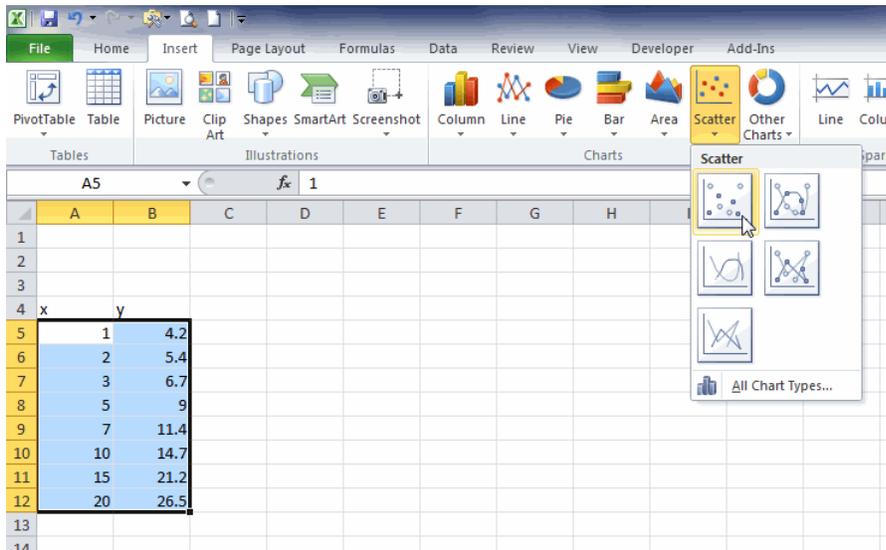


Dealing with Data in Excel 2013/2016

Excel provides the ability to do computations and graphing of data. Here we provide the basics and some advanced capabilities available in Excel that are useful for dealing with scientific data.

Graphing Data

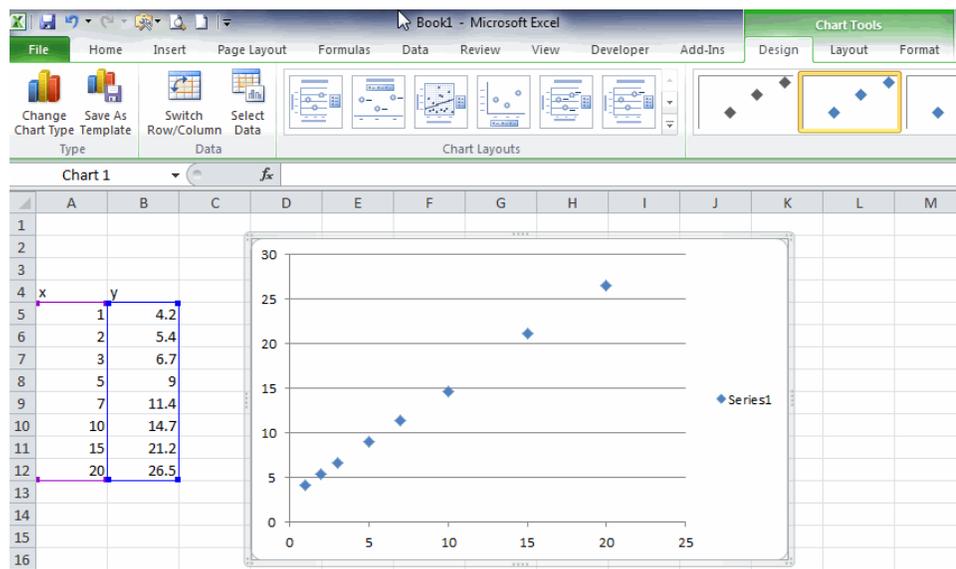
Let's consider graphing the data given below. Go to cell A5 click and hold while highlighting to cell B12.



Then go to the Insert ribbon and click on Scatter to bring down the choices. If you move the cursor around on the choices, an explanation appears.

