

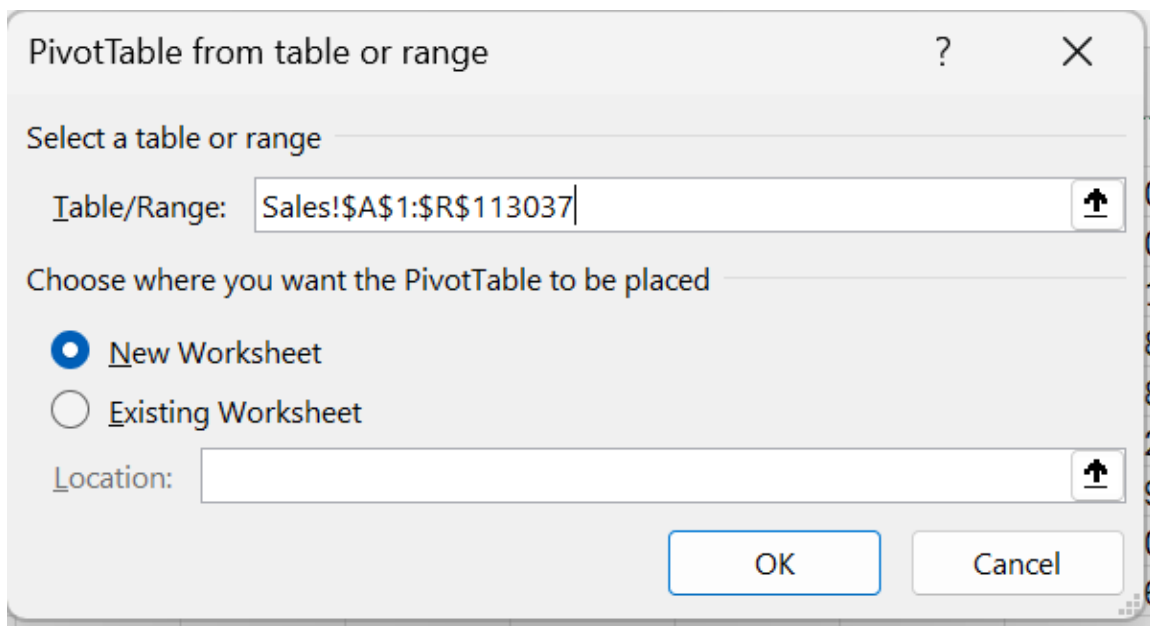
Data Analyst BootCamp

Excel

[DataSet Link](#)

Pivot Table

Insert→ Pivot Table



how to select entire data?? ctrl+shift+ rightarrow/downarrow

Row Labels

Row Labels	Revenue per country	Cost per country	Profit per country
Australia	\$ 2,13,02,059	\$ 1,45,26,029	\$ 67,76,030
Canada	\$ 79,35,738	\$ 42,18,442	\$ 37,17,296
France	\$ 84,32,872	\$ 55,52,590	\$ 28,80,282
Germany	\$ 89,78,596	\$ 56,18,601	\$ 33,59,995
United Kingdom	\$ 1,06,46,196	\$ 62,32,343	\$ 44,13,853
United States	\$ 2,79,75,547	\$ 1,69,01,903	\$ 1,10,73,644
Grand Total	\$ 8,52,71,008	\$ 5,30,49,908	\$ 3,22,21,100

PivotTable Fields

Choose fields to add to report:

Search

☐ Date
☐ Day
☐ Month
☐ Year

Drag fields between areas below:

