

LAB 1 – EXCEL FOR DATA ANALYSIS

OBJECTIVES

1. Define absolute and relative referencing in Excel formulae
2. Describe the process of linear regression in Excel (adding a trendline)
3. Describe the process using “Solver” function in excel

ABSOLUTE AND RELATIVE REFERENCING

In Excel, when we copy a formula to an adjacent cell, the formula is automatically updated, and this is called *relative referencing*. The formula can be forced to use fixed cell by using a “\$” sign in the cell name and this is called *absolute referencing*.

For example: A formula that references cell “C1” will change to “C2” when you go copy the formula one row down and it changes to “D1” when it is copied horizontally to the next column. This is relative referencing. Similarly, for the same cell “C1” absolute referencing works as follows:

\$C\$1 – Fixes both column and row. The formula will always reference the cell “C1” no matter where it is copied in the spreadsheet

\$C1 – Fixes the column to C but the row number changes when it is copied to a different row

C\$1 – Fixes the row to 1 but the column number changes when it is copied to a different column

[Note: Please complete each Task in a separate worksheet in the **same** spreadsheet]

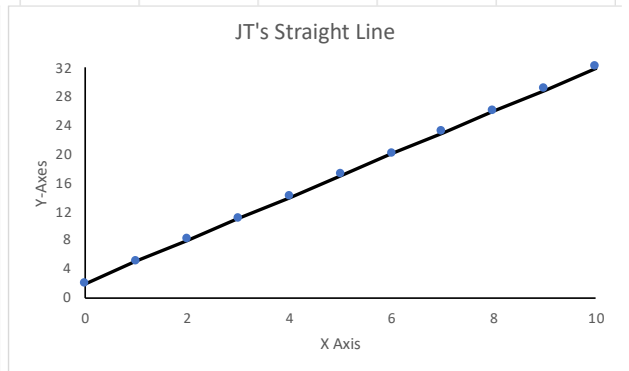
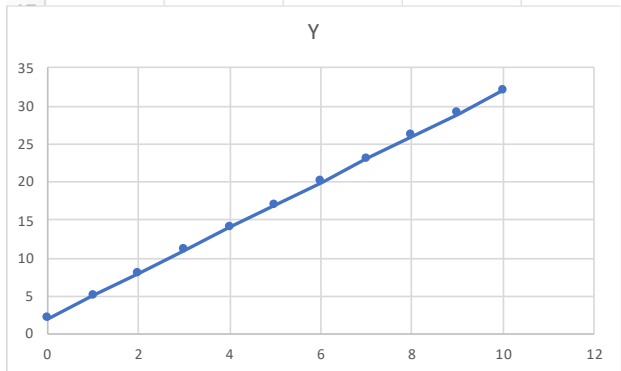
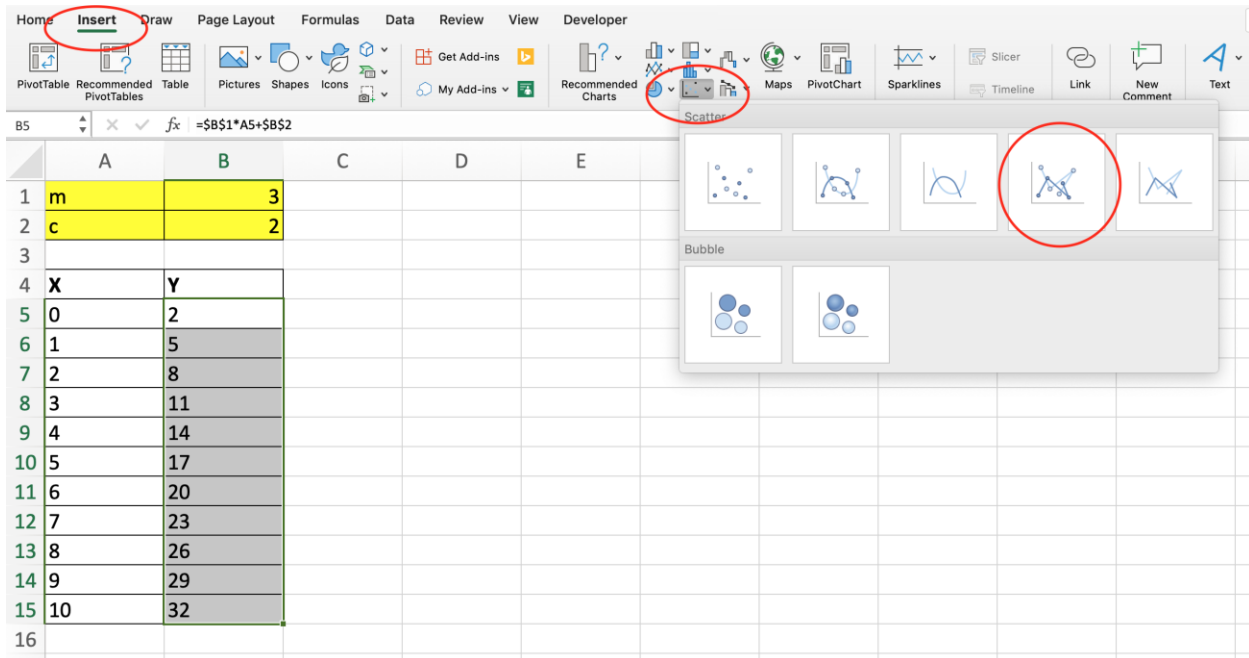
TASK 1: CREATE A PLOT FOR STRAIGHT LINE EQUATION IN EXCEL