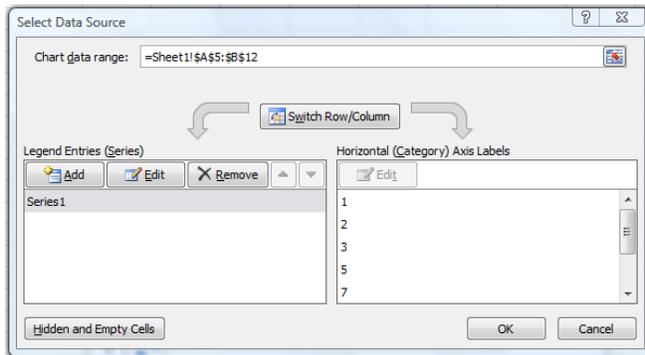
In Excel, the scatter chart is the ONLY graph where the x-axis is treated as a numerical value. (On other charts the x-axis is treated as a category, even if numbers are entered).

Most scientific plots of data, plot only the data points (“Scatter with only markers”). It is usually bad practice to “connect the dots” with lines. When you click on the scatter type, the graph should appear as seen on the screenshot on the right. Notice the data is highlighted (purple is x-axis and blue is the y-axis) and that the Chart Tools appear on the ribbon bar.

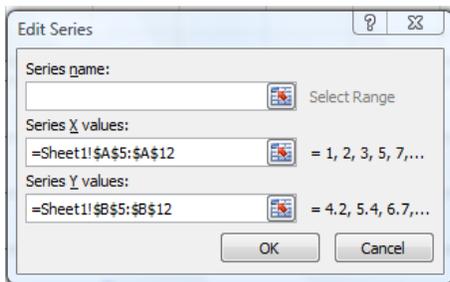


One instance where you may want to use a different type of scatter plot is if you have non-linear data such as absorption spectra: in such cases using the “Scatter/ smooth lines” option may be appropriate.

Click on the Layout ribbon (under Chart Tools, NOT “Chart Layouts” on the Design ribbon) as this is the most useful ribbon for charts. Here you can remove the gridlines, label the x and y axes, add a title, and maneuver the legend if needed. The Chart tools only appear if the graph is highlighted – click on the graph to get them. Clean up your graph at this point!



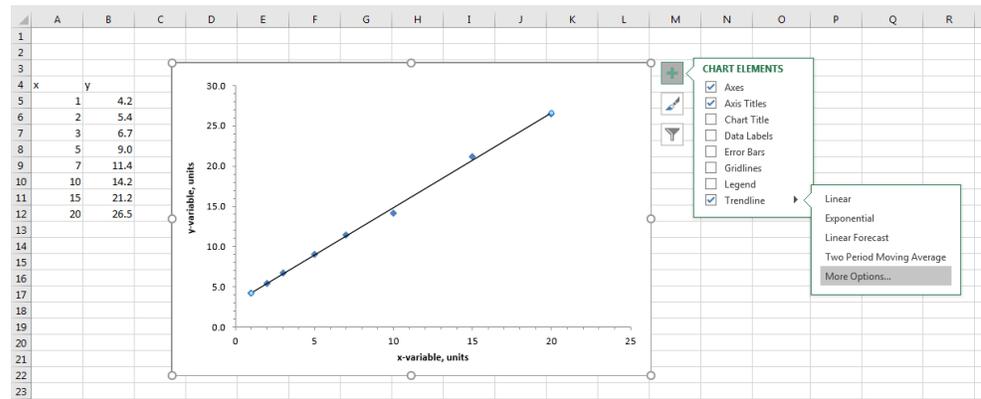
One thing you cannot do in the layout section is change the text of the legend. To do this, you must rename the data series: right click anywhere in the chart and click Select Data... and the Select Data Source menu pops up. Select the series of data you wish to rename (Series 1 here) and click Edit. You can then rename the data series. This is handy if you have more than one data set on the graph and need to show a legend.



