USC Marshall

School of Business

Full-Time MBA Program

MBA

Welcome Class of 2025 Orientation

July 20 & 21, 2023

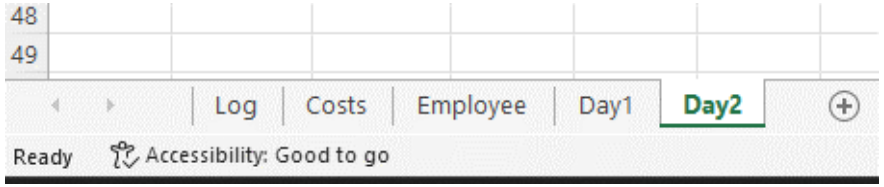
Professor: Dr. Gourab Mukherjee

- Associate Professor in the Data Sciences & Operations (DSO) Department
- PhD in Statistics, Stanford University, 2013
- Expert in
  - algorithmic development, machine learning
  - pricing and monetization policies in digital marketing
- Will teach GSBA524: Data Science for Business in term 2

- How do we monetize Digital goods?
- Innovation in durable products?
- Consulting: Fintech, Online retail

Outline

- We will learn several excel functions by answering the questions based on the dataset **datalog.xlsx**
- **Download data file from**: uscstats.github.io
- Open the file in excel
- Check that the file has 5 worksheets (look at the bottom of file):
  - Log: which is the main data.
  - Costs: which has all different costs other than material costs
  - Employee: which shows the number of employees in the factory per month
  - Day1: set of questions we will do on day 1 of the bootcamp
  - Day2: set of questions we will do on day 2 of the bootcamp.



- About the main worksheet “Log”. The columns in this worksheet are:
  - Date, factory number,
  - Sales (in units),
  - Production (in units),
  - Raw Material Cost (in local currency),
  - Forex rate ($/local currency).

Excel Basics: Cell number and Formula Bar

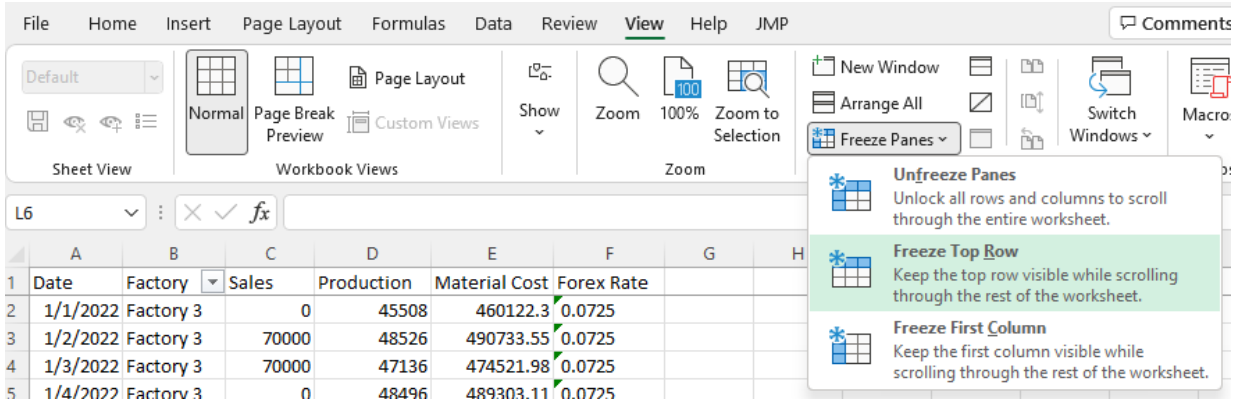
Question 1

Report the sample size (n). Sample size is the number of rows in the file except the header row.

Learn how to go down to the last row to get and how to freeze the header row.

sample size = number of rows - 1

Freeze header row:
Go to View and
click on FreezePanes and
select Freeze top row



Some Shortcuts…

Ctrl + A	Select a whole data set
Ctrl + down arrow (try it with right, left, and up arrows)	Move to the last row in the same column
Ctrl + Shift + down arrow (try it with right, left, and up arrows)	Select data in the same column

Mac: Use Command instead of Ctrl

Question 2

Q2) Report the following summary statistics for the Production column. Round to 2 decimal places

Min	Max	Average	Median	Range	Standard Deviation
=MIN(Log!D:D)					

- Every formula begins with an equality
- Every function has a name
- After the function name comes parenthesis where the inputs of the functions are provided
- Range=max – min
- Other function names here: min, max, average, median, stdev.s

Question 3: Exercise.

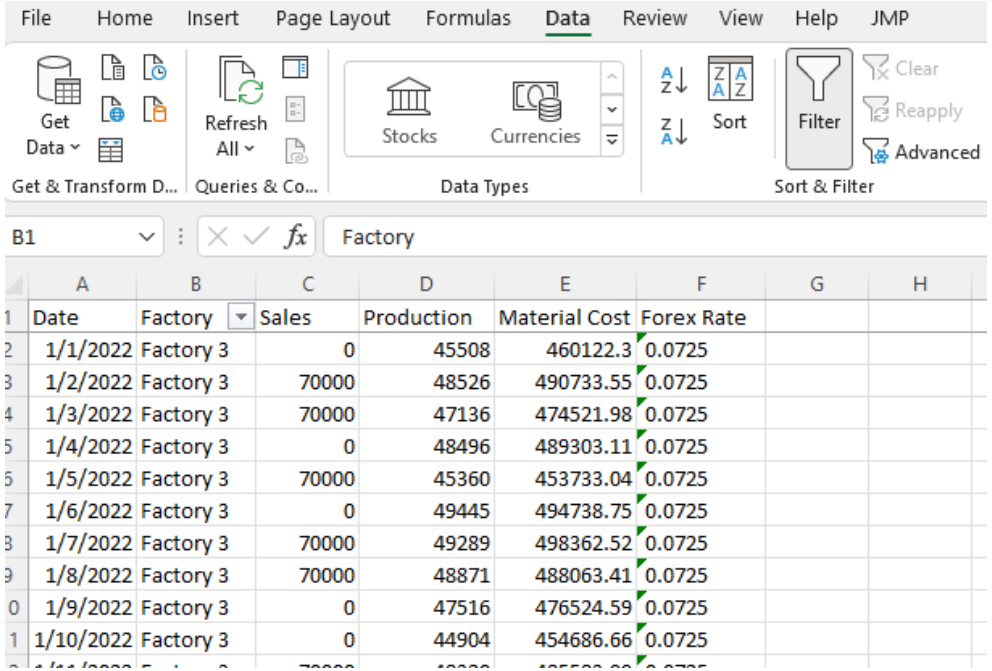
Cut, Copy, Paste

- Copy: Control C
- Cut: Control X
- Paste: Control V
  - Paste as formula
  - Paste as values
  - Paste Special
- In Mac: command instead of control

Question 4

Q4) Filter out and keep only records for factory 1. Put those in a new worksheet. Name the worksheet: "Factory 1 only". Report the average and standard deviation of the production column for this data. Round to two decimal places.