The line equation is given as: $y = mx + c$ where y is the dependent variable, x is the independent variable, m is the slope and c is the y-intercept. In this formula, m and c are constants whereas the value of x is input by the user and the y value is computed using the formula.

Consider $x = [1,10]$, $m=3$, $c=2$ and we can set this up in Excel as shown in the Figure. Note that values of m and c in cells B1 and B2 respectively are fixed using absolute referencing since we want the formula to reference the same value each time. In addition, it can be observed that the cell for x -value (A5) is not fixed. Copy the formula down to the rows underneath to get all the values of Y .

Select the values for X and Y and click the insert tab on the top. Select the scatter plot button and choose the “Scatter with lines and markers” option.

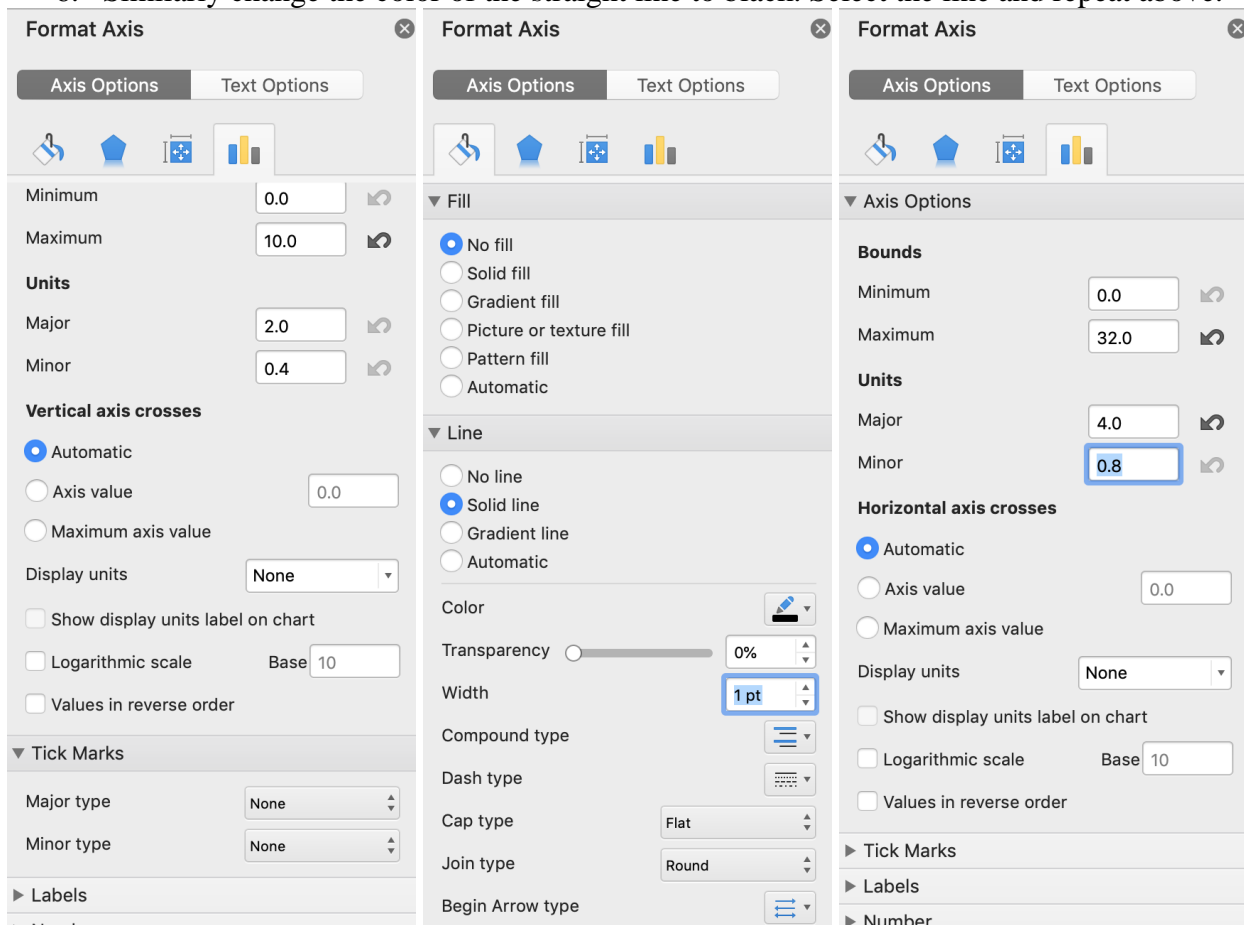
	A	B	C
1	m	3	
2	c	2	
3			
4	X	Y	
5	0	= $\$B\$1*A5+\$B\2	
6	1		
7	2		
8	3		
9	4		
10	5		
11	6		
12	7		
13	8		
14	9		
15	10		