You can also switch the x and y axes, or select a new set of data by clicking the value box next to the series X and Y values and highlighting the appropriate data.

Adding a Regression Line or Trendline – The Line of Best-fit

Regression lines are lines of best fit through the data. Excel can do a variety of fits as you will see. Click on chart to highlight it. Click on the large green plus (+) sign to the upper right of



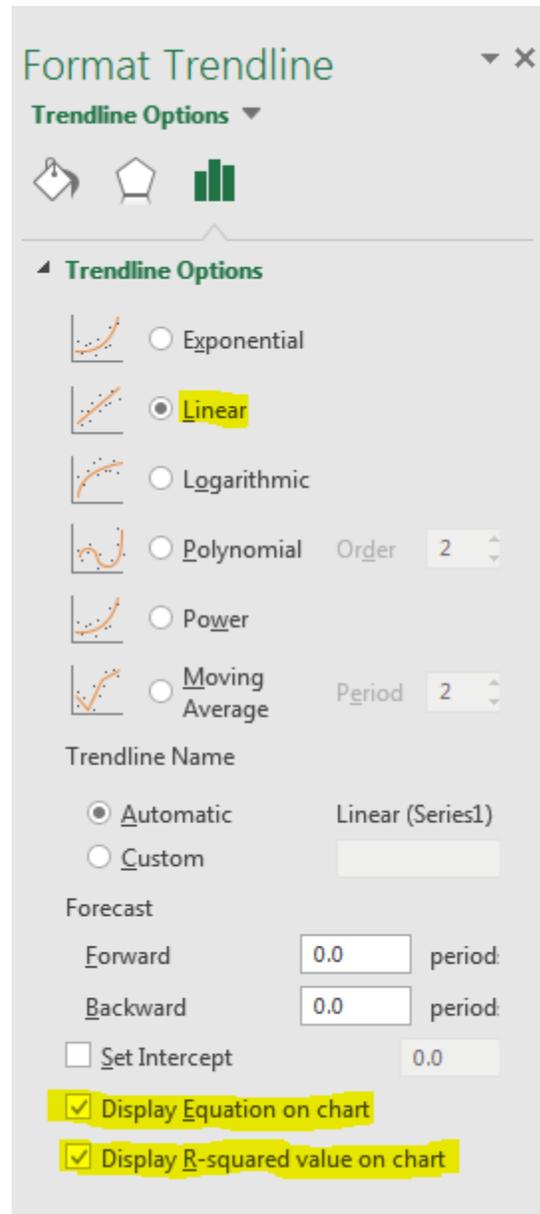
the chart to get the Chart Elements. Then click on Trendline and drop the menu down. Go to More Trendline Options... at the bottom and click on it.

This will get you the Format Trendline menu on the right side and the Trendline Options. This menu will default to a linear regression which you can change by clicking on the other types of regression fits.

You can view these on the graph. Since in most cases we want to know the equation of the model selected and how well it fits the data, go to the bottom of this menu and select Display Equation on chart and Display R-squared value on chart. The equation and r-squared value should appear on the graph.

What does R-squared tell you?

R-squared is the coefficient of determination and a measure of the goodness-of-fit of the equation to the data. A perfect fit has a value of 1.000. In simple terms, R-squared*100 tells you the percentage of the variation of the y-variable due to the variation of the x-variable.



How does linear regression work?