Filters

Columns

Rows

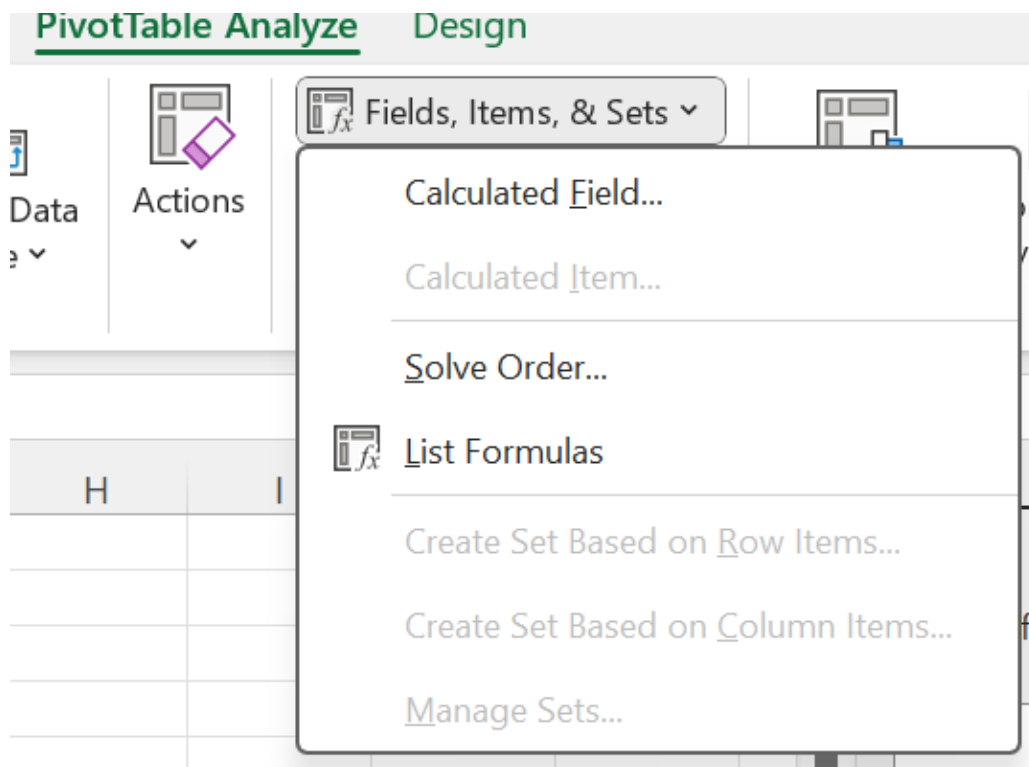
Country

Values

Revenue per co...

Defer Layout Update Update

To create a new column of your own formula create like this



calculate field click on it

Insert Calculated Field

Name: ▼ Add

Formula: Delete

Fields:

- Date
- Day
- Month
- Year
- Customer_Age
- Age_Group
- Customer_Gender
- Country

Insert Field

OK Close

and do as per your need

Formulas in Excel:

max and min

=max(range)

=min(range)

=if(condn,what to do when condn true,otherwise

eg:=IF(D2:D10>30,"old","young")

=ifs

eg:=IFS(F2:F10="Salesman","Sales",F2:F10="HR","fire",F2:F10="Receptionist","SayHi")

=LEN(C2:C10)

=LEFT(B2:B10,3)

=RIGHT(H2:H10,4) #used to extract year in dd/mm/yyyy format

date to text super helpful tool

=TEXT(H2:H10,"dd/mm/yyyy")

copy and paste this data separately and then do =right(H2:H10,4) to get the year

Trim=It just removes unwanted spaces on both sides

=trim(range)

=CONCAT(firstname," ",lastname)

=SUBSTITUTE(H2:H10,"-","/")

=SUBSTITUTE(H2:H10,"/","-",1) 1st instance

=SUBSTITUTE(H2:H10,"/","-",2) 2nd instance

=sum(range)

=SUMIF(G2:G10,">50000")

=SUMIFS(G2:G10(salary range),E2:E10(gender),"Female",D2:D10(age),">30")

=count(range)

=countif(G2:G10,">50000")

=COUNTIFS(A2:A10,">1006",D2:D10,">29")

=DAYS(end date,start date)

=NETWORKDAYS(startdate,enddate) #removes standard holidays

Xlookup

=XLOOKUP(A3,H2:H10,P2:P10)

=XLOOKUP(Value,in what array it should search, what should it do when it finds that array)

all rows are in same manner we cant modify that columns

=XLOOKUP(A3,H2:H10,O2:O10,"Notfound")

=XLOOKUP("*"&A4,H2:H10,O2:O10,"Notfound",2)

=XLOOKUP(A5&"*",H2:H10,O2:O10,"Notfound",2)

2 stands for wild card match, 0 stands for exact match and * represents firstname or lastname where are we putting

=XLOOKUP(A3,N2:N10,I2:I10,,1)

0→exact match

1→exact match or next smaller item

-1→exact match or next larger item

2→wildcard character match

=XLOOKUP(A4,N2:N10,I2:I10,,,-1)

1→search from first to last

-1→search from last to first

=XLOOKUP(I1,H1:S1,H2:S2)

I1 is february

H1:S1 is horizontal months range

H2:S2 is horizontal look of paper

=SUM(XLOOKUP(I1,H1:S1,H2:S2):XLOOKUP(J1,H1:S1,H2:S2)) SUM OF FEB AND MARCH MONTH PAPER

=VLOOKUP(A3,H1:P10,9,FALSE)

in vlookup we are looking for value in table of array and which column you want u will mention that column number in that table array here it is 9 and false for exact match

