

EXCEL FOR INTERVIEWS

Sample table

Link :- https://github.com/vanshika230/Data_Analyst_Preparation/blob/main/Dataset/ED.xlsx

A	B	C	D	E	F	G	H	I	J	K	L	M	N	
Emp ID	First Name	Last Name	Gender	Department	SALARY	Start Date	FTE	Employee type	Work location	Tenure	Work Type	Salary	Bonus Elig	Full Name
PR0007	Torrance	Collier	Female	Training	96136	7/13/2020	0.3	Wellington, New	3.94	Part time	High	Not Eligible	Torrance Collier	
PR0095	Devarat	Damarsingh	Male	Business Dev	70649	1/17/2020	1	Hyderabad, Indi	4.44	Full time	High	Eligible	Devarat Damarsingh	
PR0113	Van	Tuxwell	Female	Business Dev	80696	9/17/2018	1	Columbus, USA	5.76	Permanent	High	Eligible	Van Tuxwell	
PR0147	Minerva	Ricardot	Male	Engineering	120000	11/12/2018	1	Remote	5.61	Permanent	High	Not Eligible	Minerva Ricardot	
PR0246	Husein	Augar	Female	Marketing	67906	12/29/2020	1	Remote	3.48	Permanent	Medium	Not Eligible	Husein Augar	
PR0419	Billi	Fellgate	Female	Business Dev	68981	1/29/2019	0.8	Remote	5.4	Permanent	Medium	Eligible	Billi Fellgate	
PR0576	Lalitchandra	Vadali	Female	Legal	63705	8/5/2019	1	Hyderabad, Indi	4.88	Full time	Medium	Not Eligible	Lalitchandra Vadali	
PR0746	Hogan	Iles	Male	Accounting	114177	3/18/2020	1	Wellington, New	4.26	Full time	High	Not Eligible	Hogan Iles	
PR0770	Beryl	Burnsyde	Male	Legal	29775	10/1/2020	1	Wellington, New	3.72	Full time	Low	Not Eligible	Beryl Burnsyde	
PR0882	Jill	Shipsey	Male	Accounting	52964	4/2/2021	1	Columbus, USA	3.22	Part time	Medium	Not Eligible	Jill Shipsey	
PR0893	Vasavi	Veeravasarapu	Female	Human Resou	50310	3/30/2021	0.4	Hyderabad, Indi	3.23	Part time	Medium	Not Eligible	Vasavi Veeravasarapu	
PR0916	Inger	Chapelhow	Female	Research and	84310	11/1/2021	1	Remote	2.64	Full time	High	Not Eligible	Inger Chapelhow	
PR1055	Devasree	Fullara	Male	Accounting	35936	5/21/2018	1	Chennai, India	6.09	Full time	Low	Not Eligible	Devasree Fullara	
PR1159	Mollie	Hanway	Male	Engineering	112641	10/21/2019	1	Seattle, USA	4.67	Part time	High	Not Eligible	Mollie Hanway	
PR1211	Enoch	Dower	Male	Accounting	91645	1/27/2021	0.6	Auckland, New	3.4	Full time	High	Not Eligible	Enoch Dower	
PR1269	Eleonore	Airdrie	Female	Engineering	97105	8/17/2021	1	Columbus, USA	2.85	Full time	High	Not Eligible	Eleonore Airdrie	
PR1306	Patti	Dradey	Female	Services	84743	9/24/2020	1	Auckland, New	3.74	Full time	High	Not Eligible	Patti Dradey	
PR1346	Adolph	McNalley	Male	Business Dev	85919	2/5/2018	1	Columbus, USA	6.38	Full time	High	Eligible	Adolph McNalley	
PR1383	Addi	Studdeard	Female	Product Mana	72503	4/28/2021	0.3	Wellington, New	3.37	Full time	High	Not Eligible	Addi Studdeard	
PR1476	Jaishree	Atasi	Male	Services	12000	11/12/2018	1	Chennai, India	6.16	Full time	Low	Not Eligible	Jaishree Atasi	
PR1662	Genevra	Friday	Female	Research and	50449	11/14/2018	0.8	Auckland, New	5.61	Part time	Medium	Not Eligible	Genevra Friday	

I have named the table as hr. Press Ctrl+T and create the table. Name it in the table design option in excel.

1. EXCEL FUNCTIONS

→ BASIC MATH

Function	Parameters	Purpose	Example Formula	Example Output Column
SUM()	range	Add values	=SUM(hr[salary])	Total of all salaries
AVERAGE()	range	Mean value	=AVERAGE(hr[age])	Average employee age
COUNT()	range	Count numeric cells	=COUNT(hr[salary])	No. of numeric salary entries
COUNTA()	range	Count non-empty cells	=COUNTA(hr[department])	Filled department rows
COUNTBLANK()	range	Count blank cells	=COUNTBLANK(hr[email])	Number of missing emails
PRODUCT()	range	Multiply values	=PRODUCT(B2:B4)	Product of 3 numbers
ROUND()	number, decimals	Round to decimals	=ROUND(89.678, 1)	89.7
ROUNDUP()	number, decimals	Always round up	=ROUNDUP(4.32, 0)	5
ROUNDDOWN()	number, decimals	Always round down	=ROUNDDOWN(4.99, 0)	4
INT()	number	Nearest lower integer	=INT(5.99)	5
TRUNC()	number, num_digits	Truncate decimals	=TRUNC(6.78)	6

→ Conditional Aggregation

Function	Parameters	Purpose	Example Formula	Example Output Column
SUMIF()	range, criteria, sum_range	Sum with condition	=SUMIF(hr[department], "Finance", hr[salary])	Total salary in Finance
AVERAGEIF()	range, criteria, avg_range	Avg with condition	=AVERAGEIF(hr[gender], "Female", hr[salary])	Avg salary of female employees
COUNTIF()	range, criteria	Count matches	=COUNTIF(hr[gender], "Male")	No. of male employees
SUMIFS()	sum_range, criteria1, criteria2...	Sum with multiple filters	=SUMIFS(hr[salary], hr[gender], "Male", hr[department], "HR")	Salary of male HR employees
AVERAGEIFS()	avg_range, criteria1, criteria2...	Avg with multiple filters	=AVERAGEIFS(hr[salary], hr[gender], "Female", hr[department], "IT")	Avg salary of female IT staff
COUNTIFS()	criteria1, value1, ...	Count with multiple filters	=COUNTIFS(hr[gender], "Female", hr[department], "IT")	No. of female employees in IT

→ Lookup & Reference

Function	Parameters	Purpose	Example Formula	Example Output
VLOOKUP()	Value to search, table, col_index to return, FALSE (exact match)	Search by row	=VLOOKUP(102, hr, 3, FALSE)	Salary of employee ID 102
HLOOKUP()	Value to search, table, row_index to return, FALSE (exact match)	Search by column	=HLOOKUP("Q1", A1:D2, 2, FALSE)	Value for Q1
XLOOKUP()	lookup_value, search_col, return_col	Best modern lookup	=XLOOKUP(102, hr[emp_id], hr[salary])	Salary for emp_id 102
INDEX()	range, row_num	Value by position	=INDEX(hr[name], 3)	3rd employee name
MATCH()	value, range, 0	Position of value	=MATCH("HR", hr[department], 0)	Row number where "HR" appears