Linear regression minimizes the sum of the square of the deviations, where the deviation is the difference between the calculated y-value minus the y-data value. This is shown visually with user eye-balling a best-fit line using an interactive Excel spreadsheet at:

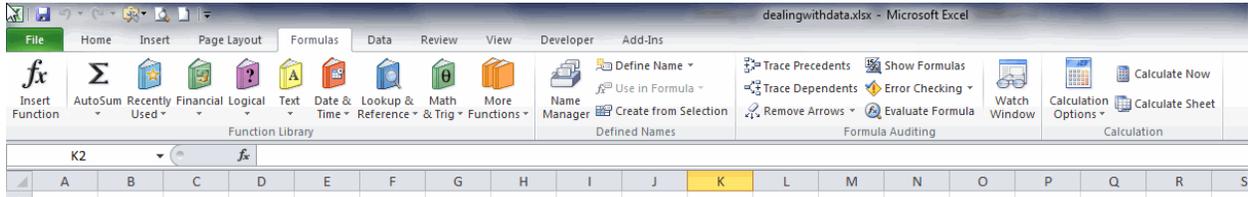
<http://academic.pgcc.edu/~ssinex/regression.xls>

Users can also explore the difference between interpolation and extrapolation with this spreadsheet:

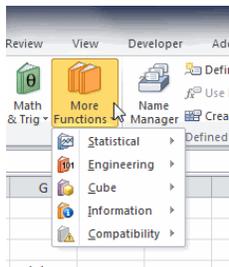
http://academic.pgcc.edu/~ssinex/interpol_extrapol.xls

Simple Computations and Formatting Data

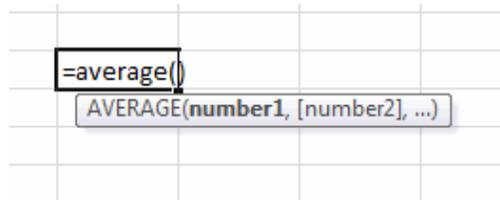
Transformation of data for analysis is very common. A large number of functions are available in the on the Formula ribbon (click anywhere in the spreadsheet outside of a chart, then click Formula) using the Function Library.



The More Functions button includes many common statistical functions such as standard deviation, median, plus others.



You can type the function directly into a cell, or you can select it from the library. Any function will have a specific syntax that will appear when the function is selected. Hence to use average, you place the range of cells to average.



If you need to calculate a quantity down a column of data, click on the first cell and go to the lower right corner (note the cursor changes) and click and hold while dragging down the column.

inches	cm
11	=O14*2.54
15	
32	
12	

inches	cm
11	27.94
15	
32	
12	

All calculations must start with an equal sign, “=” in the cell. To multiply numbers use the asterisk such as 3*2. To raise a number to a power: 5² is 5^2. Many of the functions need only a cell reference to deliver the quantity.

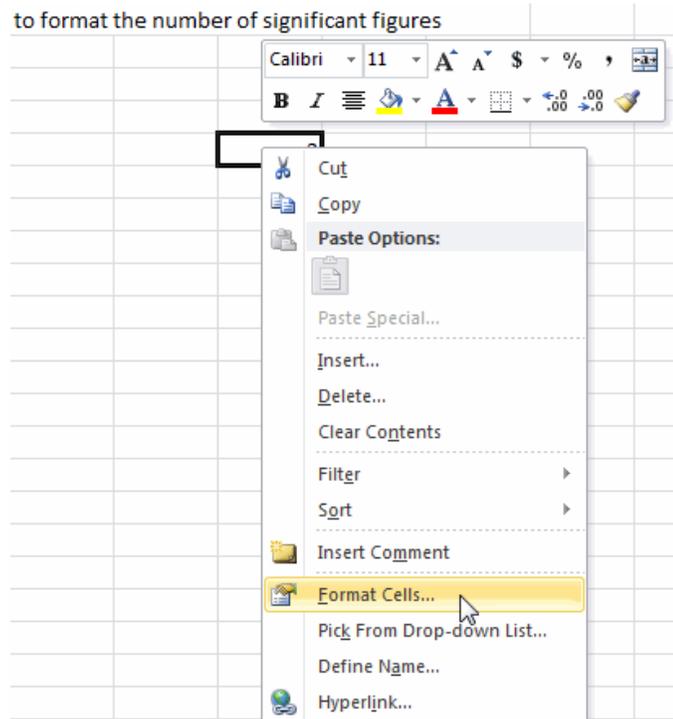
	A	B	C	D	E	F	G	H
1								
2				<u>value</u>		<u>type into column F</u>		<u>this is the result</u>
3	natural logarithm			100		=ln(D3)		4.61
4								
5	common logarithm			100		=log(D5)		2
6								
7	e ^x			1		=EXP(D7)		2.72
8								
9								

