

Linest Function (PC) | ERC

The linest function is a useful tool in Excel to fit a line ($y = mx + b$) to the data in order to identify the relationship between two variables (x and y). In addition to the fitted value, the linest function also produces uncertainty estimates for the fitted values, which we could not get from the [trendline](#) function. The linest function uses the least square procedure which fits a line to a set of data points by minimizing the sum of the squares of the residuals of the points from the curve. You can learn more about least square procedure here: <http://mathworld.wolfram.com/LeastSquaresFitting.html>

Calculating Linest Function

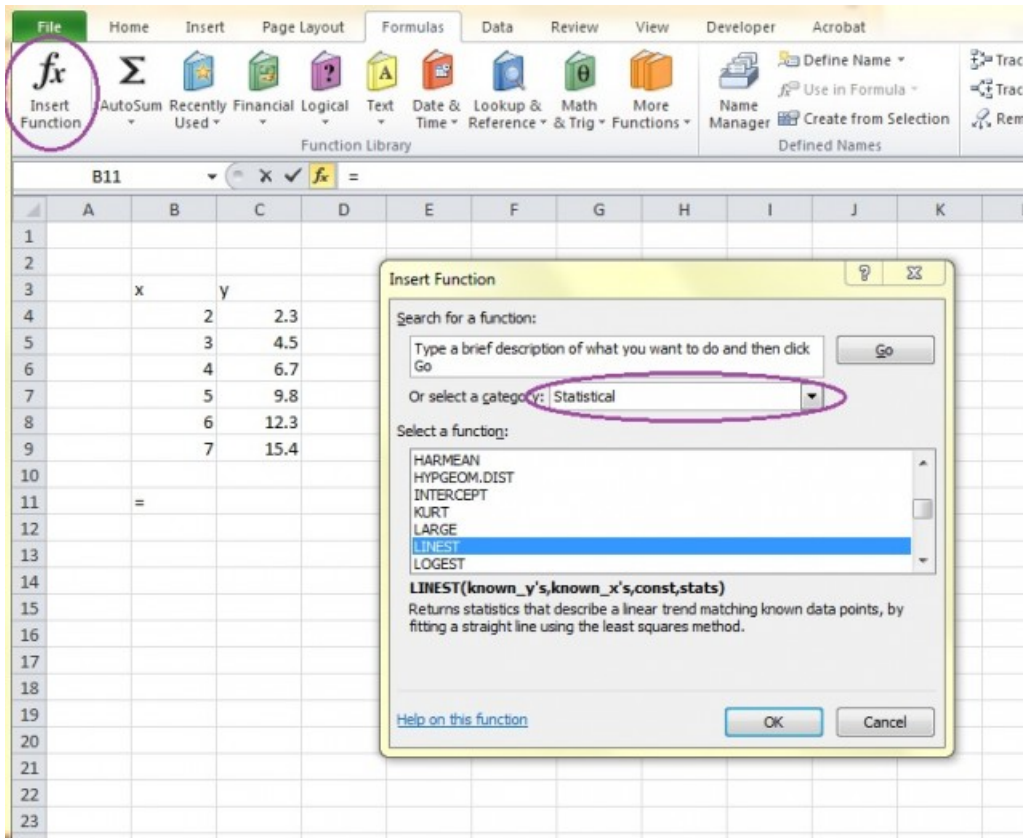
First of all, let us enter the data for which you want to calculate the linest function. Type your data in two columns, one for the x variables and one for the y.

	A	B	C
1			
2			
3		x	y
4		2	2.3
5		3	4.5
6		4	6.7
7		5	9.8
8		6	12.3
9		7	15.4
10			

Select the area that will hold the statistics of the fitted function. Unlike other Excel functions, the linest function has multiple outputs. Each output appears in a different cell. In order to see all of the outputs, you need to select the appropriate number of cells in the appropriate locations. You should draw the mouse to form a 5 row by 2 column area to hold the results, as shown below. (We will explain why and what each entry means) **You only need to form a 3 row by 2 column area.**