Conditional Formatting

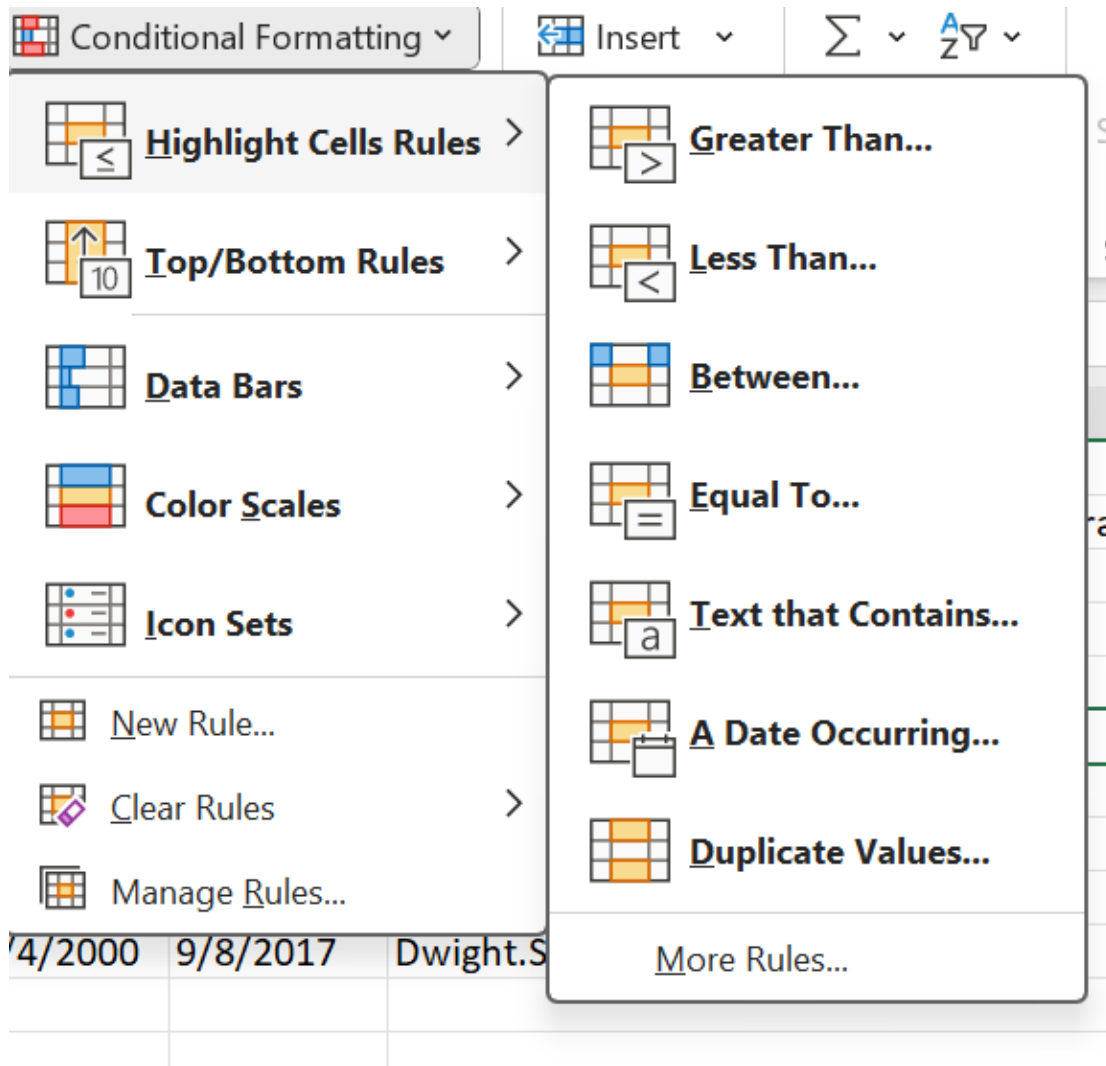
Conditional formatting can help make patterns and trends in your data more apparent. To use it, you create rules that determine the format of cells based on their values, such as the following monthly temperature data with cell colors tied to cell values.