	A	B	C
1			
2		k =	0.2
3			
4	x	y	
5	1	0.2	
6	2	= \$C\$2 * A6	

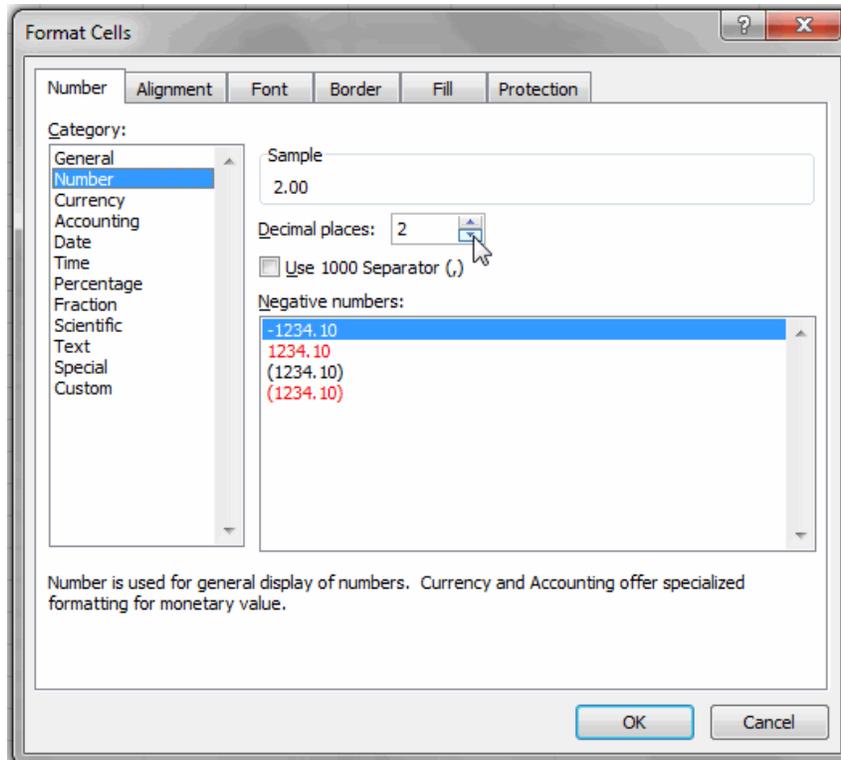
If you want to enter a constant in a separate cell (such as C2) to use in a computation, you will want to change its format to \$C\$2 (This is an absolute reference.) in the calculation cell and hence it will not change on dragging the calculation down the column.

Here we want y =kx.

To format data in cells, click on a cell or highlight a column of data and then right click and select Format Cells...



When the Format Cells menu pops up, select the Number tab. Here you can select Number and set the number of decimal places. You can select Scientific if you want to use scientific notation. In Excel the number 2.0×10^5 will show up as 2.0E5.