→ Statistical Functions

Function	Parametes	Purpose	Example Formula	Example Output
MEDIAN()	Range	Middle value	=MEDIAN(hr[salary])	Median salary
MODE.SNGL()	range	Most frequent value	=MODE.SNGL(hr[experience])	Most common years of experience
STDEV.S()	range	Std deviation (sample)	=STDEV.S(hr[salary])	Spread in salaries
VAR.S()	range	Sample variance	=VAR.S(hr[age])	Variability in age

RANK.EQ()	value, range, order	Rank of value	=RANK.EQ(72000, hr[salary], 0)	Rank of 72000 salary
CORREL()	range1, range2	Correlation	=CORREL(hr[experience], hr[salary])	Strength of relationship
PERCENTILE.INC()	range, k	Percentile	=PERCENTILE.INC(hr[salary], 0.9)	90th percentile salary
QUARTILE.INC()	range, quart	Quartile	=QUARTILE.INC(hr[salary], 3)	3rd quartile of salaries

→ Logical Functions

Function	Parameters	Purpose	Example Formula	Example Purpose
IF()	condition, true, false	Basic logic	=IF(hr[salary]>50000, "High", "Low")	Categorize salary
IFS()	Multiple condition, value pairs	Multi-branch logic	=IFS(hr[salary]>70000,"High", hr[salary]>50000,"Med", TRUE,"Low")	Label salary tier
AND()	cond1, cond2	All conditions true	=IF(AND(hr[age]>25, hr[gender]="Male"), "Yes", "No")	Check multiple conditions
OR()	cond1, cond2	Any condition true	=IF(OR(hr[department]="HR", hr[age]<30), "Flag", "Clear")	Alert on multiple cases
NOT()	cond	Reverse logic	=NOT(hr[gender]="Male")	Is not male?
SWITCH()	value, match1, val1, ...	Multiple matching logic	=SWITCH(hr[grade], "A","Top", "B","Good", "C","Average")	Grade remark

→ Text Functions

Function	Parameters	Purpose	Example Formula	Example Purpose
TEXTJOIN()	delimiter, ignore_empty, text1, text2	Join multiple columns	=TEXTJOIN(" ", TRUE, hr[first_name], hr[last_name])	Full name
CONCAT()	text1, text2	Merge strings	=CONCAT("Emp-", hr[emp_id])	ID like Emp-101
LEFT()	text, n	First n characters	=LEFT(hr[name], 3)	"Van" from "Vanshika"
RIGHT()	text, n	Last n characters	=RIGHT(hr[emp_id], 2)	"01" from "EMP01"
MID()	text, start, length	Middle substring	=MID(hr[email], 3, 4)	3rd to 6th chars
LEN()	text	Character count	=LEN(hr[name])	Length of name
TRIM()	text	Remove extra spaces	=TRIM(" hello ")	"hello"
UPPER() / LOWER()	text	Case change	=UPPER(hr[dept]) → "FINANCE"	
PROPER()	text	Title Case	=PROPER("vanshika mishra") → "Vanshika Mishra"	

TEXT()	value, format	Format dates or numbers	=TEXT(hr[joining_date],"yyyy-mm")	"2024-03"
SUBSTITUTE()	text, old, new	Replace value	=SUBSTITUTE(hr[dept], "IT", "InfoTech")	Replace dept names
FIND() / SEARCH()	find, within	Position of substring	=FIND("@", hr[email])	Where @ is in email

→ Date & Time

Function	Parameters	Purpose	Example Formula	Example Output
TODAY()	None	Returns current date	=TODAY()	2025-07-10
NOW()	None	Returns current date & time	=NOW()	2025-07-10 00:12
DAY()	date	Extracts day from a given date	=DAY("26-01-2024")	26
MONTH()	date	Extracts month from a given date	=MONTH("26-01-2024")	1
YEAR()	date	Extracts year from a given date	=YEAR("26-01-2024")	2024
WEEKDAY()	date, [return_type]	Returns weekday number (1=Sunday, 2=Monday...)	=WEEKDAY("26-01-2024", 1)	6 (Friday)
WEEKNUM()	date, [return_type]	Returns the week number in the year	=WEEKNUM("26-01-2024")	4
ISOWEEKNUM()	date	Returns ISO week number (starts on Monday)	=ISOWEEKNUM("26-01-2024")	4
DATE()	year, month, day	Creates a date from year, month, and day	=DATE(2024, 1, 26)	26-01-2024
DATEVALUE()	text	Converts a text date into an actual Excel date	=DATEVALUE("01-Jan-2024")	01-01-2024
EDATE()	date, months	Adds or subtracts months	=EDATE("21-01-2024", 5)	21-06-2024
EDATE()	date, -months	Subtracts months	=EDATE("21-01-2024", -5)	21-08-2023
EDATE()	date, years*12	Adds or subtracts years by multiplying by 12	=EDATE("21-01-2024", 5*12)	21-01-2029
DATEDIF()	start, end, "Y"	Calculates years between dates	=DATEDIF("24-01-2023", "17-02-2024", "Y")	1
DATEDIF()	start, end, "M"	Calculates months between dates	=DATEDIF("24-01-2023", "17-02-2024", "M")	12
DATEDIF()	start, end, "D"	Calculates days between dates	=DATEDIF("24-01-2023", "17-02-2024", "D")	389
YEARFRAC()	start, end, [basis]	Returns fractional years between two dates	=YEARFRAC("01-Jan-2020", "26-Jan-2024")	4.07
NETWORKDAYS()	start, end	Returns business days excluding weekends	=NETWORKDAYS("01-01-	260

			2023", "31-12-2023")	
NETWORKDAYS()	start, end, holidays_range	Business days excluding holidays as well	=NETWORKDAYS("01-01-2023", "31-12-2023", F34:F35)	258
WORKDAY()	start, days, holidays_range	Returns date after n working days	=WORKDAY("01-01-2022", 10, F34:F35)	14-01-2022
EOMONTH()	date, 0	Returns last day of current month	=EOMONTH("01-01-2023", 0)	31-01-2023
EOMONTH()	date, -1	Returns last day of previous month	=EOMONTH("02-04-2023", -1)	31-03-2023
EOMONTH()	date, 2	Returns last day two months ahead	=EOMONTH("04-08-2023", 2)	31-10-2023
+ (Date + Days)	date + n	Adds n days to a date	="21-01-2024" + 5	26-01-2024
- (Date - Days)	date - n	Subtracts n days from a date	="21-01-2024" - 5	16-01-2024

➔ Forecast

Function	Parameters	Purpose	Example Formula	Example Purpose
FORECAST()	x, known_ys, known_xs	Predict value	=FORECAST(2025, hr[salary], hr[year])	Projected salary
TREND()	known_ys, known_xs, new_xs	Linear fit	=TREND(hr[salary], hr[year], {2025})	Estimate
LINEST()	y_range, x_range	Regression	=LINEST(Ys, Xs)	Slope and intercept
GROWTH()	Ys, Xs, newX	Exponential forecast	=GROWTH(Ys, Xs, 2025)	Non-linear growth