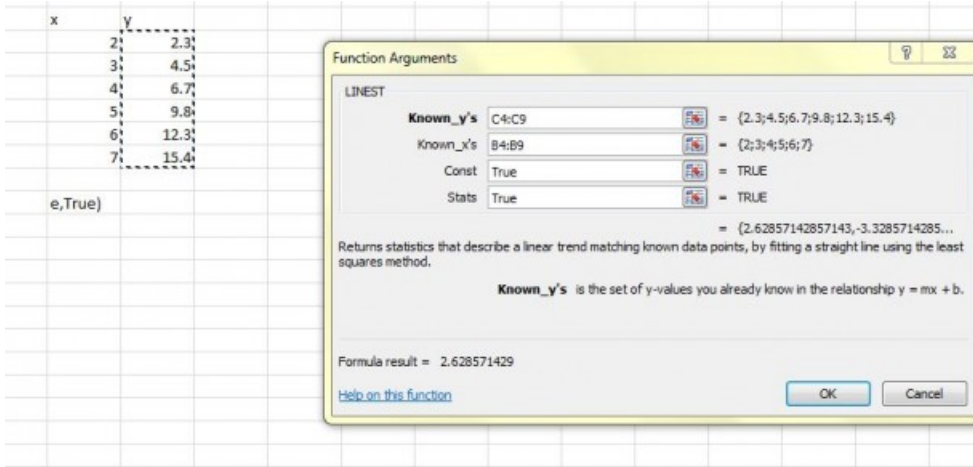
	A	B	C	D
1				
2				
3		x	y	
4		2	2.3	
5		3	4.5	
6		4	6.7	
7		5	9.8	
8		6	12.3	
9		7	15.4	
10				
11				
12				
13				
14				
15				
16				
17				

Click Formulas in the Menu and select Insert Function. In the pop up dialogue box, select the category “Statistical” from the drop down menu. And from that category, scroll down and select function “LINEST”.



Click on OK and a dialogue box will show up to allow you to choose the data you want to fit a line.

Select the y and x values you want to fit with the linest function and type in "True" for the last two dialogue boxes. The first True indicates that you wish the line to be in the form $y=mx+b$ with a non-zero intercept. The second True specifies that you wish the error estimates to be listed. **NOTE THAT IF YOU CHOOSE THE THIRD ARGUMENT, const, TO BE FALSE INSTEAD OF TRUE YOU ARE FORCING THE Y-INTERCEPT TO BE ZERO.**



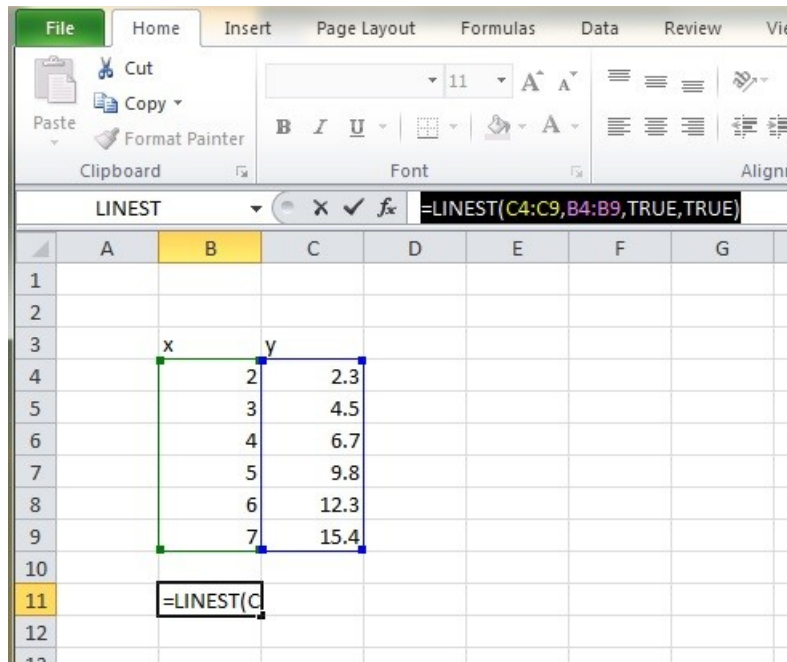
After you click on "OK", you will see a number in the first box of your 5*2 matrix. This number is the coefficient m in the equation $y=mx+b$.

x	y
2	2.3
3	4.5
4	6.7
5	9.8
6	12.3
7	15.4

2.628571	

Getting Least Square Results

The next step will give you the error estimates and other information of your linear model. Highlight the entire formula in the formula bar for the cell containing the coefficient (in this example 2.628571), including the equal sign.