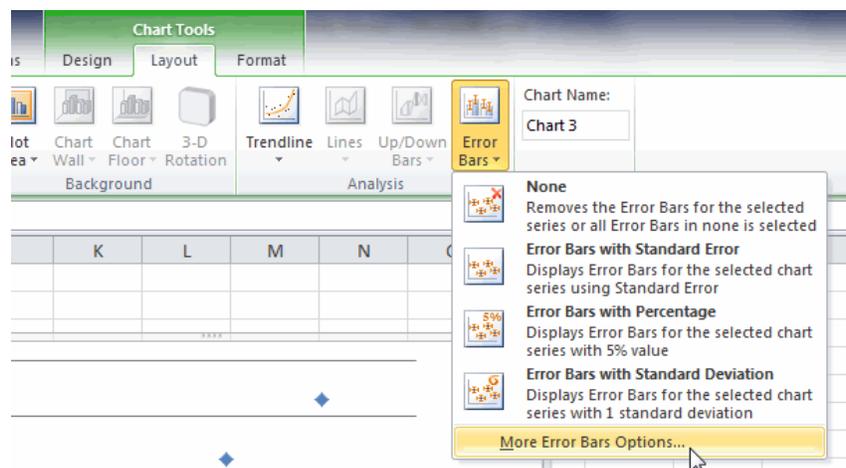


For more information on computational spreadsheets, see the following handout.

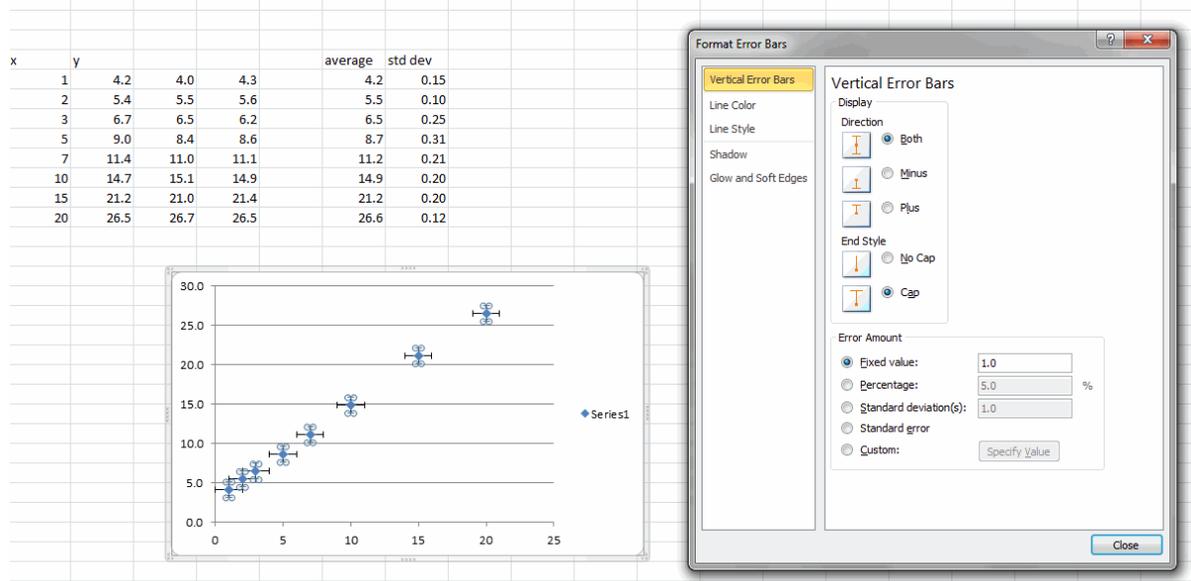
http://academic.pgcc.edu/~ssinex/excelets/Computational_Spreadsheets.pdf