- Go to Log worksheet
- Click on top of column B
- Go to data, click on the filter icon in data
- This will put a button in the first row of column B
- Click on the button and choose factory 1
- Copy all the data and paste it in a new worksheet
- Name the worksheet "Factory 1 only"

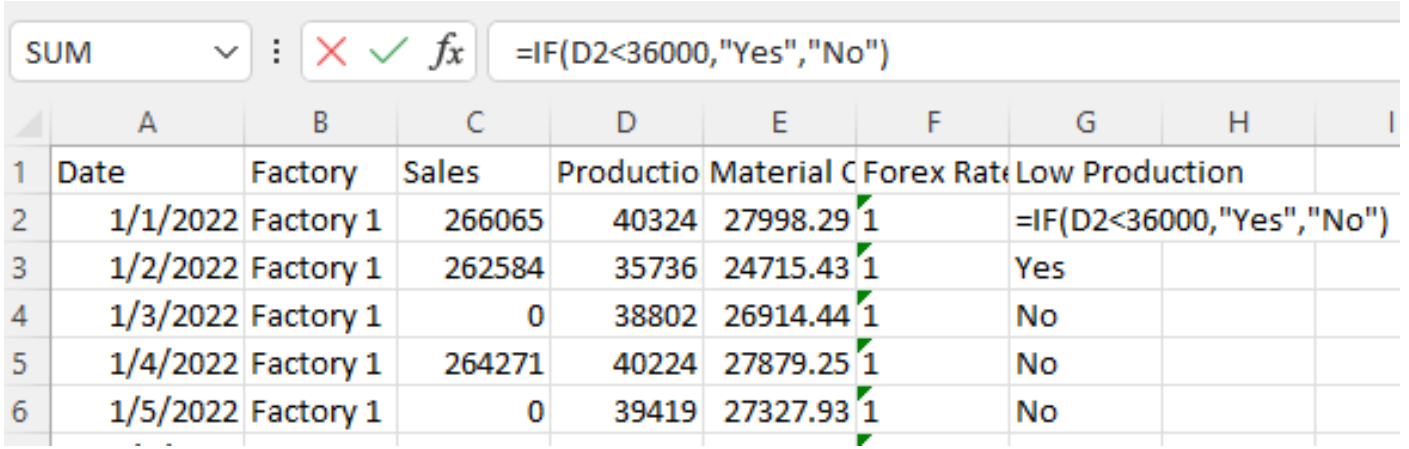


	A	B	C	D	E	F	G	H
1	Date	Factory	Sales	Production	Material Cost	Forex Rate		
2	1/1/2022	Factory 3	0	45508	460122.3	0.0725		
3	1/2/2022	Factory 3	70000	48526	490733.55	0.0725		
4	1/3/2022	Factory 3	70000	47136	474521.98	0.0725		
5	1/4/2022	Factory 3	0	48496	489303.11	0.0725		
6	1/5/2022	Factory 3	70000	45360	453733.04	0.0725		
7	1/6/2022	Factory 3	0	49445	494738.75	0.0725		
8	1/7/2022	Factory 3	70000	49289	498362.52	0.0725		
9	1/8/2022	Factory 3	70000	48871	488063.41	0.0725		
10	1/9/2022	Factory 3	0	47516	476524.59	0.0725		
11	1/10/2022	Factory 3	0	44904	454686.66	0.0725		

Question 5

Q5) Add a new column to the filtered data. Call it "Low Production". If Production was less than 36000 units, put "yes" else "no" in that column.

- Go to "Factory 1 only" worksheet
- Type the column name in cell G1
- We'll use formula IF
- This is a "cell formula"
- Do it for cell G2 first
- Autocomplete it for the remaining cells in the column by dragging or clicking the plus sign at the bottom right corner of G2 cell.



	A	B	C	D	E	F	G	H	I
1	Date	Factory	Sales	Productio	Material C	Forex Rate	Low Production		
2	1/1/2022	Factory 1	266065	40324	27998.29	1	=IF(D2<36000,"Yes","No")		
3	1/2/2022	Factory 1	262584	35736	24715.43	1	Yes		
4	1/3/2022	Factory 1	0	38802	26914.44	1	No		
5	1/4/2022	Factory 1	264271	40224	27879.25	1	No		
6	1/5/2022	Factory 1	0	39419	27327.93	1	No		

=IF(logical statement, do this if true, do this if false)

Question 6

Q6) Find the percentage of rows with positive sales in the filtered data?

=COUNTIF(Log!C:C,">0")

- First do the count of rows and then do proportion
- We use the COUNTIF function
- Two inputs are needed
  - Range of numbers
  - Criterion; remember to put criterion in " "

Question 7

Q7) Find the number of rows with positive sales and low production in the filtered data?