useful tools

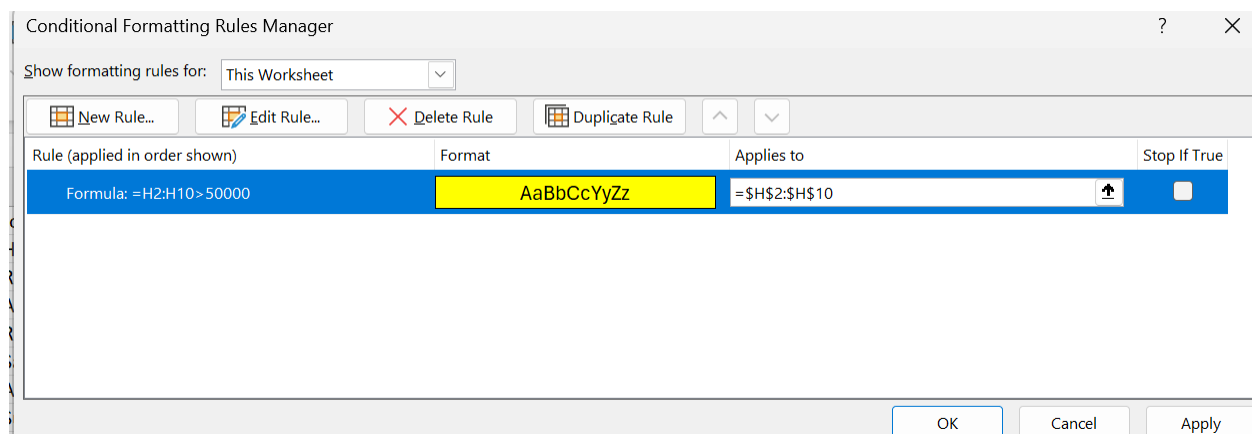
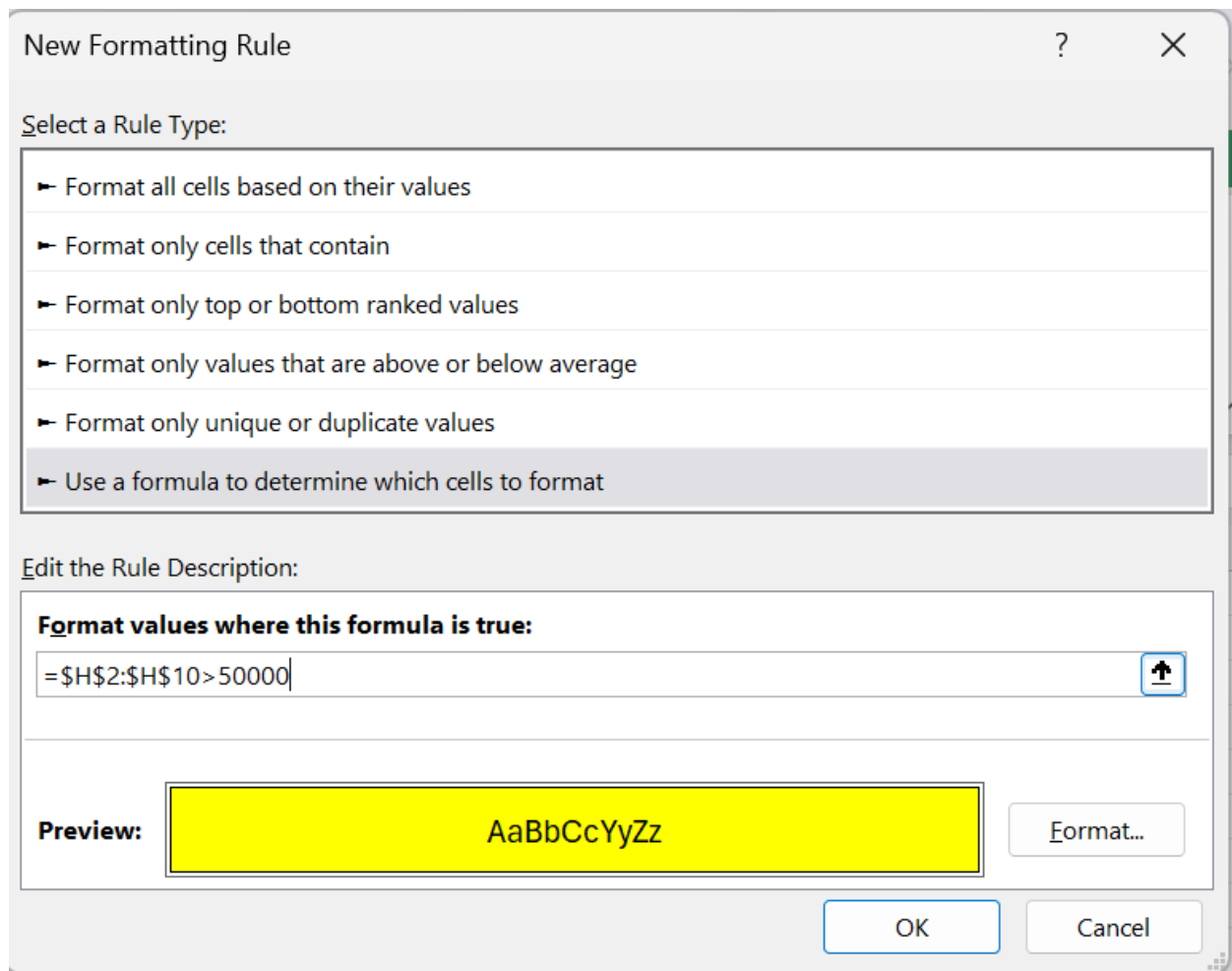
select top row and add filter to it and add conditional formatting