Error Bars on Graphs

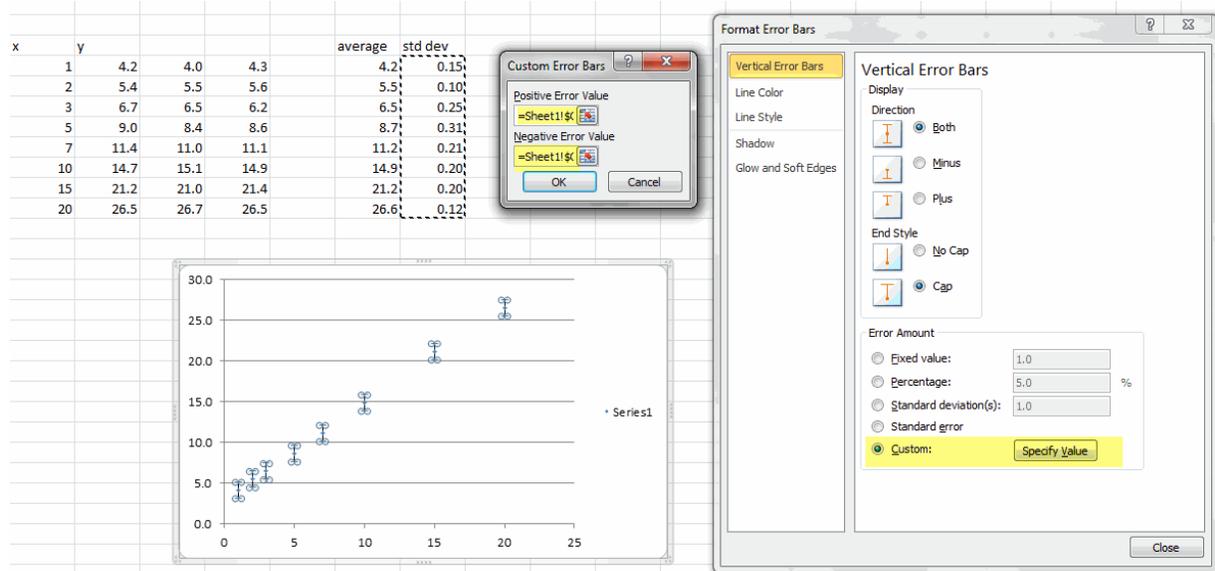
Error bars are a method of expressing the uncertainty in data on a graph. Adding error bars to a graph is done by selecting the graph, then going to the Layout ribbon, selecting the Error Bars menu and then the More Error Bars Option...



Both horizontal and vertical error bars appear with a fixed value. The vertical or y error bars opens up first as seen on the screenshot below of the Format Error Bars pop-up menu. To set the horizontal or x error bar click on the horizontal bars on the graph. If you highlight the x error bars and then hit delete, you can remove them.



We can add custom error amounts such as plus and minus one standard deviation. Here we are using the standard deviation of the three values, so the error bar is plus/minus one standard deviation around the mean value. Select the Custom at the bottom of the Format Error Bar menu and the Custom Error Bars menu pops up. Click in the Positive Error Value box and then highlight the column of data and then repeat for the Negative Error Value.



Standard Error of Slope and Y-intercept for Linear Regression with Confidence Intervals