A plot generated from the process is shown above on the left side and we will transform this generic Figure into a cleaner Figure by adding axes etc as shown on the right side,

1. Select the horizontal and vertical grid lines and press delete. This will remove them from your plot. Click on the “Chart Title” shown as “Y” and you can change it to “<Your Name>’s Straight Line”
2. Select the “Chart Design” tab and click “Add Chart Element”. Note that you have to select the Figure for the “Chart Design” tab to appear. From the “Add Chart Element” select “Axes Titles”. Add both the Primary horizontal and Primary Vertical Axes. Change the chart titles to X-axis and Y-axis
3. Double click the X-axis and you will get the “Format Axis” menu on the right side of your spreadsheet. You can also bring this menu by right-clicking on the X-axis and selecting the “Format axis” option. Change the Maximum value to 10. **[Figure below]**
4. Select the “Fill & Line” tab in the menu and select the “Solid Line” button and change the line color to black. Change the line width to 1.0 **[Figure below]**

5. Repeat the process for Y-axis. Change the line width to 1.0. In addition, change the maximum of y value to 32 and Major units to 4. **[Figure below]**
6. Similarly change the color of the straight line to black. Select the line and repeat above.

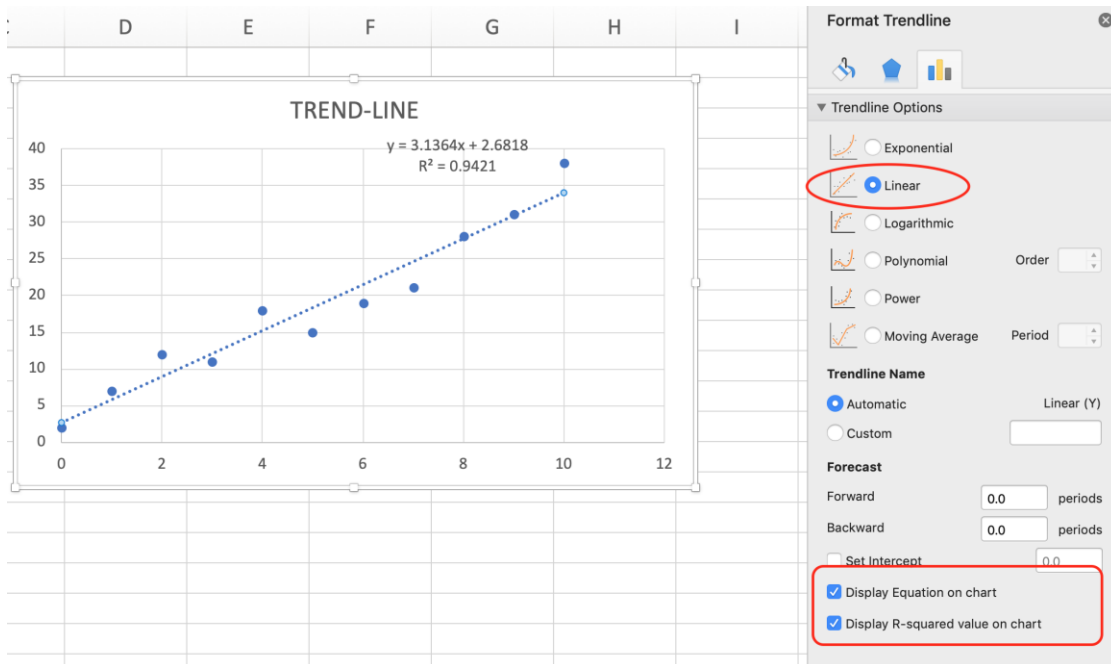
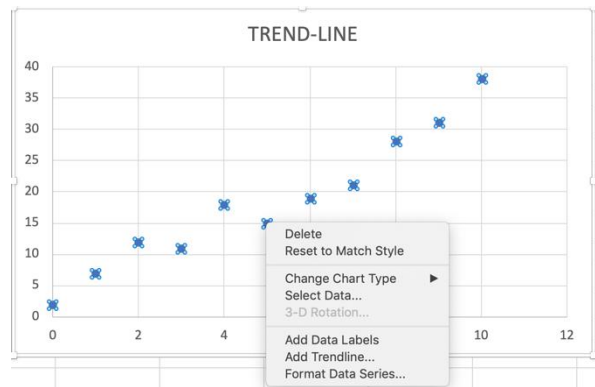


TASK 2: ADD A TRENDLINE FOR A GIVEN DATA

Consider the data below for X and Y axis. We can figure out a relationship between these two variables X and Y by adding a trendline.

X	0	1	2	3	4	5	6	7	8	9	10
Y	2	7	12	11	18	15	19	21	28	31	38