➔ Error Handling & Info

Function	Parameters	Purpose	Example Formula	Example Purpose
IFERROR()	formula, fallback	Handle errors	=IFERROR(A1/B1, "Invalid")	Prevents crash
ISNUMBER()	value	Check if number	=ISNUMBER(hr[score])	TRUE/FALSE
ISBLANK()	value	Check empty cell	=ISBLANK(hr[email])	TRUE if missing
ISTEXT()	value	Check text	=ISTEXT(hr[dept])	TRUE if text
CLEAN()	text	Remove non-printable	=CLEAN(A1)	Text only
FORMULATEXT()	cell	Show formula	=FORMULATEXT(A1)	=SUM(A1:A10)

→ FILTER

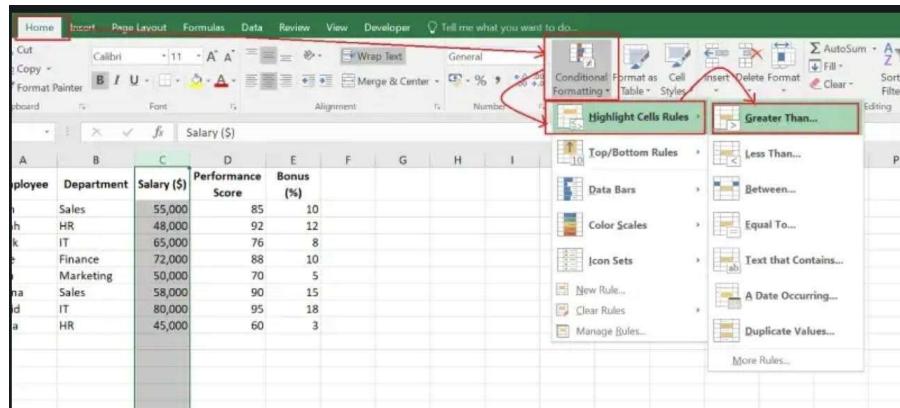
Use Case	Formula	Purpose	Example Output Column
1. Basic Filter by One Condition	=FILTER(hr[name], hr[department]="Finance")	Get names of employees in Finance department	List of Finance employees
2. Filter by Numeric Condition	=FILTER(hr[name], hr[salary]>60000)	Names with salary > 60K	High-salary employees
3. Filter by Date	=FILTER(hr[name], hr[joining_year]>2020)	Employees who joined after 2020	Post-2020 joiners
4. Filter by Month	=FILTER(hr[name], MONTH(hr[joining_date])=3)	Employees who joined in March	March joiners
5. Filter by Gender and Department (AND)	=FILTER(hr[name], (hr[department]="HR") * (hr[gender]="Female"))	Female employees in HR	Filtered HR females
6. Filter by Gender OR Department	=FILTER(hr[name], (hr[gender]="Female") + (hr[department]="Finance"))	Female or Finance dept	Union filter output
7. Filter Salaries by Department	=FILTER(hr[salary], hr[department]="IT")	Salaries in IT dept	IT salary list
8. Filter with Text Function	=FILTER(hr[name], LEFT(hr[name],1)="A")	Names starting with 'A'	A-name employees
9. Top 3 Salaries in a Dept	=TAKE(SORT(FILTER(hr[salary], hr[department]="IT"), , -1), 3)	Top 3 salaries in IT	Sorted descending
10. Filter with IFERROR (no match)	=IFERROR(FILTER(hr[name], hr[dept]="Admin"), "No records")	Handle no match cases	"No records" if none
11. Filter with ISNUMBER	=FILTER(hr[salary], ISNUMBER(hr[salary])))	Exclude errors/blanks	Valid salary values
12. Filter with Multiple Columns Returned	=FILTER(hr[[name],[salary]], hr[department]="Finance")	Name & salary of Finance dept	Multi-column filter
13. Filter Employees with Length > 6 Letters	=FILTER(hr[name], LEN(hr[name])>6)	Long names	Names like "Vanshika"
14. Filter by Not Equal	=FILTER(hr[name], hr[department]<>"IT")	All non-IT employees	Exclude IT dept
15. Filter Top K by Condition	=LARGE(FILTER(hr[salary], hr[department]="HR"), 2)	2nd highest salary in HR	Top-K from filtered

How to Do Conditional Formatting in Excel

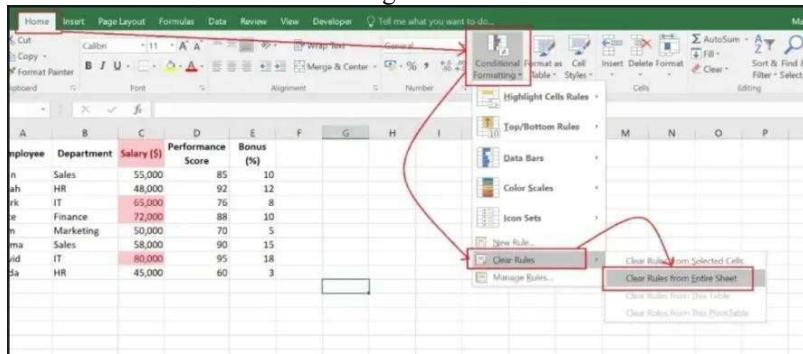
- Select the Cells to Format
- Navigate: Home Tab → Conditional Formatting
- Choose a Formatting Rule
- Set Condition and Style
- Click OK

Basic Conditional Formatting Techniques

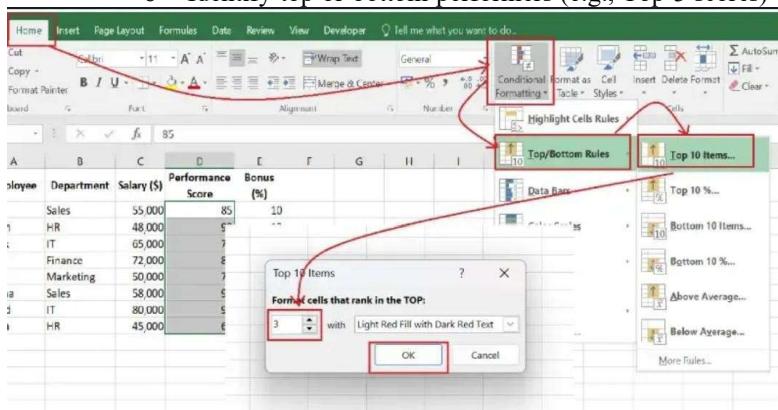
- **Highlight Cells Based on Values:**
 - Use rules like greater than, less than, between
 - Example: Highlight salaries over \$60,000 in red.