most used conditional formatting are: text that contains and duplicates

and click on filter and sort by using color

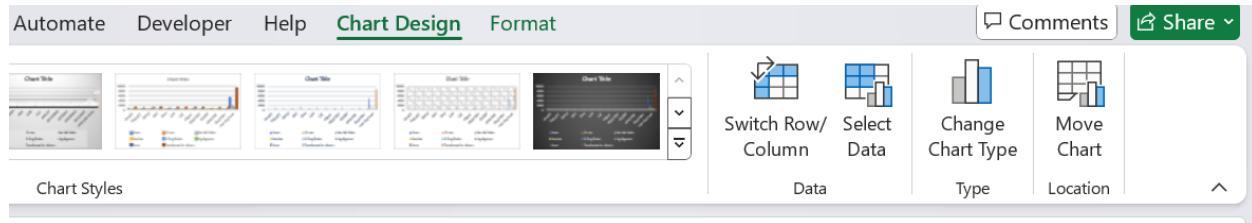


New rule writing



make sure both range matches then only it works

Charts in Excel:



some imp things:

switch row/col

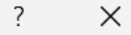
select data

change chart type

move chart

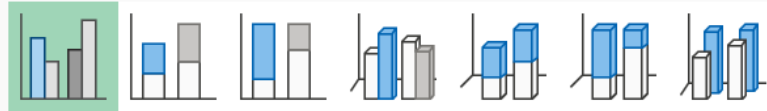
click on change chart type:

Change Chart Type

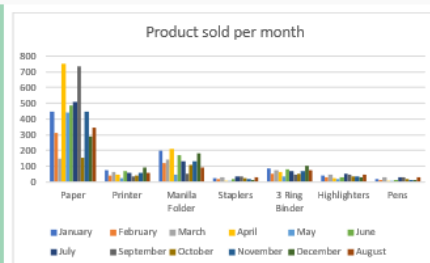


Recommended Charts **All Charts**

- Recent
- Templates
- Column**
- Line
- Pie
- Bar
- Area
- X Y (Scatter)
- Map
- Stock
- Surface
- Radar
- Treemap
- Sunburst
- Histogram
- Box & Whisker
- Waterfall
- Funnel
- Combo