- We use the COUNTIFS function
- Multiple inputs can be put in. The sequence is:
  - Range of numbers
  - Criterion; remember to put criterion in " "

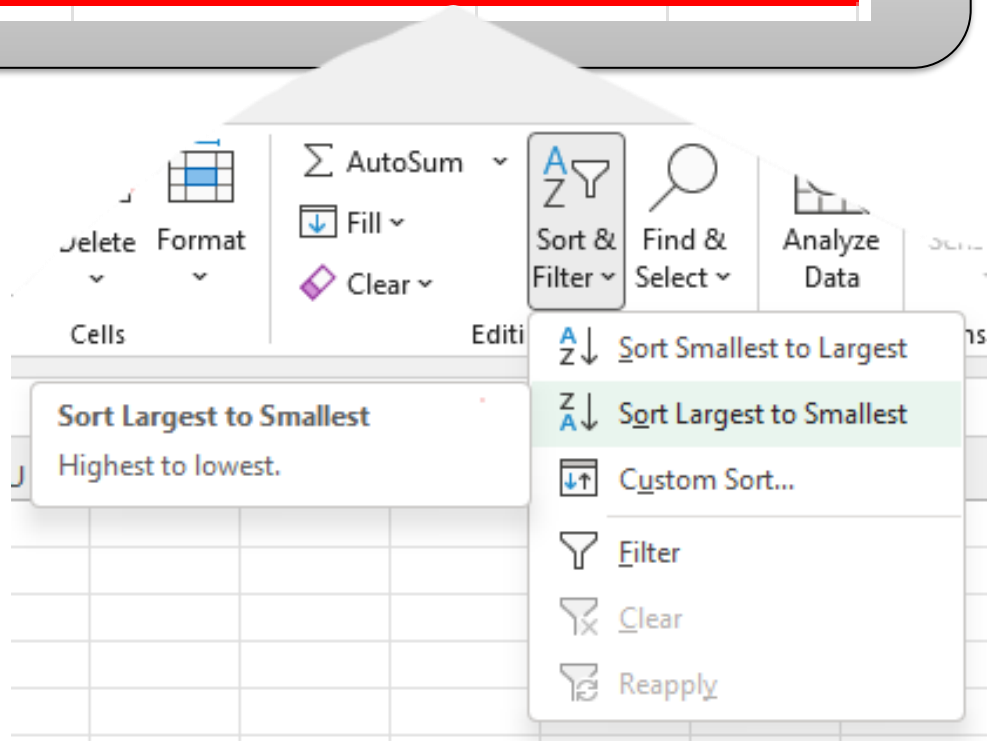
=COUNTIFS('Factory 1 only'!C:C,">0",'Factory 1 only'!D:D,"<36000")

Question 8, 9: Exercise.

Question 10

Q10) Sort the rows of the filtered data from highest to lowest production. Which month dominates the top 5 production days?

- Click on top of column D: Production and select it
- Go to sort function and click descending order
- Choose expand the selection option

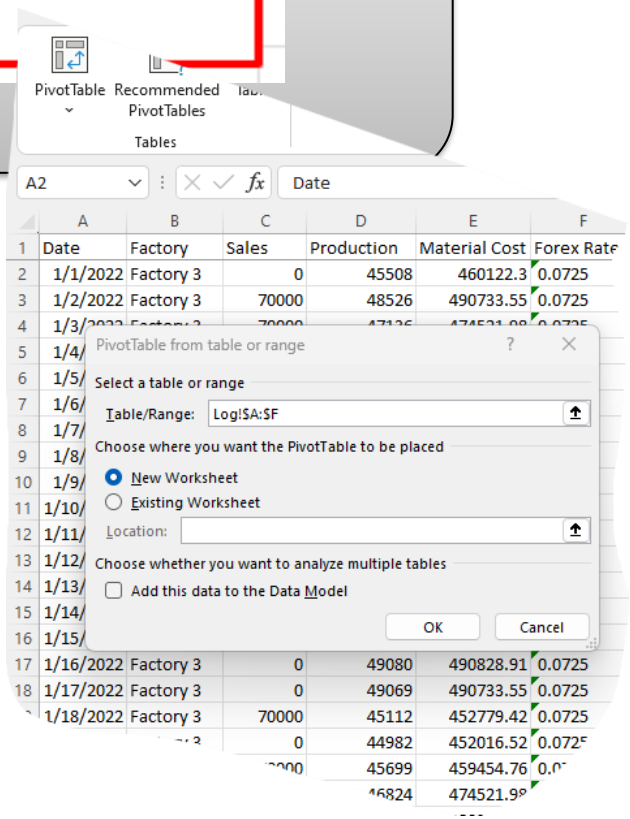


Question 12

Q12) On the complete data report the following KPI: "Total half-yearly production by factory". Paste the answers in the answer box such that it will stay even if you delete the pivot table worksheet.

We will use Pivot table –perhaps the greatest feature of excel.

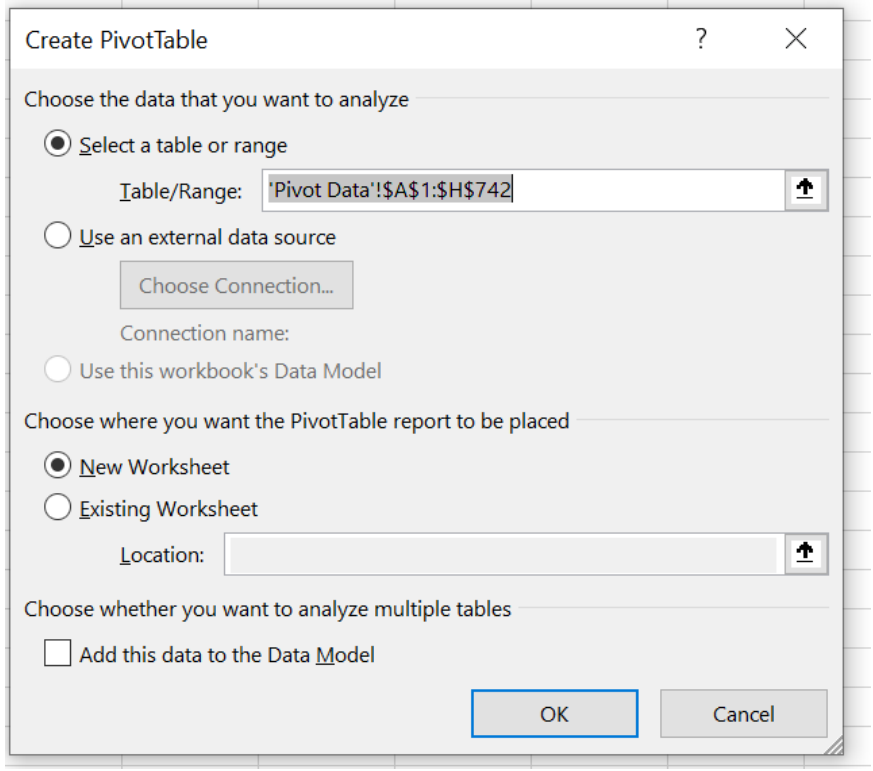
- Go to Log file
- Select all columns by pressing control+A
- Then go to the insert tab and select pivot table
- Make it in a new worksheet and name it accordingly