- **Clear Rules:**
 - Remove formatting from selected cells or entire sheet

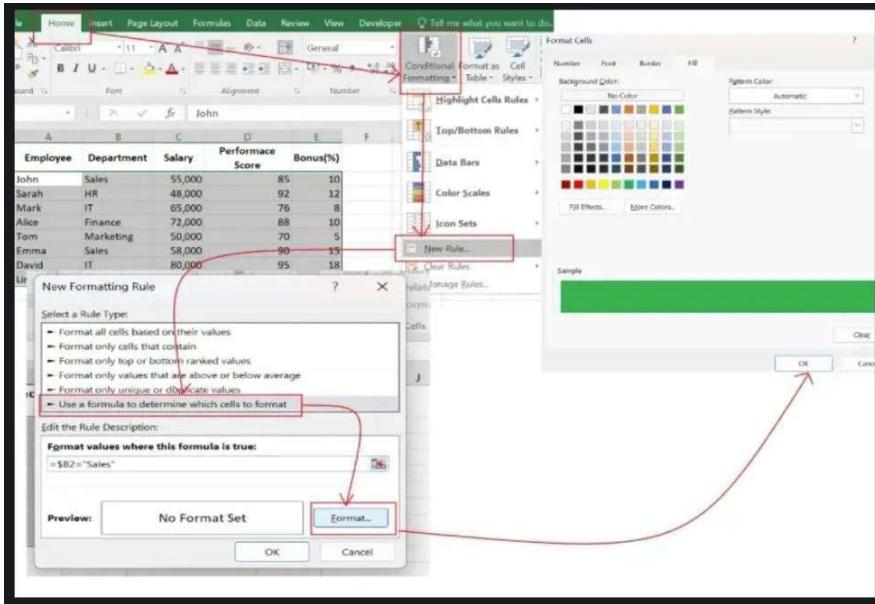


- **Top/Bottom Rules:**
 - Identify top or bottom performers (e.g., Top 3 scores)



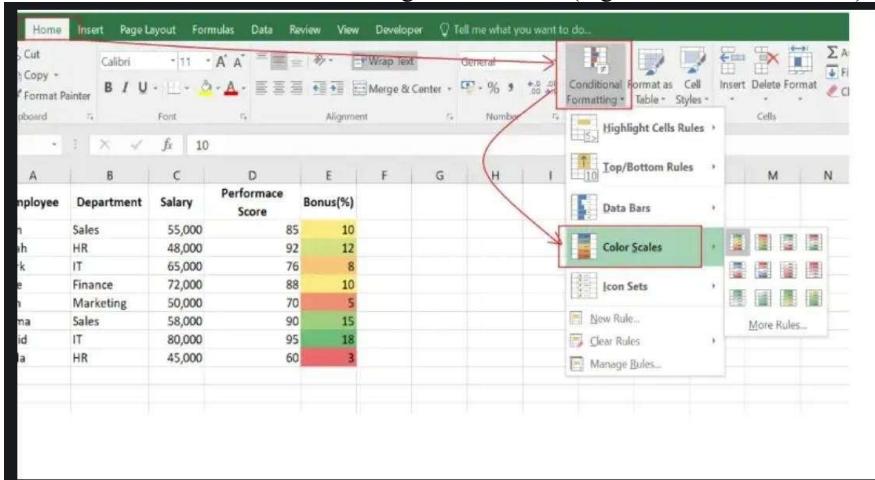
- **Conditional Formatting with Formulas:**

- Create custom rules with formulas
- Example: Highlight employees in the Sales department with $=\$B2="Sales"$



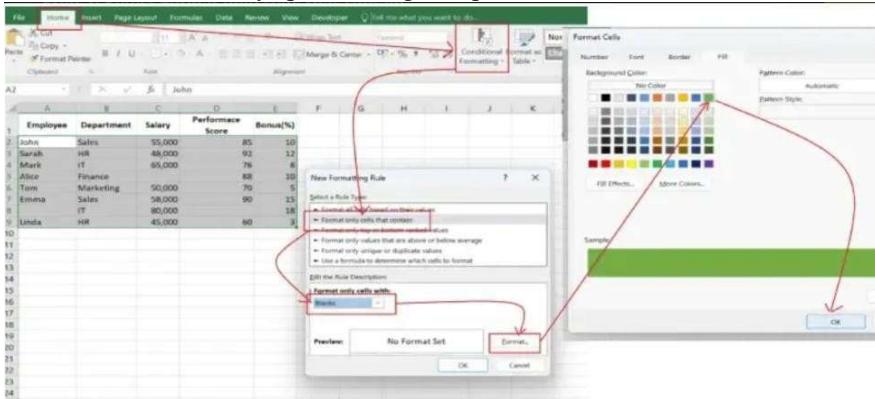
- **Using Color Scales:**

- Visualize data with gradient colors (high, medium, low values)



- **Highlighting Blank or Duplicate Cells:**

- Quickly spot missing or repeated data



Editing and Removing Conditional Formatting

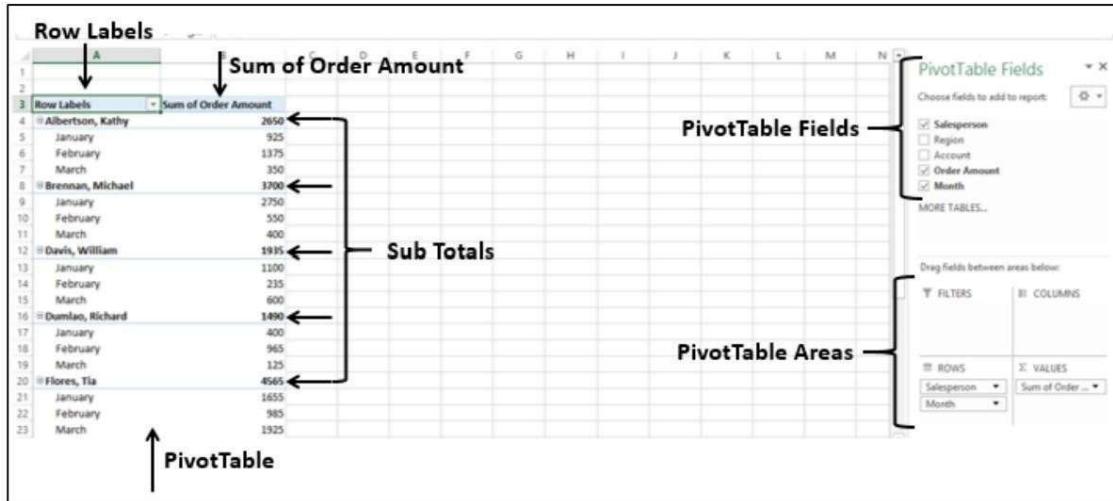
- **Edit Rules:**

- Home Tab → Conditional Formatting → Manage Rules
- Modify settings and apply

- **Remove Rules:**

- Clear Rules option or use keyboard shortcut (Alt + H + E + F)

PIVOT TABLES



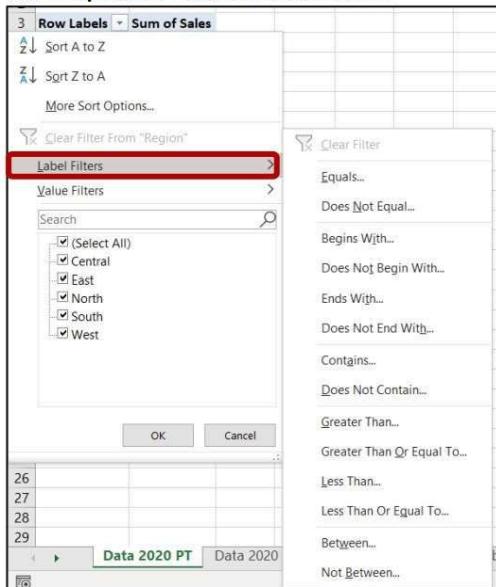
Filtering Data

- Purpose: To view specific data by narrowing down visible information.
- Filter Button in B1:**
 - Displays automatic filter button in PivotTable.
 - Shows “Multiple Items” when filtering multiple items.
 - Label outside the PivotTable to clarify applied filters.
 - Always clear filters before continuing work.