1. Follow the process to draw a scatter plot for the above dataset. Choose the first option with basic scatter plot with just markers [no lines or curves]
2. Choose the “Linear option” from the available options since the markers show a linear trend. Also, select the “Display equation” and the “Display R-squared value” options. The R-square value tells us about the goodness of the fit. The fitness of the trendline is better of the R-squared value is close to 1.



3. Complete this by adding the axes as it was done in Task 1.

TASK 3 – USING THE EXCEL SOLVER FUNCTION

The solver option in Excel is a powerful optimization tools that helps to solve for unknown variables in any equation.

In this task we will compute the depth of a flow in an open channel (y) for a given flowrate (Q) using the Manning’s equation given as:

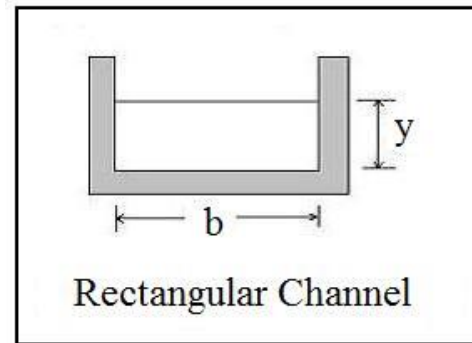
$$Q = \frac{1}{n} AR^{2/3} S_o^{1/2}$$

Where n is the Manning’s roughness coefficient, A is the wetted-area of the rectangular channel, R is the hydraulic radius which is the ratio of wetted-area (A) and wetted-perimeter (P) of the channel and So is the slope of the channel. For a rectangular channel with a width of b and a flow depth of y, area (A), wetted-perimeter (P) and hydraulic radius (R) are given as follows:

$$A = by; P = b + 2y; R = \frac{A}{P} = \frac{by}{b + 2y}$$

Substituting the above parameters in the Q equation above, we have

$$Q = \frac{1}{n} (by) \left(\frac{by}{b + 2y} \right)^{2/3} S_o^{1/2}$$



The above equation is a non-linear equation with unknown “y”

Consider a rectangular channel with the following parameters (given information):

Q = 5 m³/s;
n = 0.013;
b = 5 m;
So=0.0015;