	A	B	C	D	E	F
1	Date	Factory	Sales	Production	Material Cost	Forex Rate
2	1/1/2022	Factory 3	0	45508	460122.3	0.0725
3	1/2/2022	Factory 3	70000	48526	490733.55	0.0725
17	1/16/2022	Factory 3	0	49080	490828.91	0.0725
18	1/17/2022	Factory 3	0	49069	490733.55	0.0725
	1/18/2022	Factory 3	70000	45112	452779.42	0.0725
			0	44982	452016.52	0.0725
				45699	459454.76	
					474521.98	

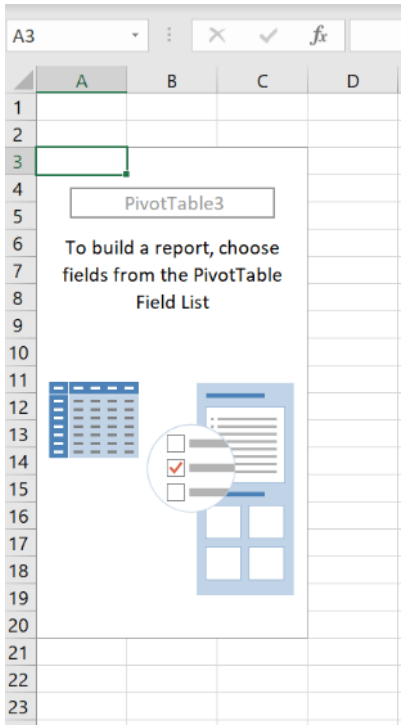
Pivot Table Dialog Box

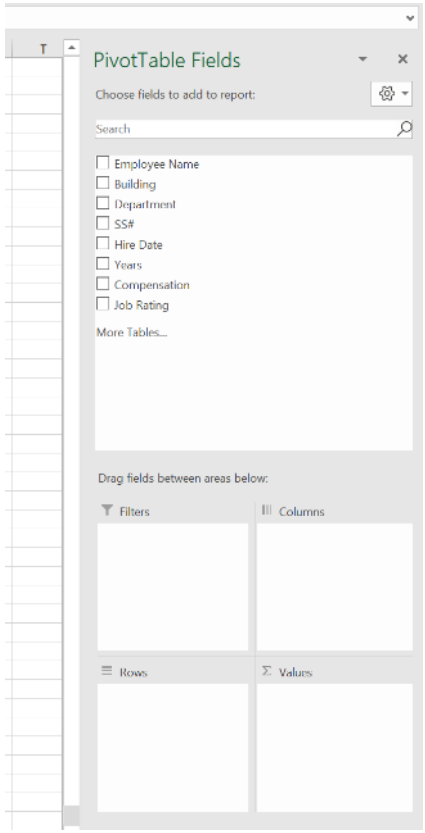
Don’t change anything and click OK



Pivot Table Controls

You should see these two controls…

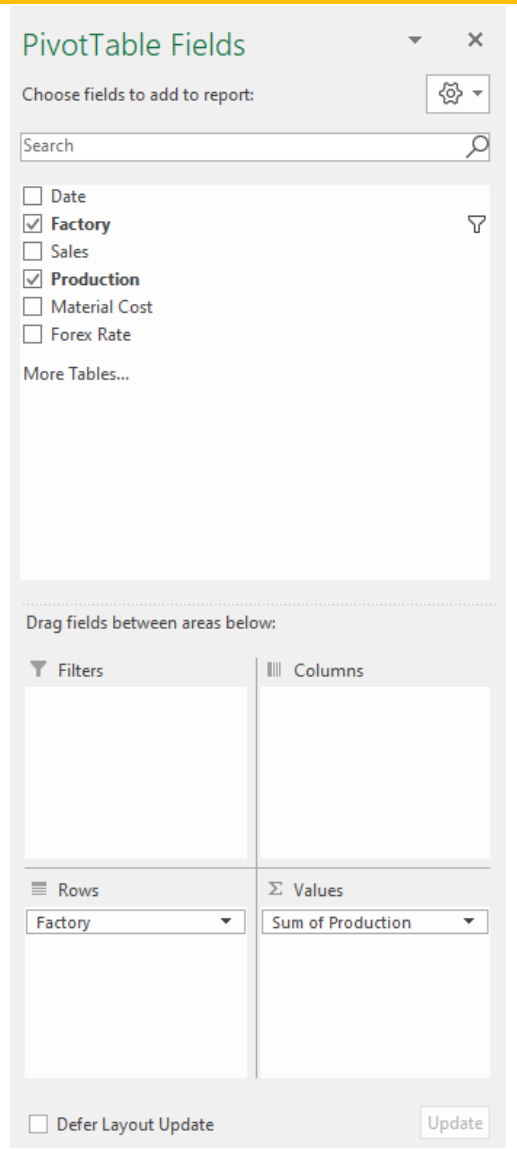




Question 12 continued…

- Put factory in rows and production in the Sigma values
- Use sum function in Sigma values
- Copy and past the table as values in the answer field