Row and Column Filters

- Accessing Filters:**
 - Click the drop-down arrow next to Row or Column Labels.
 - Use Label Filters for custom filtering.
 - Traditional filters available throughout Excel.

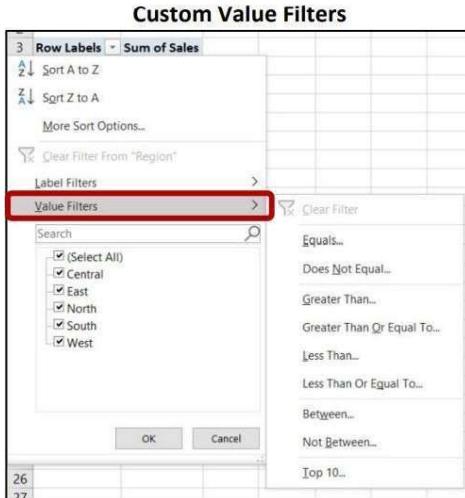
Filter Options > Custom Label Filters



- Clear Filter**
- Equals...
- Does Not Equal...
- Begins With...
- Does Not Begin With...
- Ends With...
- Does Not End With...
- Contains...
- Does Not Contain...
- Greater Than...
- Greater Than Or Equal To...
- Less Than...
- Less Than Or Equal To...
- Between...
- Not Between...

Value Filters

- Options Specific to Values:**
 - Includes “Top 10...” filter.
 - Custom value filters help when filtering long lists.



Search Box Filter (Office 365 Excel)

- Helps locate a single entry in long lists.
- Steps:
 - Start typing in the Search box → AutoComplete suggests entries.
 - Press **Enter** to apply the filter.
 - To add more items, type again and select "**Add current selection to filter.**"
 - To clear filters:
 - Click **Select All** or **Clear Filter** option.

Right-Click Filtering

- Options:
 - Right-click an item → Select **Filter**.
 - Choose:
 - **Keep Only Selected Items**
 - **Hide Selected Items**
- Keyboard Shortcuts:
 - **Ctrl** → Select noncontiguous items.
 - **Shift** → Select a contiguous range.

Filters in PivotTables

- Filters impact overall PivotTable results.
- **Filter Area Box:**
 - Fields can be placed in the **Filter area** but are not mandatory.

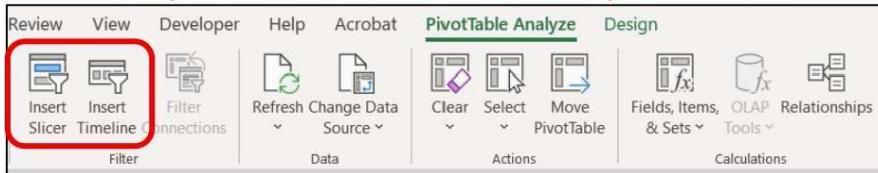
Multiple Filters in PivotTables

- Stack fields in the Filter area for multi-filter selection.
- Change filter layout to horizontal if needed.
- Steps:
 1. Go to **Analyze Tab > PivotTable Group > Options**
 2. Navigate to **Options > Layout & Format > Filter fields per column**

Slicers – Advanced Filters

- **Purpose:** Create multi-filter dashboards for dynamic data filtering.
- **Benefits:**
 - Can filter multiple PivotTables with the same data source.
 - Easy visualization and interaction.
- **How to Insert a Slicer:**
 1. **Analyze Tab > Sort & Filter Command Group > Insert Slicer**
 2. Select columns to filter (e.g., Region, Product).
 3. Apply color, display multiple columns, adjust settings via the **Slicer Tab**.

Analyze Tab > Sort & Filter Command Group > Insert Slicer



Managing Slicers

- **Clear Filters:**
 - Click **Clear Filter** button on the slicer.
 - Use **Multi-Select** button to filter multiple items without **Ctrl**.
- **Remove Slicers:**
 - Right-click a slicer → Select **Remove slicer**.
 - To remove multiple, **Ctrl+click** slicers → Press **Delete**.

Slicer Filter Connections (Connecting Slicers to Multiple PivotTables)

- **Steps to Connect:**
 1. Click on a PivotTable.
 2. Insert a **Slicer** and select relevant fields.
 3. Click **Filter Connections** → Select PivotTables to filter.
- **Marking Tables for Filtering:**
 1. Click the worksheet tab containing the slicers.
 2. Select slicer → **Slicer Tab > Report Connections**.
 3. Mark tables to filter → Click **OK**.
- **Disconnecting Slicers from Tables:**
 - Reverse the above steps → Remove checkmarks from connected tables.

Timeline Slicer (For Time-Based Data Filtering)

- **Purpose:** Filters PivotTables based on time.
- **How to Use:**
 - Select the **time element** (upper-right of the Timeline Slicer).
 - Drag or click the timeline lever to adjust dates.
 - PivotTable updates automatically.
 - Clear filter before leaving the file.
 - To remove: Select the filter → Press **Delete**.

Timeline Slicer

	Row Labels	Sum of Sales
4	Row Labels	Sum of Sales
5	Central	516163
6	East	615828
7	North	672584
8	South	650130
9	West	1536342
10	Grand Total	3991047

Order Date

Q1 2020 MONTHS ▾

2020

JAN FEB MAR APR MAY JUN JUL

Grouping Dates

- Right-click on any **row label value (Column A)** → Select "Group".
- The **Grouping Dialog** appears, showing the **oldest and most recent date**.
- Select a **time increment** (e.g., Quarters, Months).
- The **Order Date** field in the Rows area allows:
 - A **hierarchical layout** (e.g., Quarters and Months).
 - Collapse and Expand buttons** on the left of the Analyze tab.

The screenshot shows two Excel tables side-by-side. The left table, titled 'Grouped by Months', shows sales data from January to December 2020. The right table, titled 'Grouped by Quarters and Months', shows sales data grouped by quarter (Q1, Q2, Q3, Q4) and then by month within each quarter. Both tables have a 'Sum of Sales' column at the bottom.

Month	Sum of Sales
Jan	277788
Feb	202566
Mar	152682
Apr	203455
May	177272
Jun	178934
Jul	172949
Aug	221727
Sep	242653
Oct	149245
Nov	83230
Dec	225944
Grand Total	2288445

Quarter	Sum of Sales
Q1	633036
Jan	277788
Feb	202566
Mar	152682
Q2	559661
Apr	203455
May	177272
Jun	178934
Q3	637329
Jul	172949
Aug	221727
Sep	242653
Q4	458419
Oct	149245
Nov	83230
Dec	225944
Grand Total	2288445

Grouping Numbers

- Helps **simplify** large numeric datasets by creating **ranges** instead of listing every value.
- Grouped numbers display as **numeric text values**.

Steps to Group Numbers

- Initial **PivotTable** displays a list of individual numbers.
- Right-click on any **number in Column A** → Select "Group".
- Grouping Dialog** appears with the smallest and largest number.
- Adjust the **range of values** as needed.
- Click **OK** to apply the grouping.
- Example: **Sales grouped into ranges** (e.g., Coleman has sales in the largest range).

The screenshot shows a PivotTable and its corresponding 'Grouping' dialog box. The dialog box shows 'Starting at' 12600 and 'Ending at' 53921, with a 'By:' step of 10000. The PivotTable to the right shows sales data for five categories: Appleby, Carney, Close, Coleman, and Jacobs. The data is grouped by a range of sales values, with the largest group being 12600-22599.