Clustered Column

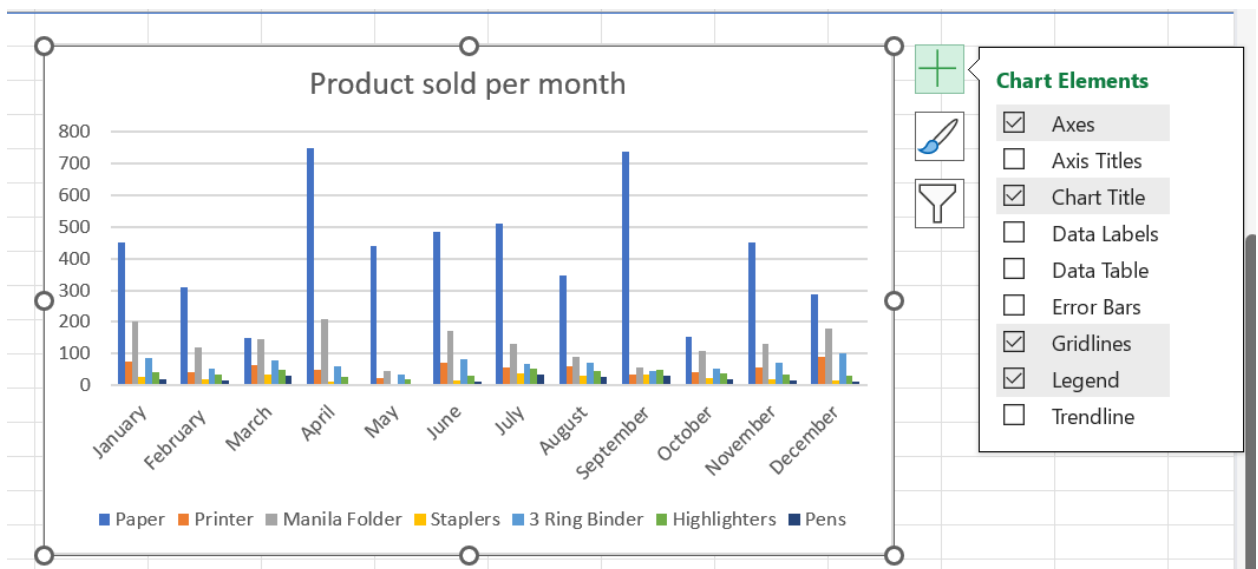


OK

Cancel



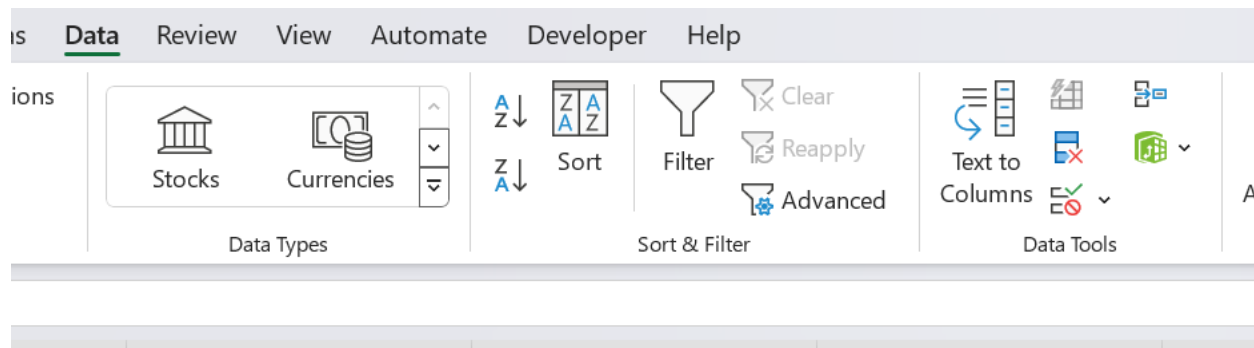
Choose color accordingly



choose labels accordingly

Cleaning data in excel:

1.Remove duplicates



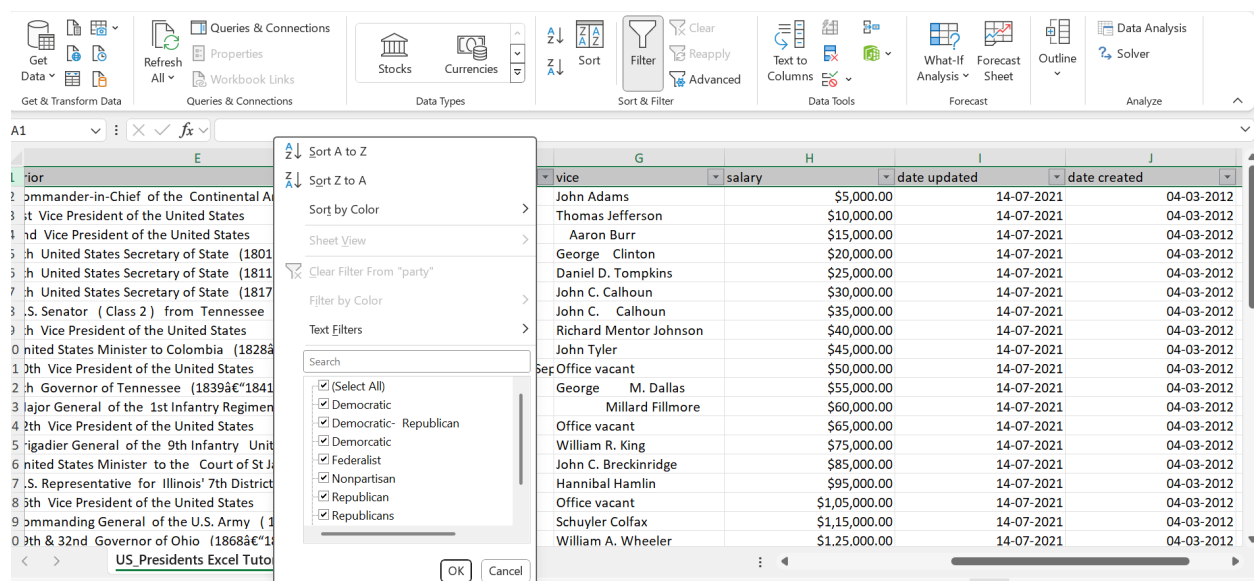
select entire data go to data and in data tools remove duplicates it will remove