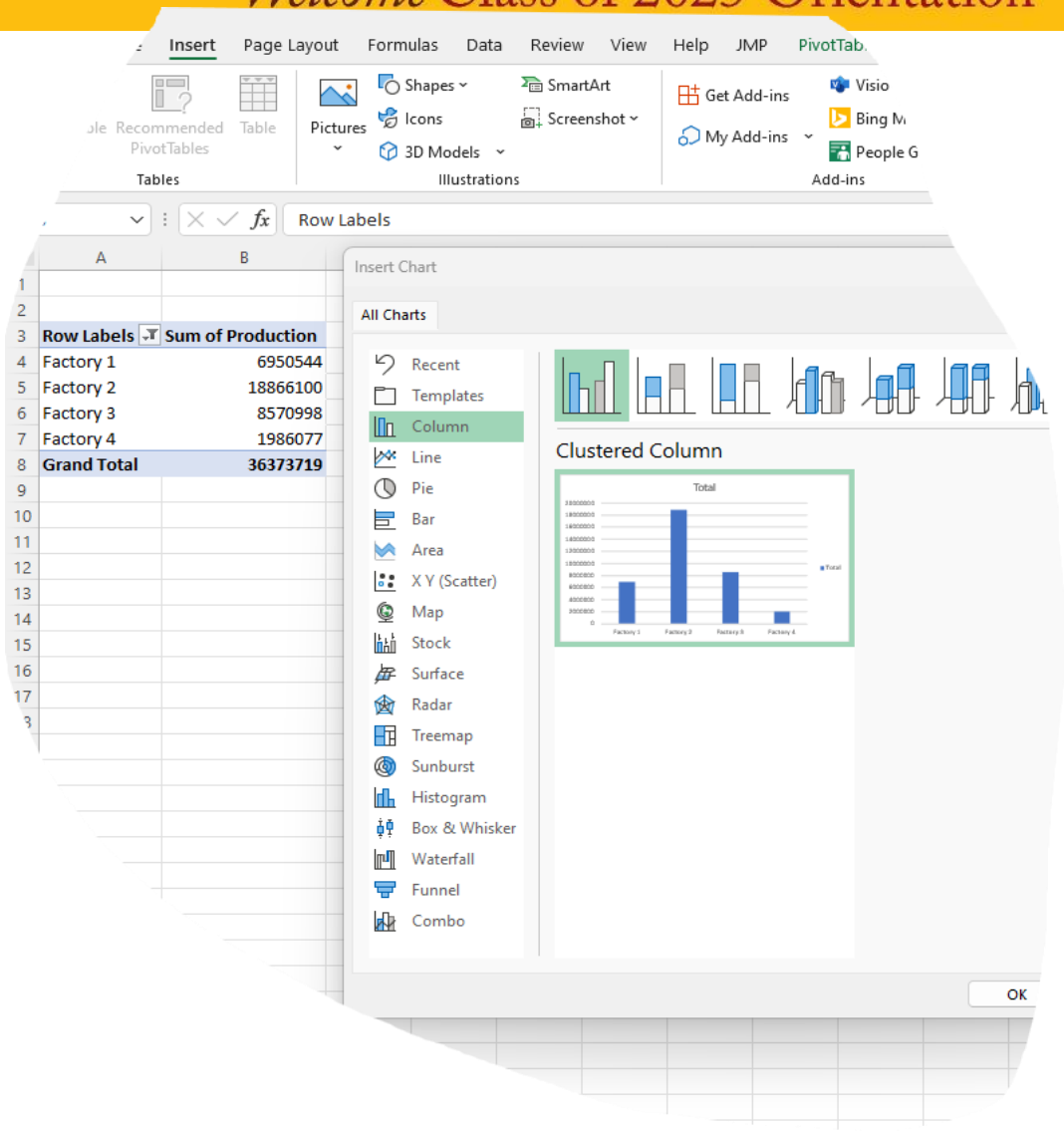
Row Labels	Sum of Production
Factory 1	6950544
Factory 2	18866100
Factory 3	8570998
Factory 4	1986077
Grand Total	**36373719**



Question 13

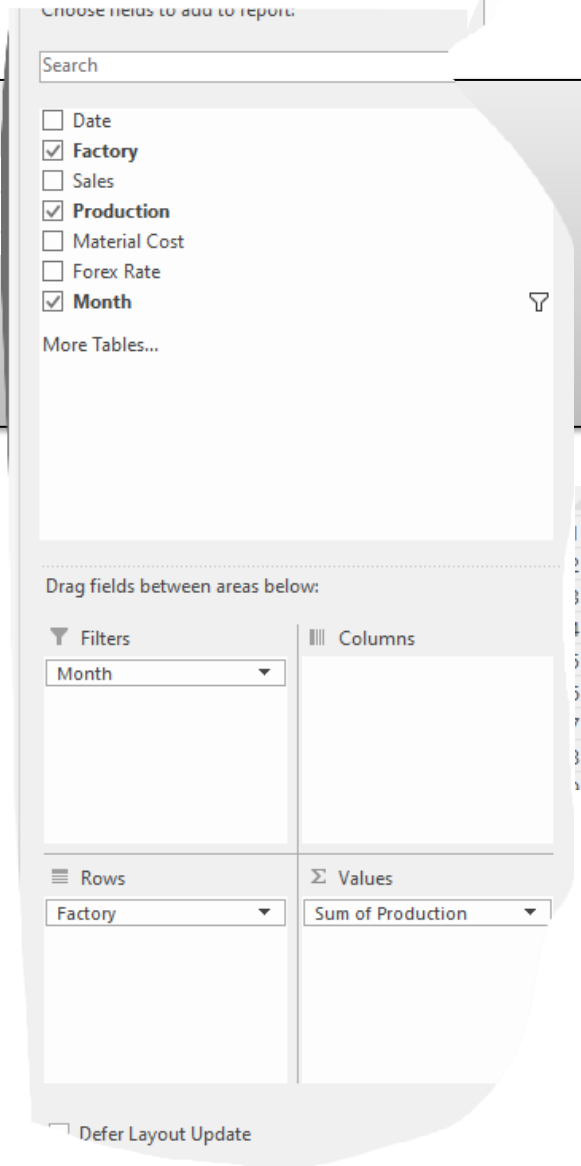
Q13) Make chart of the table made in question 12. Make a bar chart as well as a pie chart.

- Select the pivot table
- Click on the insert tab
- Then, click on pivot chart
- There are several kinds of charts available.