Category	Appleby	Carney	Close	Coleman	Jacobs	Grand Total
12600-22599	65,097	81,857	139,224	60,894	33,284	380,356
22600-32599	55,902	56,999	240,115	27,818	111,496	492,330
32600-42599	110,852	77,885	338,248	153,462	79,483	759,930
42600-52599	89,626	93,332	280,161	45,000	93,789	601,908
52600-62599				53,921		53,921
Grand Total	321,477	310,073	997,748	341,095	318,052	2,288,445

Custom Text Groups

- Allows **manual grouping** of selected data.
- Useful for categorizing data like:
 - States into regions**
 - Products into product types**

Steps to Create Custom Text Groups

- Select multiple **contiguous** (Shift) or **non-contiguous** (Ctrl +key) items.
- Right-click → Select "Group".
- Enter a **custom label** to replace the automatic label.
- The **new group name** appears in the **filter list** for selection.
- Collapse/expand groups using:
 - Collapse button** next to the group name.

- Analyze tab > Active Field group.

Original PivotTable			Grouped PivotTable		
3	Row Labels	Sum of Sales	3	Row Labels	Sum of Sales
4	Apples	127,662	4	Fruit	362,814
5	Bananas	150,837	5	Apples	127,662
6	Green Beans	172,704	6	Bananas	150,837
7	Lettuce	238,684	7	Oranges	84,315
8	Oranges	84,315	8	Vegetables	550,837
9	Peas	139,449	9	Green Beans	172,704
10	Grand Total	913,651	10	Lettuce	238,684
			11	Peas	139,449
			12	Grand Total	913,651

Important Notes

- Custom groups introduce new terms that are not in the original dataset.
- Ensure to document custom changes for better collaboration.
- Ungrouping:
 - Click Ungroup button or Right-click > Ungroup.
 - Works for all date, numeric, and text fields.
 - Text groups can be individually ungrouped.

Summarize Values By

- Default Calculation Behavior:
 - Numeric fields → Sum.
 - Text fields → Count.
 - If a numeric field has text or blanks, the function defaults to Count.
- Right-click on a numeric cell → Summarize Values By → Choose function.
- Additional settings available via:
 - Analyze Tab > Calculations Command Group > Summarize Values By.
 - Value Field Settings dialog for quick adjustments.

The screenshot shows a Microsoft Excel spreadsheet with a PivotTable. A right-click context menu is open over a cell containing the value '633'. The menu path 'Summarize Values By' is highlighted with a red box. To the right, the 'Value Field Settings Dialog' is displayed, also with its title bar 'Value Field Settings' highlighted by a red box. The dialog shows the 'Source Name' as 'Amt' and the 'Custom Name' as 'Sum of Amt'. Under the 'Summarize value field by' section, the 'Sum' option is selected. Other options include Count, Average, Max, Min, and Product. At the bottom right of the dialog are 'OK' and 'Cancel' buttons.

Multiple Subtotals

- Horizontal Layout:
 - The same numeric field can be reused multiple times with different calculations.
 - Right-click on a field → Value Field Settings → Choose function.
 - Remove fields by dragging them out of the Values box.
- Stacked (Vertical) Layout:
 - Display multiple subtotal functions at once.
 - Select a parent item → Field Settings → Subtotals & Filters Tab → Custom → Choose functions (Sum, Count, Average).
 - Remove via the same Field Settings path.

Analyze > Active Field > Field Settings >
Subtotals & Filters Tab > Subtotals > Custom > Select function(s)

Field Settings Dialog

Sum of Sales	Column Labels	Row Labels	Apples	Oranges	Grand Total
Qtr1		Jan	82662	68603	151265
		Feb	53031		53031
		Mar		30620	30620
Qtr1 Sum	135693		99223		234916
Qtr1 Count	6		6		12
Qtr1 Average	22615.5	16537.16667	19576.33333		
Qtr2		Apr	40418		40418
		May		35194	35194
		Jun	57307	30655	87962
Qtr2 Sum	97725		65849		163574
Qtr2 Count	4		4		8
Qtr2 Average	24431.25	16462.25			20446.75

Follow the same path to remove these settings.

Show Values As

- Provides pre-built calculations for comparative analysis.
- Right-click on a numeric value → Show Values As → Choose from:
 - % of Grand Total
 - % of Row/Column Total
 - Difference From
 - Running Total In
 - Rank
 - Index (Weighted Average Calculation)

Right-click on a value > Show Values As

Show Values as: % of Grand Total Results

Row Labels	Sum of Sales	% Grand Total
Central	516,163	26.67%
East	615,828	26.67%
North	672,584	26.67%
South	483,870	20.00%
Grand Total	2,288,445	100.00%

% of Row Totals

Row Labels	Sum of Sales	Column Labels	Apples	Oranges	Bananas	Peaches	Pears	Grand Total
Central			18.25%	12.06%	17.76%	16.82%	35.11%	100.00%
East			18.35%	11.32%	22.96%	19.40%	27.98%	100.00%
North			17.10%	9.97%	22.28%	22.79%	27.86%	100.00%
South			7.91%	10.54%	23.61%	25.54%	32.40%	100.00%
Grand Total			15.75%	10.92%	21.73%	21.11%	30.49%	100.00%

Specific Calculations:

- Parent Row/Column Calculations:

- Right-click on an inner field value → **Show Values As > % of Parent Row Total.**
- **Running Totals:**
 - Displays cumulative values.
 - Right-click → **Show Values As > Running Total In.**
 - The **Grand Total** is displayed in the **final period's cell**, with the Grand Total row empty.
 - Hide the Grand Total label via **Design Tab > Grand Totals > Off.**
- **Index Calculation:**
 - Compares values against totals to calculate a weighted average.
 - Right-click → **Show Values As > Index.**

Original PivotTable

Row Labels	Central	East	North	South
Apples	94206	112980	114984	38276
Bananas	91668	141408	149868	114252
Oranges	62272	69698	67031	50980
Peaches	86799	119457	153296	123588
Pears	181218	172285	187405	156774

Index Results

Row Labels	Central	East	North	South
Apples		1.16	1.16	1.09
Bananas		0.82	1.06	1.03
Oranges		1.10	1.04	0.91
Peaches		0.80	0.92	1.08
Pears		1.15	0.92	0.91
				1.06

- **Rank Calculation:**

- Converts numbers into ranked order (smallest to largest or vice versa).
- Right-click → **Show Values As > Rank.**

Row Labels	Sum of Sales
Apples	4
Bananas	2
Oranges	5
Peaches	3
Pears	1

- **To Disable Custom Calculations:**

- Right-click → **Show Values As > No Calculation.**

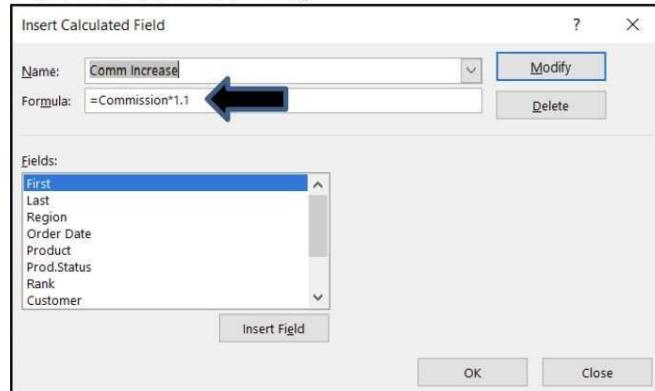
To Create a Calculated Field:

1. Click within the PivotTable data.
2. Analyze > Calculations Command Group > Fields, Items, & Sets button > Calculated Field
3. The Insert Calculated Field dialog appears.
4. Enter a name for the field, such as Comm Increase.
5. Enter the formula for the new field, inserting existing fields from the field list as needed.
6. The new field will appear in the field list. The new column will appear with the calculations in the PivotTable. The new field is part of the PivotTable's list and can be turned on and off.

