SUMIF.

=SUMIFS(sum\_range, criteria\_range1, criteria1, [criteria\_range2, criteria2], ...)

Example =SUMIFS(C2:C5, A2:A5, "Apple", B2:B5, "East")

Countif = count with a single condition =COUNTIF(range, criteria)

Conutifs = count with multiple conditions =COUNTIFS(range1, criteria1, [range2, criteria2])

Days = calculate **days between dates =DAYS(B1, A1)**

**Click on the corner of cell when we have to drag and drop while copying cells**

NETWORKDAYS = used to **calculate the number of working days** (excluding weekends and optionally holidays) between two dates **=NETWORKDAYS (start \_date, end \_date, [holidays])**

Data Management and Cleaning:

We can auto-fit rows and columns using options given in Format options by selecting whole cols and rows.

We can remove blank rows by using go to option which is under Find & Select Option to select the blank rows and Delete option to delete them.

We can remove blank cells by using go to option under Find & Select Option to select the blank cells and Ctrl+ Enter +zero to place zeros in the blank cells.

We can clear formatting which is unnecessary using the clear option under Home tab.

We can remove duplicates by using Remove Duplicates option under Data Tab.

For converting text to number which is in number, we would write “1” in blank cell copy it select the row which has to be converted then using Paste option we would multiply “1” with whole content of the row to convert it into number.

We can reduce decimal places by using Decrease Decimal Option.

Number Formatting = Through Format cells dialog box which is in Number we use accounting and currency where difference is that in accounting we incline digits and put dollar signs whereas in other we does not incline digits.

Date Formatting = In this we format by using Dates option.

Country Formatting = We would create a new col beside it use proper function to standardize the format trim function to remove extra spaces clean function is used to **remove non-printable characters from text. To delete the original col which is full of errors we’ll copy the whole col then using paste option we would paste values so that its formula is removed**

**Using replace option we can replace under Find & Select option.**

**We can check the spell by Spelling Option under the review Tab.**

**To convert your excel data into table, we would do this by Table Option under Insert Tab.**

**Table Auto-Expand**

A **Pivot Table** in Excel is a powerful tool that lets you **summarize, analyze, explore, and present** large amounts of data quickly—without formulas. It’s especially useful in data cleaning and reporting

Convert table into pivot table by using Summarize with Table Option under Design Tab.

Pivot Chart = First change the format of currency values by using Value Field Setting Option by right clicking. Remove the grand total by using Grand total under Design Tab. Create Pivot Chart using Pivot Chart Option.

Gross Sales = Remove the side measuring scale and put it on every data series.

Format Data Series = Right click on the data series and use Format Data Series Option.

Custom Number Formatting = Using Value Field Setting.

Format Data Labels

We would import text file which has data to excel and use delimiter to draw rows and columns in the data using Text Import Wizard dialog box.

**Flash Fill** helps you automatically fill in values based on patterns **without writing formulas**. It's ideal for tasks like splitting names, combining fields, extracting parts of text.

**Data** tab → Click **Flash Fill.**

**Under Data Tab in Get Data option to get the data. We can combine two almost similar files to one so that the file get updated.**

**Power Query** is Excel’s built-in tool for **importing, cleaning, transforming, and reshaping data** without formulas or VBA. It’s extremely useful for repetitive data-cleaning tasks and can handle large datasets with ease.

Power Query Editor – From here we would combine the files.

Split Columns, Trim Columns

We can convert Power Query into Pivot Table using Close & Load Option. Refresh Option

We can change the currency in the pivot table.

We can drop down the sub-list in the list in pivot table

In the Value Field Setting we can change the values of numerical figures and the col name too.

We can sort the entities in the pivot table.

We can create calculated field by using calculated field option under Field, Items &Sets

Selecting the whole worksheet by Ctrl + Enter +Right Arrow +Down Arrow.

Applying Filters in Pivot Table

XLOOKUP is a modern and powerful Excel function used to search a range or array for a match and return a corresponding value from another range or array.