Question 14

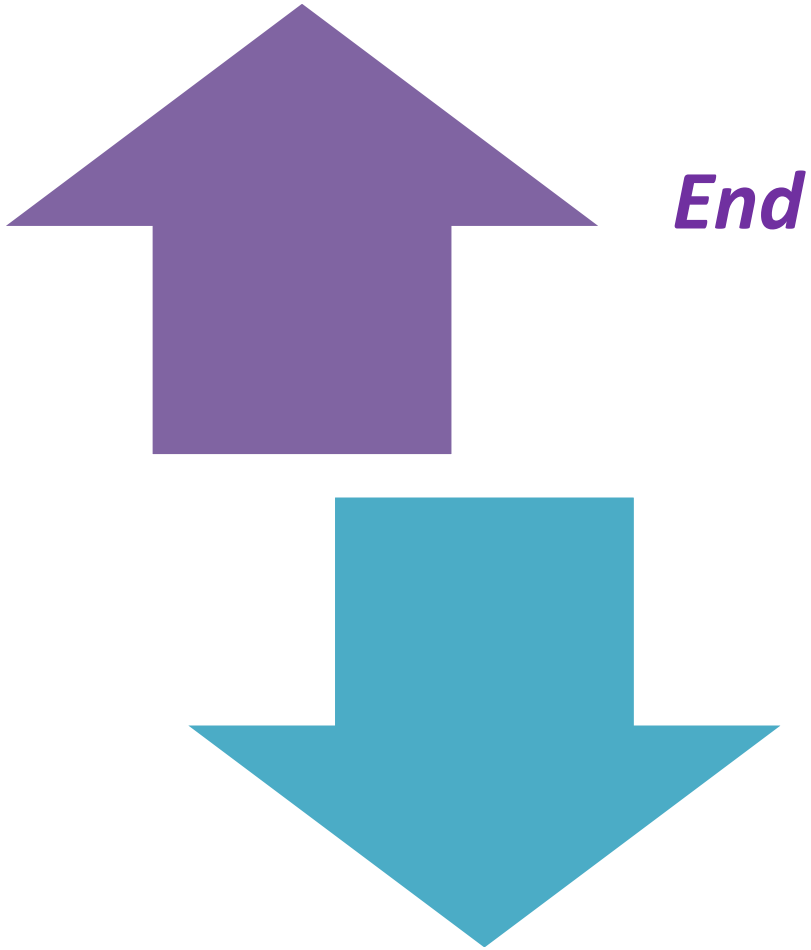
Q14) Report the "total half-yearly production by factory" but only for the first quarter.



A	B
Month	(Multiple Items)
Row Labels	Sum of Production
Factory 1	3450897
Factory 2	9385200
Factory 3	4262575
Factory 4	985698
Grand Total	18084370

- Make a new column in the Log file called months.
- Then, redo the pivot table again with all columns.
- We use the filter tab in the Pivot table.
- We filter on months.
- Note that, month 1,2,3 constitute the first quarter.

Question 15, 16: Exercise.



End of Session 1

Will continue in session 2 tomorrow

Question 17

Q17) Report the "total half-yearly production by factory table" in question 12 but as % shares of each factory.

Factory 1	
Factory 2	
Factory 3	
Factory 4	

- Use the pivot table done before in Q12
- I find it easier to copy-paste the table as values and not as formula in the side
- Use cell locking (anchoring) to find proportions
- Here we use absolute cell locking
- Use formatting to make percentages

Row Labels	Sum of Production
Factory 1	6950544
Factory 2	18866100
Factory 3	8570998
Factory 4	1986077
Grand Total	36373719

Row Labe	Sum of Production	
Factory 1	6950544	=E4/$E$8
Factory 2	18866100	0.518674
Factory 3	8570998	0.235637
Factory 4	1986077	0.054602
Grand Tot	36373719	1

Anchoring a cell (Cell Locking in Excel)

- Consider cell B2 (say)
- What is the $ around B2 mean?
- $B$2 tells Excel to stay fixed on B2 even if you copy and paste this formula to another cell

Absolute Referencing

- allows you to tell Excel to stay on a cell and not move it, even if you copy and paste to another cell
- **Shortcut for Absolute Referencing:**
  - While writing the formula: F4
  - To edit an existing formula: F2 (to edit the formula) and F4 (to add in absolute referencing)
- Or, you can type $ signs on your own

Relative Referencing

- This is called **relative referencing**. When you copy and paste a formula, Excel adjusts it so that:
  - If you copy and paste one column over, Excel will change the formula so that it is one column over (but keep the row the same)
  - If you copy and paste one row down, Excel will change the formula so that it is one column down (but keep the column the same)

Question 18

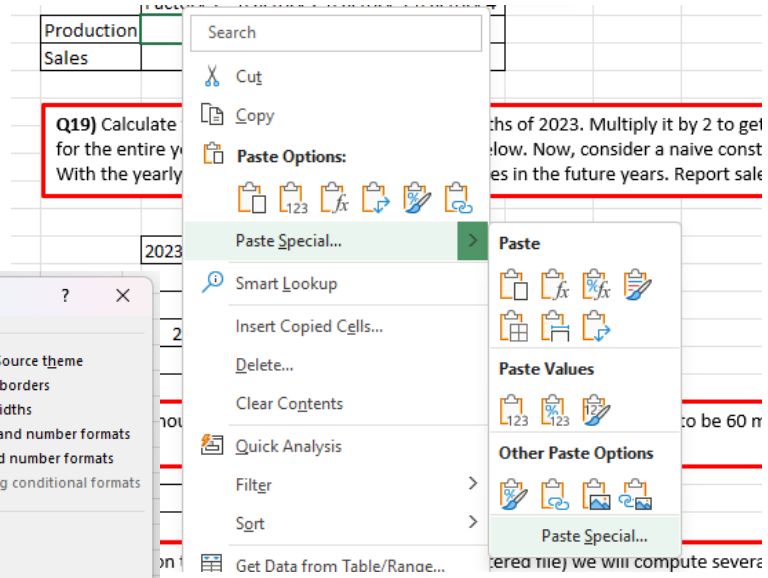
Q18) Report a table with total half-yearly production and total half-yearly sales as percentages contributed from each factory. Print it in the format specified in the answer box.

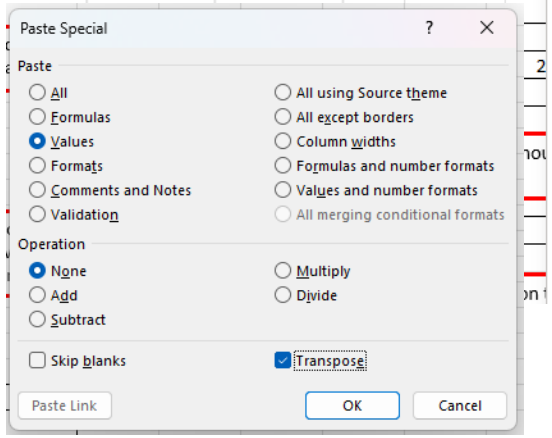
	Factory 1	Factory 2	Factory 3	Factory 4
Production				
Sales				

- Make pivot table again
- Use cell locking but not relative locking than absolute
- Copy paste your answer as paste special as transpose

Row Labels	Sum of Production	Sum of Sales
Factory 1	6950544	16407540
Factory 2	18866100	10262180
Factory 3	8570998	6720000
Factory 4	1986077	3253265
Grand Total	**36373719**	**36642985**

Row Labe	Sum of Pr	Sum of Sales		
Factory 1	6950544	16407540	=F4/F$8	0.447768
Factory 2	18866100	10262180	0.518674	0.280059
Factory 3	8570998	6720000	0.235637	0.183391
Factory 4	1986077	3253265	0.054602	0.088783
Grand Tot	36373719	36642985	1	1





Question 19

Q19) Calculate the total of all sales in the first 6 months of 2023. Multiply it by 2 to get an estimate of sales for the entire year of 2023 and report it in the box below. Now, consider a naive constant growth model. With the yearly growth % given below predict the sales in the future years. Report sales in millions.