Unknown: y = ?
START WITH AN ASSUMPTION THAT y=1 m

	A	B	C
1	Q	5	m ³ /s
2	b	5	m
3	n	0.013	
4	So	0.0015	
5	y	1	assume
6			
7	A	=B2*B5	Type the formula shown in column “B”
8	P	=B2+2*B5	
9	R	=B7/B8	
10			
11	Q	=1/B3*B7*(B9^(2/3))*SQRT(B4)	

Let us set up the spreadsheet as follows: [Note: Figure above shows formulae used in column C and all the values are computed using an assumption that y = 1m. The solver function will calculate the complete an automatic trial-and-error process until Q = 5 m³/s]

1. The spreadsheet is now setup to complete a manual trial and error process. Change the “y” value in cell B5 until you get Q = 5 m³/s in B11.

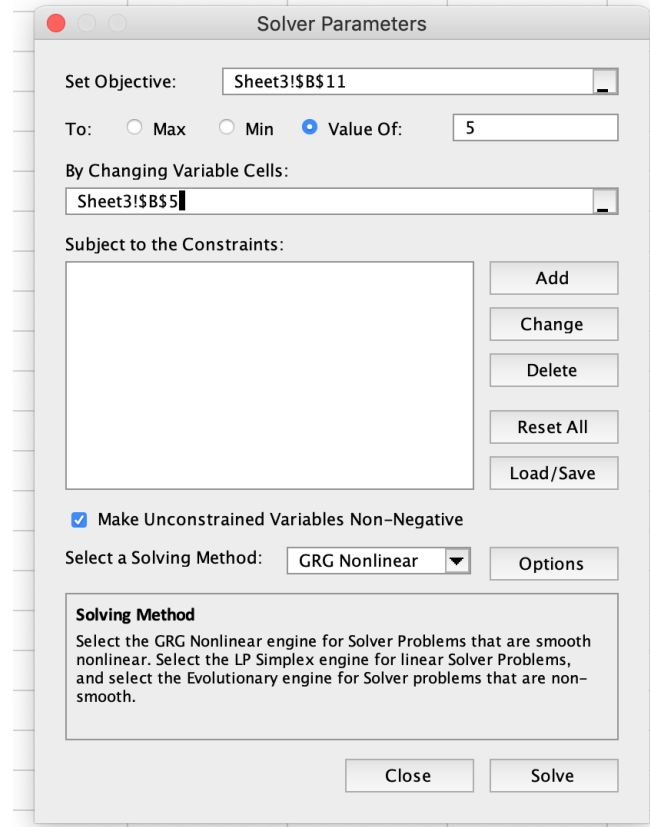
2. This is quite inefficient, and Excel can do this process faster and more accurately. Let us set up the solver. The solver add-in excel needs to be enabled first. The steps to enable solver are available at the link below:

<https://support.office.com/en-us/article/Load-the-Solver-Add-in-in-Excel-612926fc-d53b-46b4-872c-e24772f078ca#OfficeVersion=macOS>

3. In this problem, our goal to get a flowrate (Q) of $5 \text{ m}^3/\text{s}$ by changing the depth (y). Therefore, set objective to cell B11, which is the cell for our computed flowrate. This flowrate should be $5 \text{ m}^3/\text{s}$ as it is given information. Set this to a value of 5

4. We will obtain this value by changing B5 value like you had done manually in step 1. So, set “By Changing Variable Cells” to B5.

5. Click “Solve” button. This should give you a value of y value 0.563438 m and a depth of 4.9999971 which is very close to the original value of $5 \text{ m}^3/\text{s}$.



DELIVERABLES [Individual submission by Slack]

1. Completed Excel file with all the tasks. Each task should be in a separate sheet.
2. Repeat “Task 3” by computing the depth (y) for a **b value of 3 m** and. [complete this in a new sheet]
3. Generate an Excel plot for parabola equation: $y = ax^2$ for $x = [1,10]$ and $a = 4$ [complete this in a new sheet]

CHECKLIST [All your answer should be “yes” or fix it]

- | | |
|--|------------|
| 1. Is your submission one Excel file? | [yes no] |
| 2. Did you complete all three tasks? | [yes no] |
| 3. Did you complete (2) in deliverables section in a separate sheet? | [yes no] |
| 4. Did you complete (3) in deliverables section in a separate sheet? | [yes no] |
| 5. Did you answer yes to all questions above? | [yes no] |