PivotTable with New Calculated Field

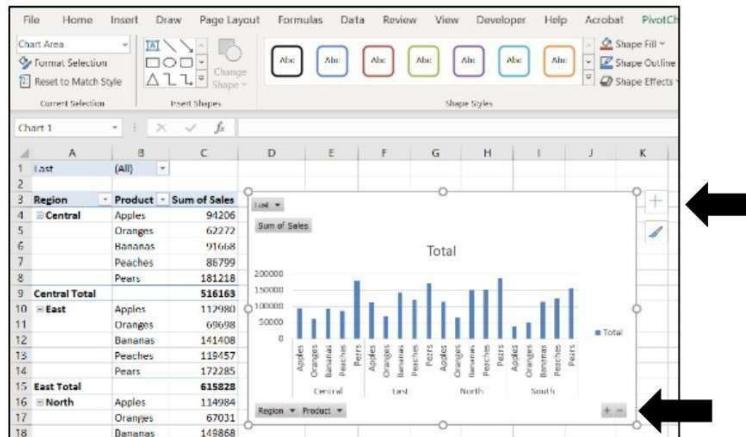
Row Labels	Sum of Commission	Sum of Comm Increase
Central	61939.56	\$68,134
East	73899.36	\$81,289
North	80710.08	\$88,781
South	58064.4	\$63,871
Grand Total	274613.4	\$302,075

Insert Calculated Field Dialog

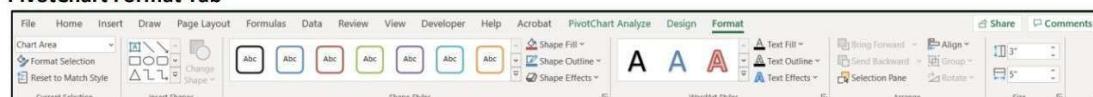


Column data = Series Axis (Y-Axis).

Row data = Category Axis (X-Axis).



PivotChart Format Tab



Creating a PivotChart

- Click anywhere inside the **PivotTable** to start.
- Navigate to **Analyze Tab > Tools Command Group > PivotChart Button**.
- Select a chart type in the **PivotChart** dialog.
- The chart appears on the same sheet as the **PivotTable** and updates dynamically.

PivotChart Features

- Filter buttons** allow filtering at the chart level.
- Changing the PivotTable** affects the PivotChart and vice versa.
- Drill-down buttons** (+/-) help expand or collapse hierarchical data.
- Row Data (X-Axis)** – Category Axis.
- Column Data (Y-Axis)** – Series Axis.

PivotChart Customization

- PivotChart Format Tab** contains:
 - PivotChart Analyze, Design, and Format** ribbons.
 - Additional commands available from **Insert Tab**.
- Chart Elements and Styles buttons** provide more formatting options.
- Dragging fields into the **Column Area (Legend)** redraws the chart with new data.

Managing PivotCharts

- **Delete a chart** – Click the border and press **Delete**.
- **Move a chart** – Click the border or white space → Drag when **four-way arrow** appears.
- **Resize a chart** – Click a **corner border**, hold **Shift**, and drag.
- **Print a chart**
 - Select the chart → **Print** (for chart only).
 - Click outside the chart → **Print** (for both PivotTable and chart).

Instant Chart Creation

- Press **F11** → Creates a default chart on a new sheet.
- Press **Alt+F1** → Creates a chart on the same sheet as the PivotTable.

Sorting and Filtering Data in PivotCharts

- Click a **filter button** → Use **Sort command**.
- Sorting in the PivotTable affects the chart order.
- Use **manual sorting** if automatic sorting is insufficient.
- Filters for **labels, numbers, and dates** work the same as in PivotTables.

Changing Chart Types

- Readability is key when selecting a chart type:
 - **Column Charts (Vertical)** – Compare series across categories.
 - **Bar Charts (Horizontal)** – Ideal for lengthy labels.
 - **Line Charts** – Useful for spotting trends.
 - **Pie Charts** – Show parts of a whole (use only totals and labels).
- To change a chart type:
Design Tab > Type Command Group > Change Chart Type
- Not all conversions work well, but changes are **undoable**.

Customizing Pie Charts

- To add **percentages** to a pie chart:
Chart Elements button > Data Labels > More Data Label Options > Percentage
- To add a **chart title**:
Chart Elements button > Chart Title > Above Chart

Setting a Default Chart Type

- In **All Chart Types** or **Change Chart Types** dialogs:
 1. Right-click the desired chart type.
 2. Click **Set as Default Chart**.
- This will be used when pressing **F11** or **Alt+F1**.

Saving & Reusing Chart Templates

- Save a customized chart for future use:
 1. Right-click the chart → **Save as Template**.
 2. Name it and save (default path: **Microsoft > Templates > Charts**).
- Apply a saved template:
 1. Select a chart to modify.
 2. **Design Tab > Change Chart Type** OR right-click the chart → **Change Chart Type**.
 3. Open **Templates Folder** and choose the saved format.

Power query

1. Load Data into Power Query

- **Where to find it:**
 - Select the data range.
 - Go to **Data → Get & Transform Data → From Table/Range**.
 - If prompted, confirm the selection by clicking **OK**.
-

2. Remove Duplicates

- **Where to find it:**
 - In **Power Query**, select the relevant column(s).
 - Go to **Home → Remove Duplicates**.
-

3. Handle Missing Data

Replace Missing Values

- **Where to find it:**
 - Select the column.
 - Go to **Transform → Replace Values**.
 - Enter the value to replace nulls (e.g., replace blanks with "N/A" or 0).

Remove Rows with Null Values

- **Where to find it:**
 - Select the column.
 - Go to **Home** → **Remove Rows** → **Remove Blank Rows**.
-

4. Split Columns (For Delimited Data)

- **Where to find it:**
 - Select the column (e.g., Full Name).
 - Go to **Transform** → **Split Column** → **By Delimiter**.
 - Choose a delimiter (e.g., space for splitting "John Doe" into First and Last Name).
-

5. Trim and Clean Text

Trim Extra Spaces

- **Where to find it:**
 - Select the column.
 - Go to **Transform** → **Format** → **Trim**.