Growth	5%

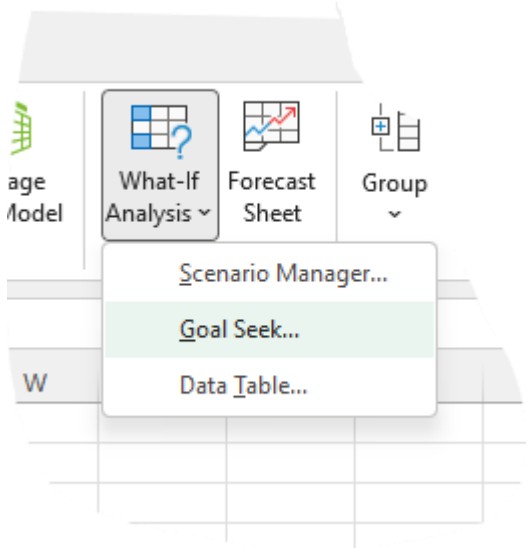
	Predicted Sales				
2023 sales	2024	2025	2026	2027	2028
=(1+$F$22)*B26			42.42	44.54	46.77

- Use sum function to get half-yearly Sales in 2023
- Multiply it by 2 to get 2023 Sales
- Use multiplication and cell locking to fill the predicted sales table

Question 20

Q20) What should be the growth % in question 19 if we would like the sales to be 60 milllion at the end of 2028.

- Under “Data” tab you will see “What-if Analysis”
- Select the Goal Seek function from there



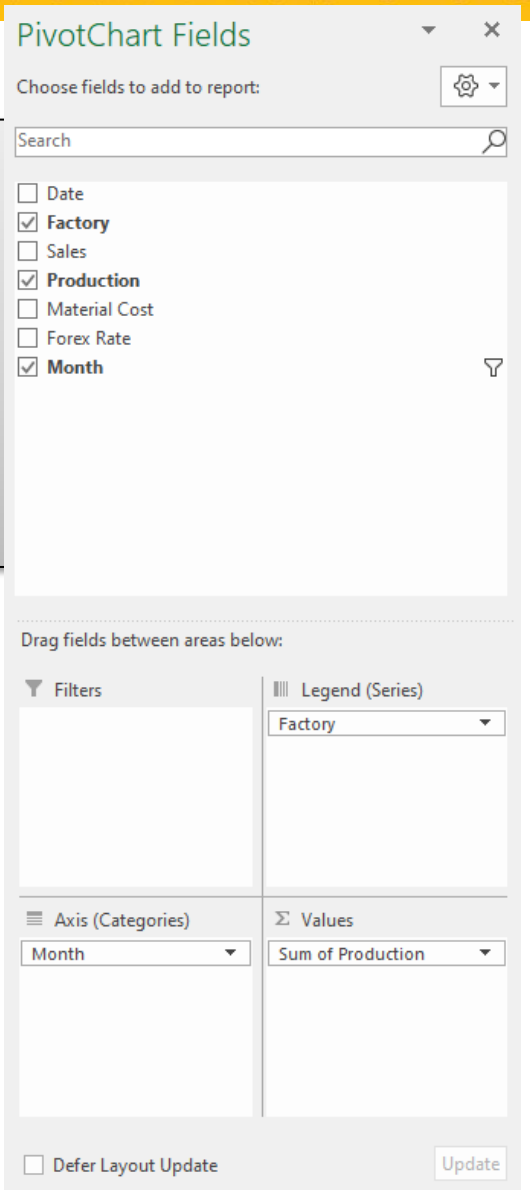
Goal Seek is a built-in What-If Analysis tool in Microsoft Excel. It determines what value to enter in an input cell to get the desired result in a formula cell. It helps you find the missing number when you know some of the numbers involved in a calculation. Goal Seek uses the trial and error method to achieve the desired result. It only works if there's only one input value.

Question 21 & 22

Q21) Based on the complete data file (and not the filtered file) we will compute several Key Performing Indices (KPIs) related to productivity, cost-analysis and sales. For this we will use pivot table. Calculate the following.
a) Total production by factory by month
b) Total sales by factory by month ($)
Keep them in new worksheet and name it appropriately.

Q22) Make a dashboard containing the above 2 charts. See if any trends are visible.

Use pivot table and pivot chart.



Question 23

Q23) Use Vlookup to fill up two new columns in the "log" worksheet. These two new columns should show the (a) operating cost (b) labor cost per laborer. Report both indices in local currency.

Vlookup – Matching items in a list

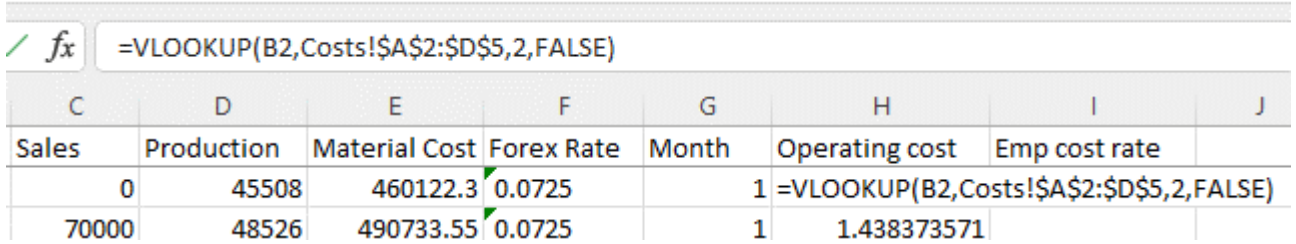
Syntax = Vlookup(What to Find, Where to Look, Item to Return, True/False)

This is what you are looking for

This is the range Excel data that the data you are looking for is within. The data you are looking for must be in the leftmost column

After the item is found in the range, you can return any item. You specify the item to return by typing its offset number from the leftmost column

False – The item you are looking for mush be an exact match
True – The item you are looking for can be an approximate match



C	D	E	F	G	H	I	J
Sales	Production	Material Cost	Forex Rate	Month	Operating cost	Emp cost rate	
0	45508	460122.3	0.0725	1	=VLOOKUP(B2,Costs!$A$2:$D$5,2,FALSE)		
70000	48526	490733.55	0.0725	1	1.438373571		

*Vlookup is a cell formula. For automation, you may have to lock the table in the 2nd input

Question 24

Q24) Use Vlookup again. This is however more complicated. Make a new column "employee" which shows the number of employees for each row.