Hold down Control and Shift and press Enter.

Excel will add “{ }” brackets around the formula. You cannot type in the “{ }” yourself since Excel will treat the brackets as part of the formula. Then your least square results will show up in the 5*2 matrix.

	A	B	C	D		
1						
2						
3		x	y			
4			2	2.3		
5			3	4.5		
6			4	6.7		
7			5	9.8		
8			6	12.3		
9			7	15.4		
10						
11	slope	<i>m</i>	2.628571	-3.32857	intercept	<i>b</i>
12	error	<i>u_m</i>	0.084997	0.409106	error	<i>u_b</i>
13	R-squared	<i>R²</i>	0.995835	0.355568	standard deviation of y	<i>s_y</i>
14	F-statistic		956.3842	4	degrees of freedom	
15	regression sum of squares		120.9143	0.505714	residual sum of squares	
16						

u_m is the standard uncertainty of slope (4-21)
u_b is the standard uncertainty of intercept (4-21)
s_y is the standard deviation of y (the fit) (4-19)

THESE TWO ROWS ARE NOT NEEDED

Error estimates for slope and intercept are useful for analysis in a variety of subjects.

Graphing Linest Function

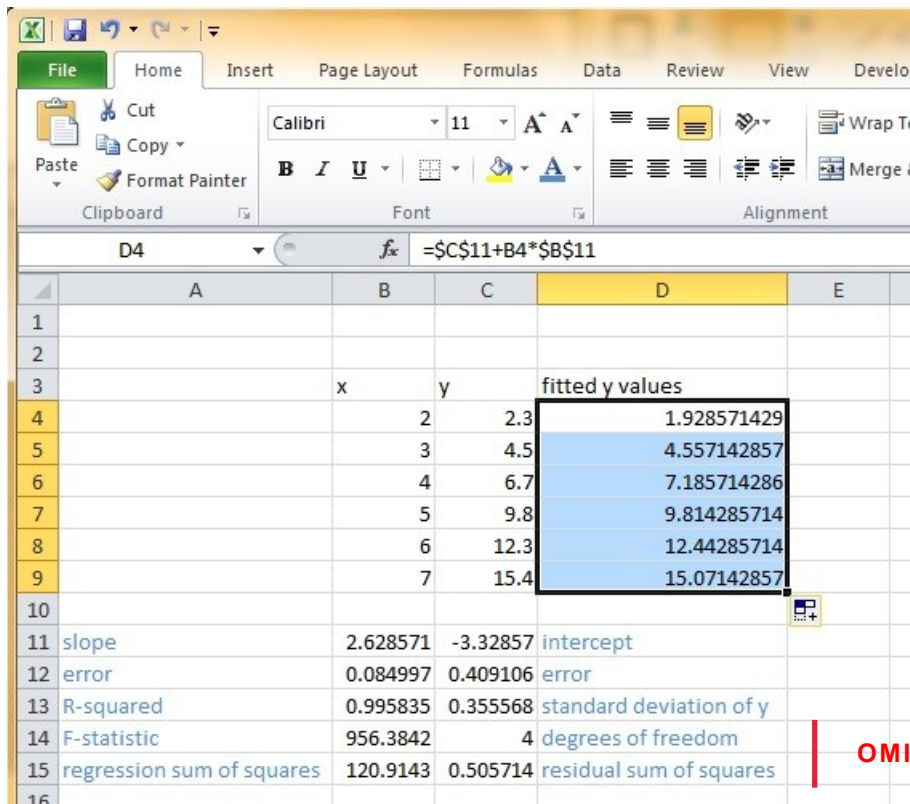
The next step, after getting all the estimates, would be to graph the data points and the fitted line. Notice that the fitted line plotted from the linest function will be the same line as the linear trendline. If you do not need the exact value of fitted y, you could skip the next step and simply add a trendline to the scatterplot of x and y.

We first calculate the fitted values. In the column after the y values, we created the fitted y values which are equal to x values * slope + intercept.