Remove Non-Printable Characters

- **Where to find it:**
 - Select the column.
 - Go to **Transform** → **Format** → **Clean**.
-

6. Change Data Types

- **Where to find it:**
 - Select the column.
 - Go to **Transform** → **Data Type** (Choose: Text, Date, Number, etc.).
-

7. Unpivot Data (For Normalization)

- **Where to find it:**
 - Select the columns that need to be transformed into row format.
 - Go to **Transform** → **Unpivot Columns**.
-

8. Group and Aggregate Data

- **Where to find it:**
 - Select the column to group by.
 - Go to **Transform** → **Group By**.
 - Choose an aggregation function like **Sum**, **Count**, or **Average**.
-

9. Load Cleaned Data Back to Excel

- **Where to find it:**
 - Once done, click **Close & Load** → Choose **Load to Worksheet/Table**.
-

Date Modification with Power Query

Age Calculation

- Determines the number of days between a specified date and the current date.
- Replaces the current date with a second date column if needed.

Steps:

1. Select the date column.
 2. Navigate to **Add Column** → **Date** → **Age**.
 3. Transforms date into the Days.Hours:Minutes:Seconds format.
-

Date Only

- Removes the time portion from the date column.

Steps:

1. Select the date column.
 2. Navigate to **Add Column** → **Date** → **Date Only**.
 3. Transforms date into Date format (e.g., 12-01-2015).
-

Year Extraction

- Extracts only the year from the date column.

Steps:

1. Select the date column.
 2. Navigate to **Add Column → Date → Year**.
 3. Transforms date into Year format (e.g., 2015).
-

Month Extraction

- Extracts only the month from the date column.
- Available options: start of the month, end of the month, days in the month, and month name.

Steps:

1. Select the date column.
 2. Navigate to **Add Column → Date → Month** (e.g., Days in a Month).
 3. Displays the number of days in a month (e.g., 31/30).
-

Quarter Extraction

- Extracts the quarter from the date column.
- Options: quarter of the year, start of the quarter, and end of the quarter.

Steps:

1. Select the date column.
 2. Navigate to **Add Column → Date → Quarter**.
 3. Displays the quarter (e.g., Q1).
-

Week Extraction

- Extracts only the week from the date column.
- Options: week of the year, week of the month, start of the week, and end of the week.

Steps:

1. Select the date column.
 2. Navigate to **Add Column → Date → Week**.
 3. Displays the week of the year (e.g., Week 3).
-

Day Extraction

- Extracts only the day from the date column.
- Options: day, day of the week, day of the year, start of the day, end of the day, and name of the day.

Steps:

1. Select the date column.
 2. Navigate to **Add Column → Date → Day** (e.g., Name of Day).
 3. Displays the day name (e.g., Monday).
-

Parse Date

- Converts a text-formatted date back into a proper date format.

Steps:

1. Convert the date to text.
 2. Navigate to **Add Column → Date → Parse**.
 3. Transforms text back into date format (e.g., 12-01-2015).
-

Subtract Days

- Subtracts one date column from another.

Steps:

1. Select two date columns.
 2. Navigate to **Add Column → Date → Subtract Days**.
 3. Displays the difference in days (e.g., -169).
-

Combine Date and Time

- Merges date and time columns into one.

Steps:

1. Select date and time columns.
 2. Navigate to **Add Column → Date → Combine Date and Time**.
 3. Displays merged result (e.g., 12-01-2015 12:02:12).
-

Earliest and Latest Days

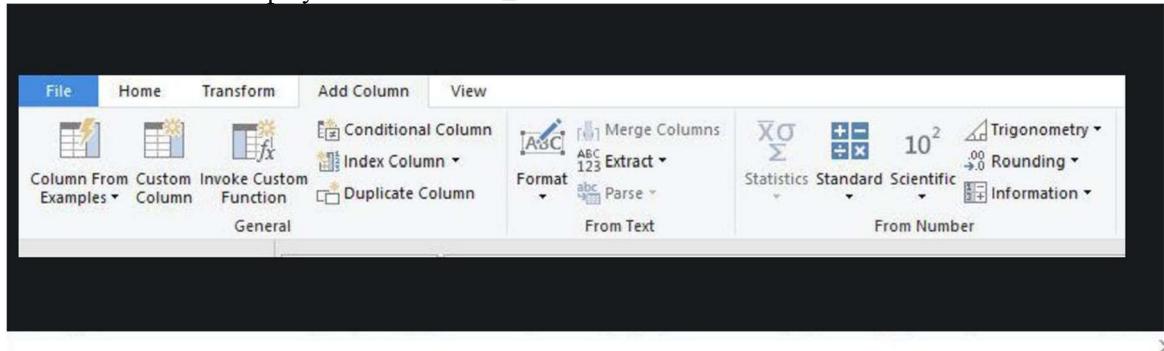
- Finds the earliest or latest date value.
- Only applicable for **Transform**, not **New Column**.

Creating a Conditional Column

- Creates a new column based on conditions.

Steps:

1. Navigate to **Add Column → Conditional Column**.
2. Enter the **New Column Name** (e.g., Pass/Fail).
3. Set conditions:
 - o If Marks > 40, output **Pass**.
 - o If Marks < 40, output **Fail**.
4. Click **OK** to generate the new column.
5. The column displays **Pass** for marks ≥ 40 and **Fail** for marks < 40 .



Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name

Pass/Fail

	Column Name	Operator	Value	Output
If	Marks	is greater than	ABC 123 40	Then ABC 123 Pass
Else If	Marks	is less than	ABC 123 40	Then ABC 123 Fail

Add Clause

Else

ABC

123

OK

Cancel

= Table.AddColumn(#"Changed Type", "Pass/Fail", each if [Marks] > 40 then "Pass" else if [Marks] < 40 then "Fail")				
	1 ² 3 Roll No.	A ^B C Student name	1 ² 3 Marks	ABC 123 Pass/Fail
1	101	Geeta	60	Pass
2	102	Rajni	78	Pass
3	103	Deepak	21	Fail
4	104	Sonu	75	Pass
5	105	Priya	16	Fail
6	106	Ashu	62	Pass
7	107	Muskan	55	Pass
8	108	Pawan	30	Fail
9	109	Monu	31	Fail
10	110	Mandeep	66	Pass

How do you remove outliers from a dataset?

Answer: Use statistical measures like the interquartile range (IQR).

Example: Values 10 50 100 500

Find Q1, Q3, and IQR: =

QUARTILE(A1:A4, 1) → 30.

$\text{QUARTILE(A1:A4, 3) } \rightarrow 125.$
Outlier threshold: $Q3 + 1.5 * \text{IQR} \rightarrow 325.$

The screenshot shows a Microsoft Excel spreadsheet with data in columns A and B. Row 1 contains headers 'Name' and 'Gender'. Rows 2 through 6 contain data: Vanshika (Female), Rahul (Male), Anushka (Female), Varun (Male), and Riva (Female). A dropdown menu is open over cell B2, showing the value 'Female'. Below the menu, the 'Data Validation' dialog box is displayed. The 'Settings' tab is selected. Under 'Validation criteria', the 'Allow' dropdown is set to 'List', and the source is '\$D\$2:\$D\$3'. The 'Ignore blank' and 'In-cell dropdown' checkboxes are checked. At the bottom of the dialog, there is a checkbox for 'Apply these changes to all other cells with the same settings' and three buttons: 'Clear All', 'OK', and 'Cancel'.

Shortcut to Remove Blank Rows Using "Go To Special"

1. Select your data (or press **Ctrl + A** to select all).
2. Press **Ctrl + G**, then click "Special...".
3. Select "Blanks" and click **OK**.
4. Press **Ctrl + -** (Control + Minus) to open the **Delete** dialog.
5. Choose "Entire Row" and press **Enter**.