2. Create a new column

=PROPER(C2) for proper naming convention and plus sign double click to apply for all(shortcut)

proper is proper case like lower and upper case

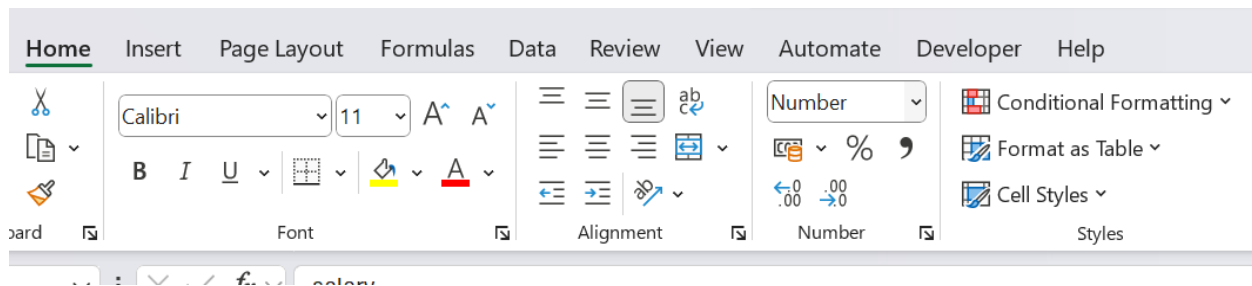
3.



select two spelling mistakes and update mistake one which is more biased

4.=trim(name)to look good without spaces

5. with dollar sign u get problem in sql so get rid of that by selecting column and in home select it as number instead of a dollar sign



6.date issue

select date columns and select short date

Conditional Formatting

Insert

Delete

Format

Cells

Editing

Sensitivity

Add-in

	J	K
	date updated	date created
5000	14-07-2021	04-03-2012
10000	14-07-2021	04-03-2012
15000	14-07-2021	04-03-2012
20000	14-07-2021	04-03-2012
25000	14-07-2021	04-03-2012
30000	14-07-2021	04-03-2012
35000	14-07-2021	04-03-2012
40000	14-07-2021	04-03-2012
45000	14-07-2021	04-03-2012
50000	14-07-2021	04-03-2012
55000	14-07-2021	04-03-2012
60000	14-07-2021	04-03-2012
65000	14-07-2021	04-03-2012
75000	14-07-2021	04-03-2012
85000	14-07-2021	04-03-2012
95000	14-07-2021	04-03-2012
105000	14-07-2021	04-03-2012
115000	14-07-2021	04-03-2012

General

No specific format

Number

Currency

Accounting

Short Date

Long Date

Time

Percentage

Fraction

More Number Formats...

7. once any new column with feature edited now copy that col and paste to earlier col as paste as value and proceed to not get any error and delete the new col which created recently because u had already values pasted in existing col