Vlookup – Matching items in a list

Syntax = Vlookup(What to Find, Where to Look, Item to Return, True/False)

- This is what you are looking for
- This is the range Excel data that the data you are looking for is within. The data you are looking for must be in the leftmost column
- After the item is found in the range, you can return any item. You specify the item to return by typing its offset number from the leftmost column
- **False** – The item you are looking for mush be an exact match
  True – The item you are looking for can be an approximate match

- This is complicated as naively looking it seems that you need to specify two items (Month and Factory) in "what to find" tab to get the correct employee number
- However, we can only input 1 item in "what to find". So, we break up the problem.
- We put in month in "what to find" tab and play around with "item to return" tab
- We make a new column called factory index in the Log worksheet that tells us the column number of the factory in the table in the employee worksheet

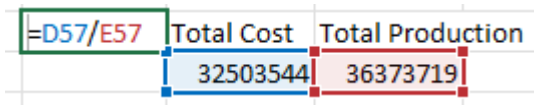
Factory index	Emp no.
=IF(B2="Factory 1", 2, 0) + IF(B2="Factory 2", 3, 0) +IF(B2="Factory 3", 4, 0) +IF(B2="Factory 4", 5, 0)	

Month	Operating cost	Emp cost rate	Factory index	Emp no.
1	1.438373571	633.3667918	4	=VLOOKUP(G2,Employee!$A$2:$E$7,Log!J2,FALSE)
1	1.438373571	633.3667918	4	45

Question 25

Q25) Compute the total daily cost in USD for each row. Note, total cost = (material cost + operating cost rate *production+ employee cost rate *number of employees) *Forex rate.
What would be a rough break-even price for each unit?

- Calculate total costs using formula
- Break even point (ignoring other costs) = Total Cost/ Total Production



Question 26

Q26) Report the following KPIs and make a dashboard of the charts. Report any visible trends.

a. Total production by factory by month (units)
b. Total cost by Factory by month ($)
c. Total sales by factory by month (units)
d. Average production per employee by Factory by Month (in units /worker)
e. Average cost per unit Factory by month ($)

Keep them in new worksheet and name it appropriately.

- Make new columns in Log file “Production/employee”, “Cost ($)/Production
- Then, do pivot table and chart

Question 27: Exercise.

Closing comments: Helpful Excel Practices

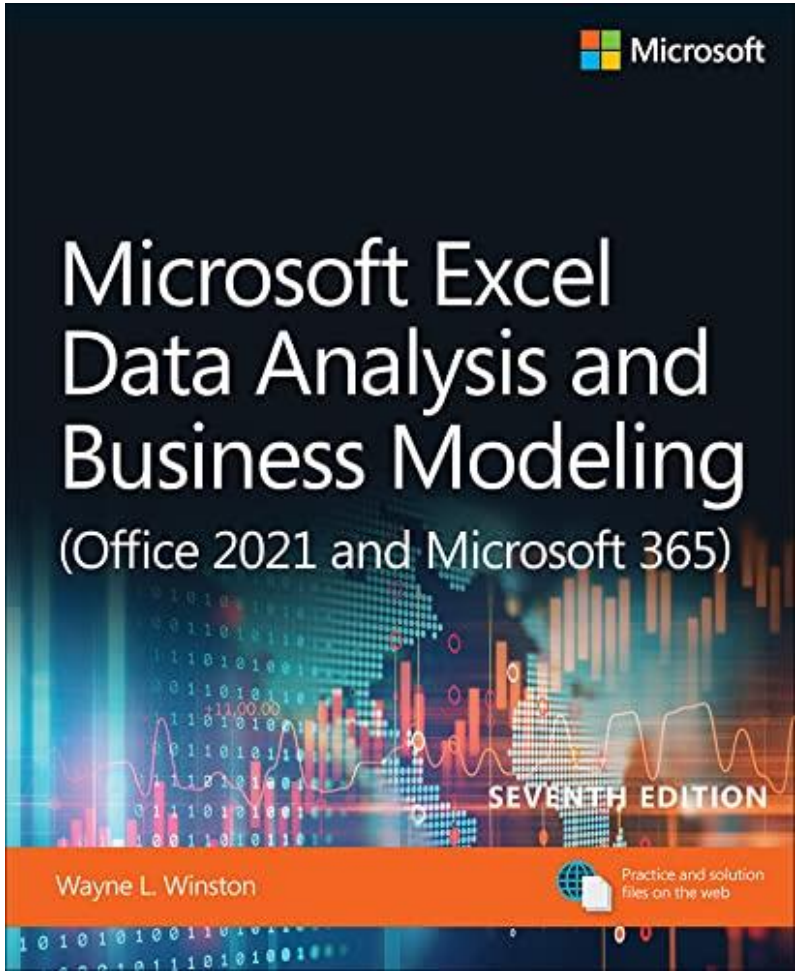
- Set up your model so that you anticipate changes
- **Never hardcode anything that might change (like assumptions or inputs)**
- Remember: The smarter you are today (with building your Excel model), the lazier you can be in the future. Help your future self be lazy. Especially if your boss or clients are fickle
- To edit a function:
  - PC: F2 (you might have to press your FN key and then F2 depending on your computer)
  - Mac: Control U

Closing comments: Helpful Excel Practices

- Using Colors to Communicate: In investment banking, PE and VC, the convention is to use colors to communicate in Excel. Typically,
  - Black: Formulas
  - Blue: Inputs (company-specific numbers or assumptions)
  - Green: Links to other worksheets

  You only change the blue cells and leave the black cells alone
- Using SHORTCUTS makes you go faster! They will save so much time and more importantly, it's so much cooler

REFERENCE:



Q&A

Contact:

gourab@usc.edu

See you on 21st August ☺