=XLOOKUP( lookup\_value , lookup\_array , return\_array , [if\_not\_found], [match\_mode], [search\_mode])

|  |  |
| --- | --- |
| lookup\_value | The value you want to search for. |

|  |  |
| --- | --- |
| lookup\_array | The array or range where you want to search for the lookup\_value. |

|  |  |
| --- | --- |
| return\_array | The array or range containing the values to return. |

|  |  |  |
| --- | --- | --- |
| if\_not\_found *(optional)* | | Value to return if no match is found |
| match\_mode *(optional)* | 0 = exact match (default), -1 = exact or next smaller, 1 = exact or next larger, 2 = wildcard | | |

|  |  |
| --- | --- |
| search\_mode *(optional)* | 1 = first-to-last (default), -1 = last-to-first |
|  |  |

In wildcard mode “\*”&

Xlookup for multiple rows is done by selecting multiple rows , Xlookup Exact Match here we use match mode , Xlookup Search Order here we use search mode , Xlookup Horizontal

If XLOOKUP returns an array across **multiple columns**, you can wrap it in SUM:

Example : =SUM(XLOOKUP("Apple", A2:A4, B2:D4))

VLOOKUP (Vertical Lookup) is one of the most commonly used functions in Excel to **find a value in the first column of a table and return a value from a specified column in the same row**.

=VLOOKUP(lookup\_value, table\_array, col\_index\_num, [range\_lookup])

|  |  |
| --- | --- |
| lookup\_value | The value to search for in the **first column** of the table |
| table\_array | The range of cells that contains the data (e.g., A2:D10) |
| col\_index\_num | The **column number** in the table from which to retrieve the value |
| range\_lookup *(optional)* | FALSE = exact match, TRUE = approximate match (default is TRUE) |

Under Data Tab where there is Filter option this is used to sort by selecting the whole row like sort by color so that the highlighted one comes to top of the col.

Manage the rules in the Conditional Formatting Option under Home Tab

Color scales, Highlight Cells, using formula in the manage rules option, troubleshooting to correct the formula based formatting, creating a rule, editing a rule.

Switch data between x axis and y axis in switch row & col option under design tab

Change the chart style under design tab.

Filter Option in charts, move chart to another sheet using Move Chart Option, Change Chart Type Option is used to change the type of chart.

Add Chart elements, chart editing option in the chart.

We would copy the data to another worksheet so that we would not touch the original data.

Ctrl + A is used to select the whole data in the worksheet

We first clean the data, edit and format the data, make its pivot table which we have to visualize

The data and the pivot table are interconnected and we use the refresh option under the Data tab.

We use Pivot Charts for making dashboards using Pivot Tables, we make dashboards which are those charts, we create a header while we are creating that Dashboard

A **Slicer** in Excel is a **visual filter tool** that lets you easily filter data in **PivotTables, Pivot Charts**, or **tables** using buttons, rather than drop-down menus. It is under Analyze Tab.`

Report Connections under Options Tab is used to manage which Pivot tables Slicer is connected.

Invisible Grid Lines under View Tab.

By double clicking on the Home Tab we can hide the Menu Bar.

Lifecycle of Data Analysis using Excel:

Extract Data (Power Query Editor) ---> Cleaning PQE---> Transformation PQE-ETL Too---> Data Modeling (Power Pivot & Measures) ----> Data Analysis (Pivot Tables) -----> Dashboard Pivot Chart (Executive Summary).

Power Pivot is an advanced data modeling and analysis add-in for **Microsoft Excel**, primarily used for:

**Data modeling** is the process of organizing and structuring data so it can be stored, retrieved, and analyzed efficiently. In tools like **Power Pivot** (or databases like SQL), it means:

Combine data from multiple sources (Excel tables, databases .Create relationships between tables.

We use folder while collecting data as we can add new file in it and it can map with other file as well.

Column profiling in the Power Query Editor is done for only some columns we have to add columns for profiling further.

Ctrl + X = Cut

Conditional Columns, Replace Values, Extract Options, Format options, Custom Column, we can change the format of data in the column, Date option, Time option, Close & Load Too option, Developer Tab from File Tab 🡪 options, Power Pivot Tab from Developer Tab from COM Add-ins options, to go from pivot table to power query editor from Get Data Option under Data Tab , Power Pivot 🡪 Manage , Value Filters in Pivot Table

We can copy pivot table or we can create a new table, we can change using another graph by copy paste , pivot field list of pivot table , we will not use the data in its raw form we would use the table of that data , Slicer Tab while using slicer , report connections in slicer , slicer styles.

In **Excel**, a **measure** is a calculation used in **PivotTables**, especially when working with **Power Pivot** or the **Data Model**. It is a powerful tool for analyzing large datasets using **DAX (Data Analysis Expressions)** formulas. A **measure** is a **calculated field** used to perform **aggregations ,** it is under Power Pivot.

We can copy paste in the Dashboard , we can add icons through Icon Option ,

A **KPI** is a measurable value that shows how effectively a person, team, or organization is achieving a key business objective. Organizations use KPIs to help individuals and departments focus on what matters most and to track progress towards goals.

To push the dashboard to Github = select all the charts using Ctrl , right click then option appears to lock the worksheet , from format option protect the worksheet.