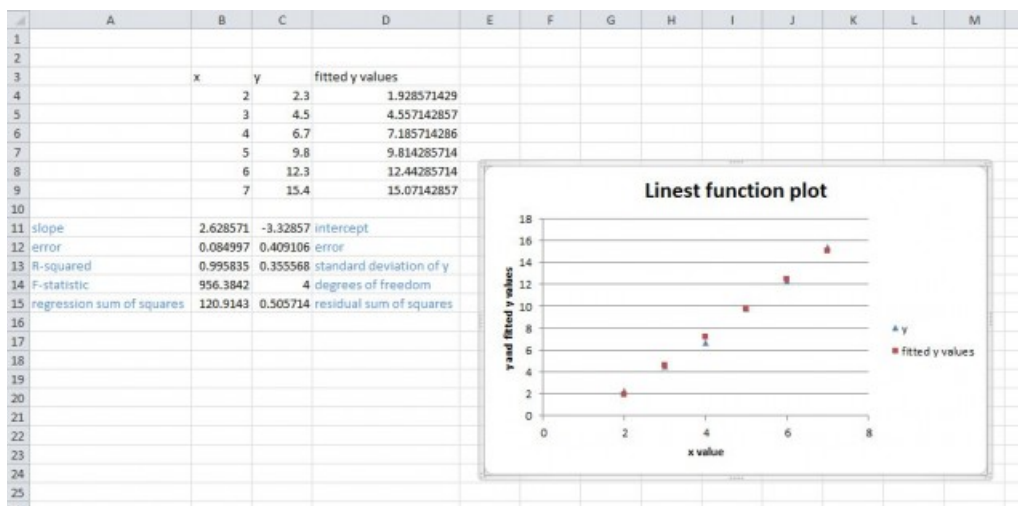
	A	B	C	D
1				
2				
3		x	y	fitted y values
4			2	2.3
5			3	4.5
6			4	6.7
7			5	9.8
8			6	12.3
9			7	15.4
10				
11	slope	2.628571	-3.32857	intercept
12	error	0.084997	0.409106	error
13	R-squared	0.995835	0.355568	standard deviation of y
14	F-statistic	956.3842	4	degrees of freedom
15	regression sum of squares	120.9143	0.505714	residual sum of squares
16				

OMIT

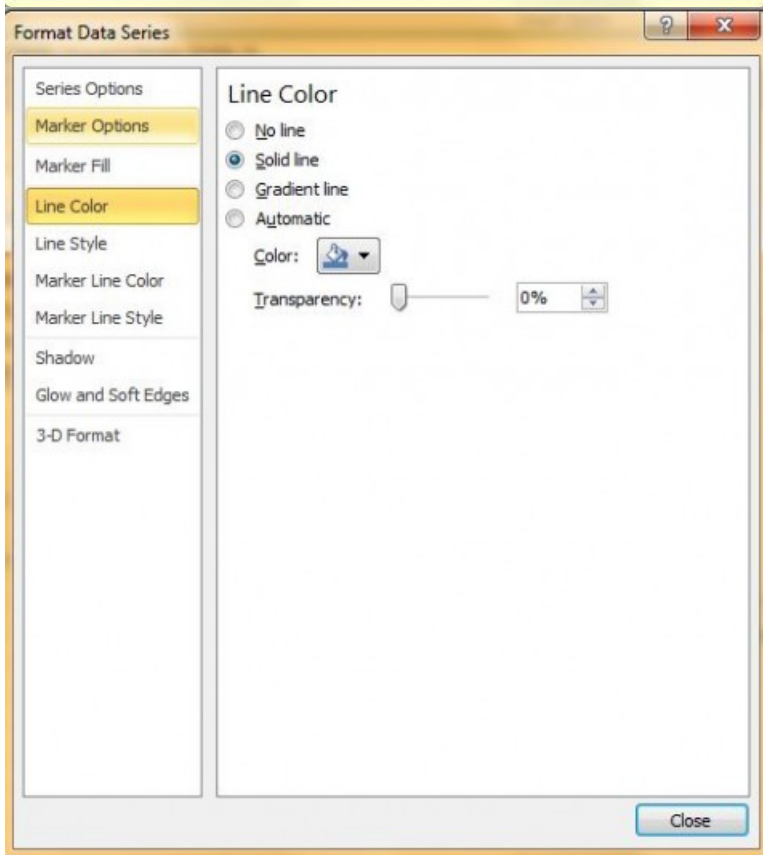
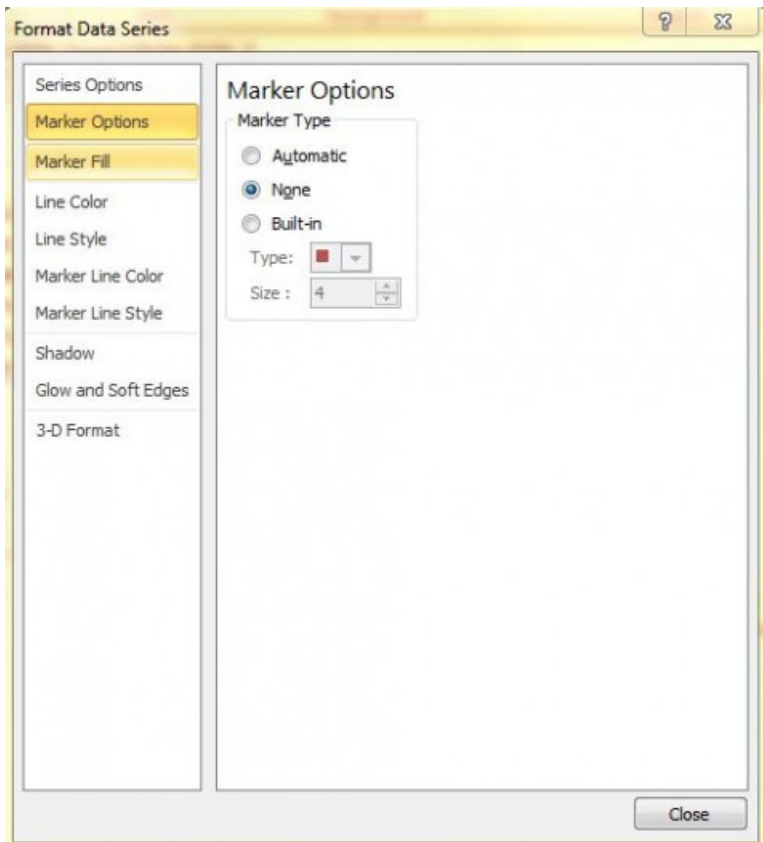
Since we are using the same slope (cell C11) and same intercept (cell B11) for all fitted values, we use absolute reference for the two cells. (Recall absolute reference [here](#)) Then we use auto fill to get all fitted y values. (Recall auto fill [here](#))