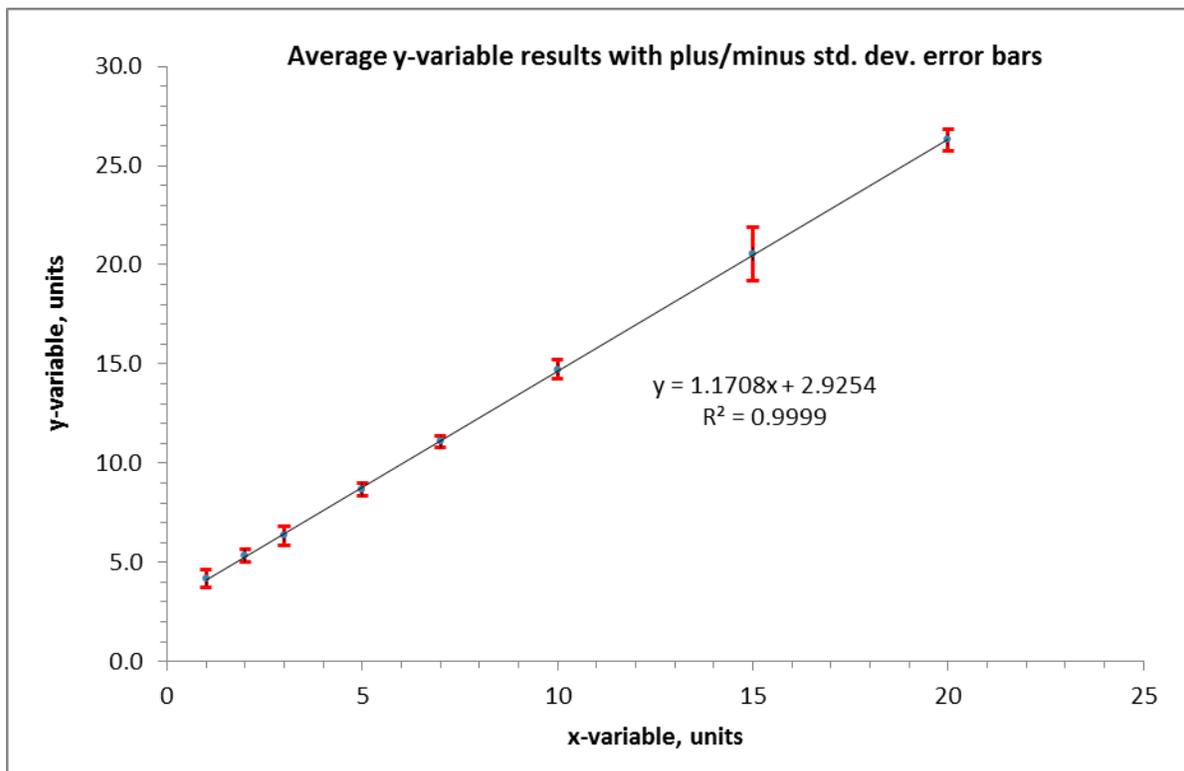
This is a prebuilt spreadsheet to which you can add data that uses LINEST. This spreadsheet calculates the standard error of the slope and y-intercept plus plots confidence intervals on the graph.

http://academic.pgcc.edu/~ssinex/excelets/regression_2.xls

Putting a Excel Graph into a Word or PowerPoint Document

The easiest way is to click on a graph and then right click and select copy in Excel. Move to Word or PowerPoint and paste. You can drag a corner to resize. An alternate method is to use the Snipping Tool in Windows 7 or your favor screen capture software to capture the graph as an image and save as gif or jpeg and name the file.

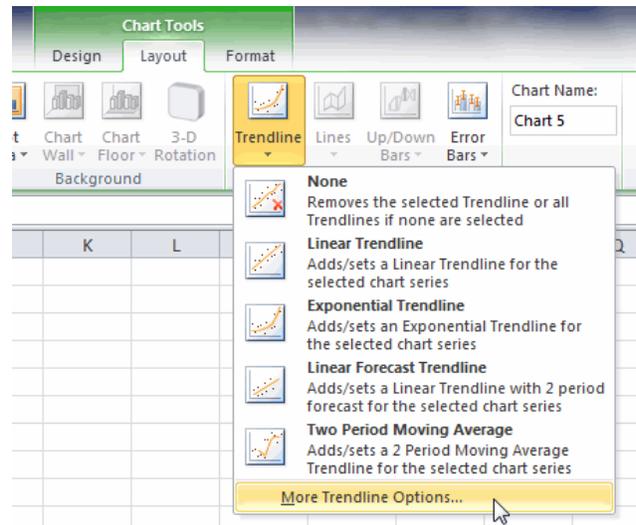
Here is an Excel graph with y-variable error bars on the data points (markers), axes labeled, title, and a linear regression line and equation. Try to keep the colors simple and all items readable on printing especially if using a black ink printer. This is what a graph should look like!!!



Appendix

Adding a Regression Line or Trendline in Excel 2010

Regression lines are lines of best fit through the data. Excel can do a variety of fits as you will see. Click on Trendline and drop the menu down. Go to More Trendline Options... at the bottom and click on it.



This will get you the Format Trendline pop-up menu. You may need to move it off the graph. This menu will default to a linear regression which you can change by clicking on the other types of regression.