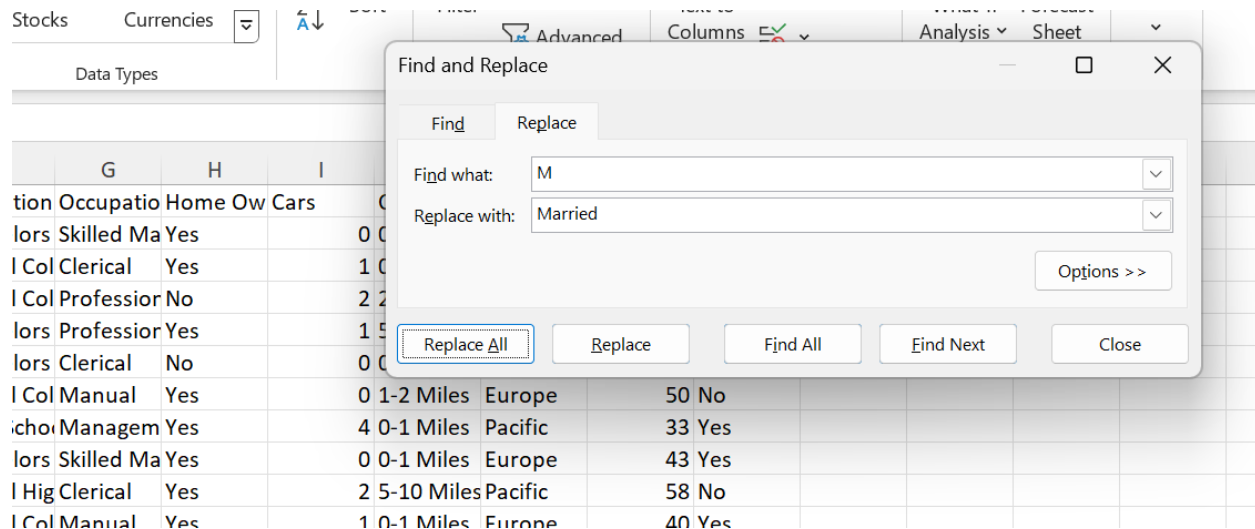
8.delete all unnecessary cols

Excel Project

create new working worksheet,dashboard,pivot table sheets under excel file

1.select data and go to data tab and remove duplicates

2.ctrl+H on col



3. similarly do for all cols which req this

4. for income col add dollar sign infront of that

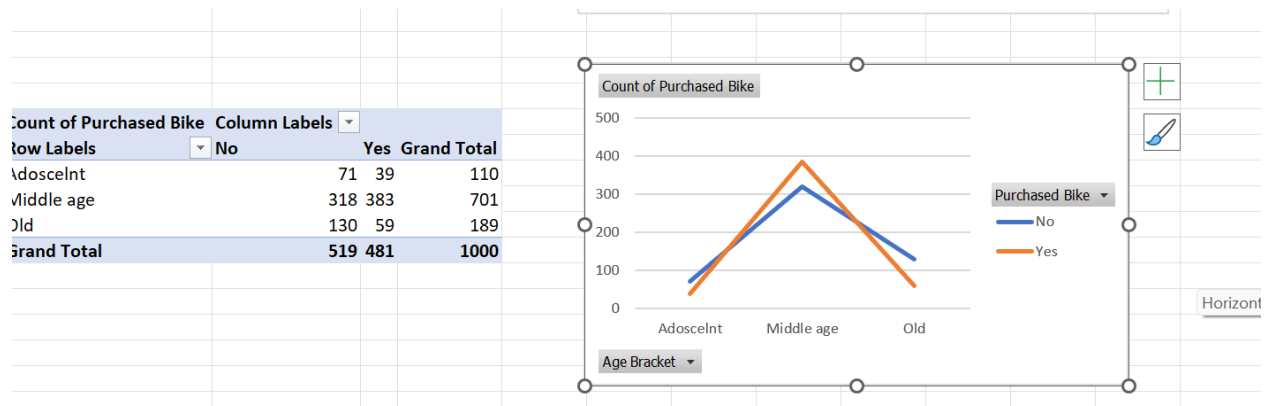
5. Create a filter for top row and check each and every col if any thing unusual change it

6. now add age bracket as age has so many values

=IF(L2>=55,"Old",IF(L2>=31,"Middle age",IF(L2<31,"AdosceInt","invalid")))

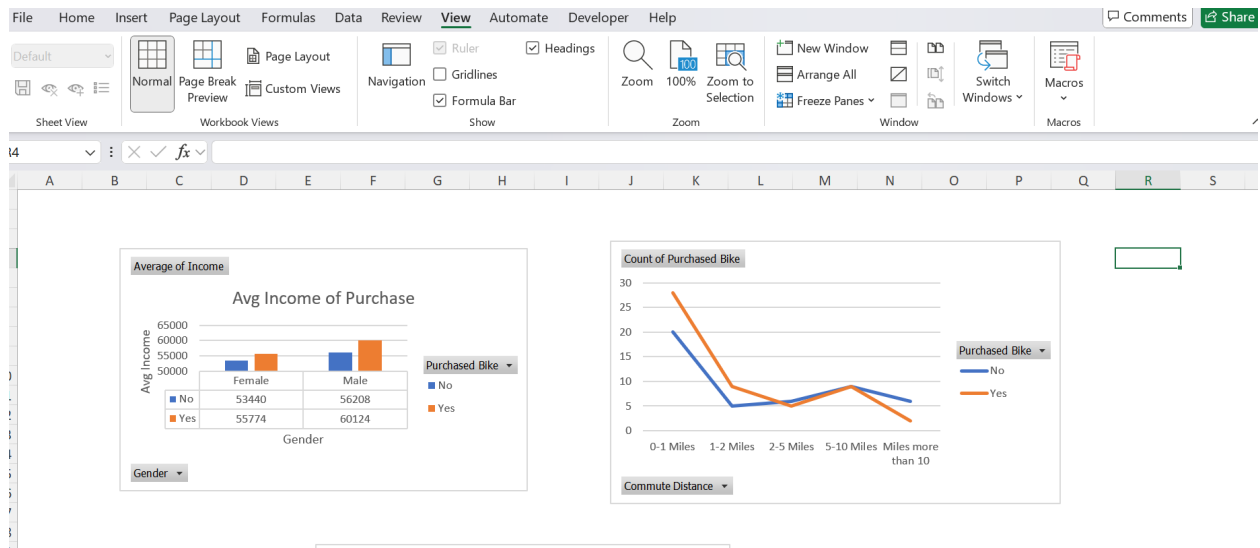
Count of Purchased Bike Column Labels			
Row Labels	No	Yes	Grand Total
0-1 Miles	20	28	48
10+ Miles	6	2	8
1-2 Miles	5	9	14
2-5 Miles	6	5	11
5-10 Miles	9	9	18
Grand Total	46	53	99

here asc/dsc order not working so go to original sheet and modify 10+ miles replace by Miles more than 10 ctrlh to do that



age brackets will help correct gaining of information

In dashboard tab copy all the graphs from pivot tab and paste from view gridlines uncheck



select on chart and click on slicer and make sure to report connections with whatever pivot table u want and do accordingly