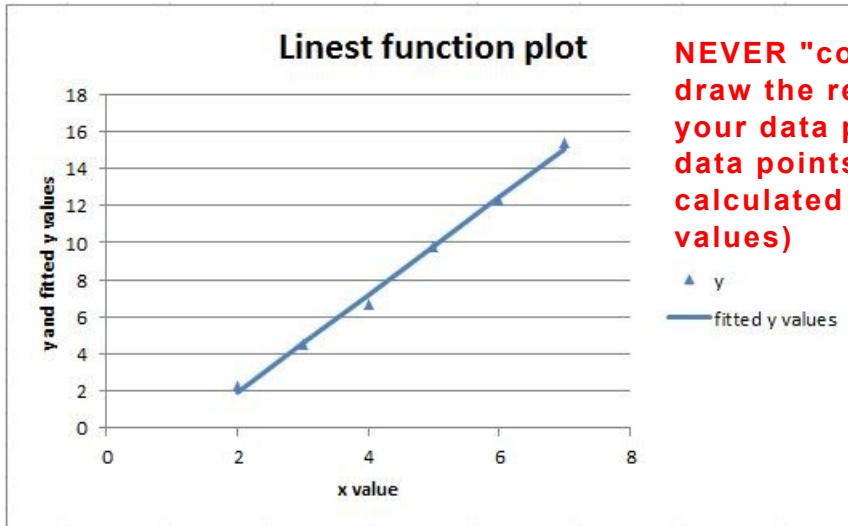
Now we have all information and we are ready to graph the results. Select the x, y and fitted y values and make a scatter plot first. (Recall how to create a chart [here](#))



You can add titles, axis names and edit legend. (To get [help](#) editing charts) Then we replace the plotting symbols for the fitted y values with connecting lines. In order to do this, we double click on one of the fitted y value data points. The Format Data Series dialogue box will pop up. Change the marker type to none and line color to be solid line.



Click on close, and you will get the plot with fitted y values connected by a line.



NEVER "connect the dots" - always draw the regression line through your data points with the actual data points included (not those calculated from the fit - the fitted y values)

▲ y
— fitted y values