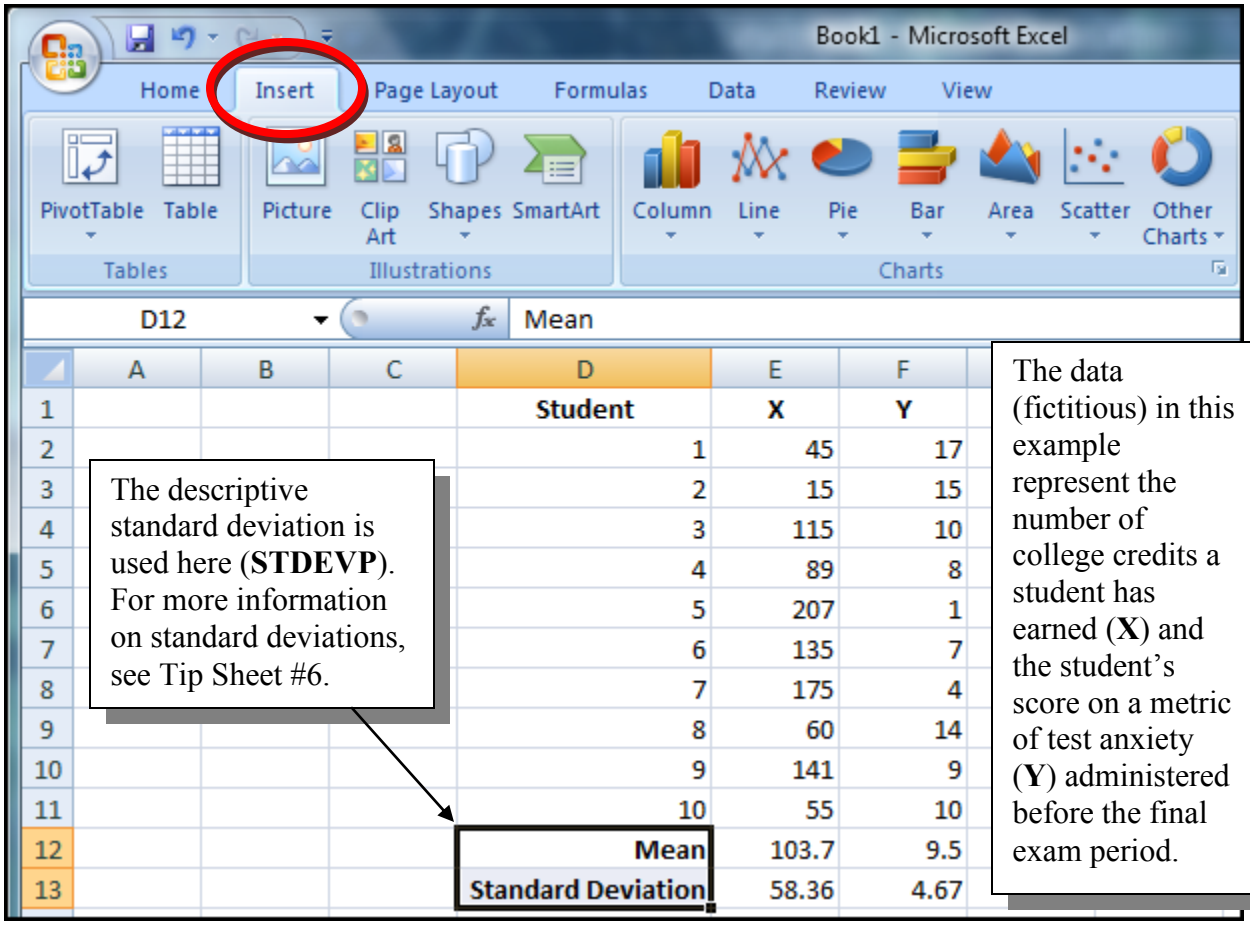
SAMLab Tip Sheet #8

The Linear Regression Equation and the Standard Error of the Estimate

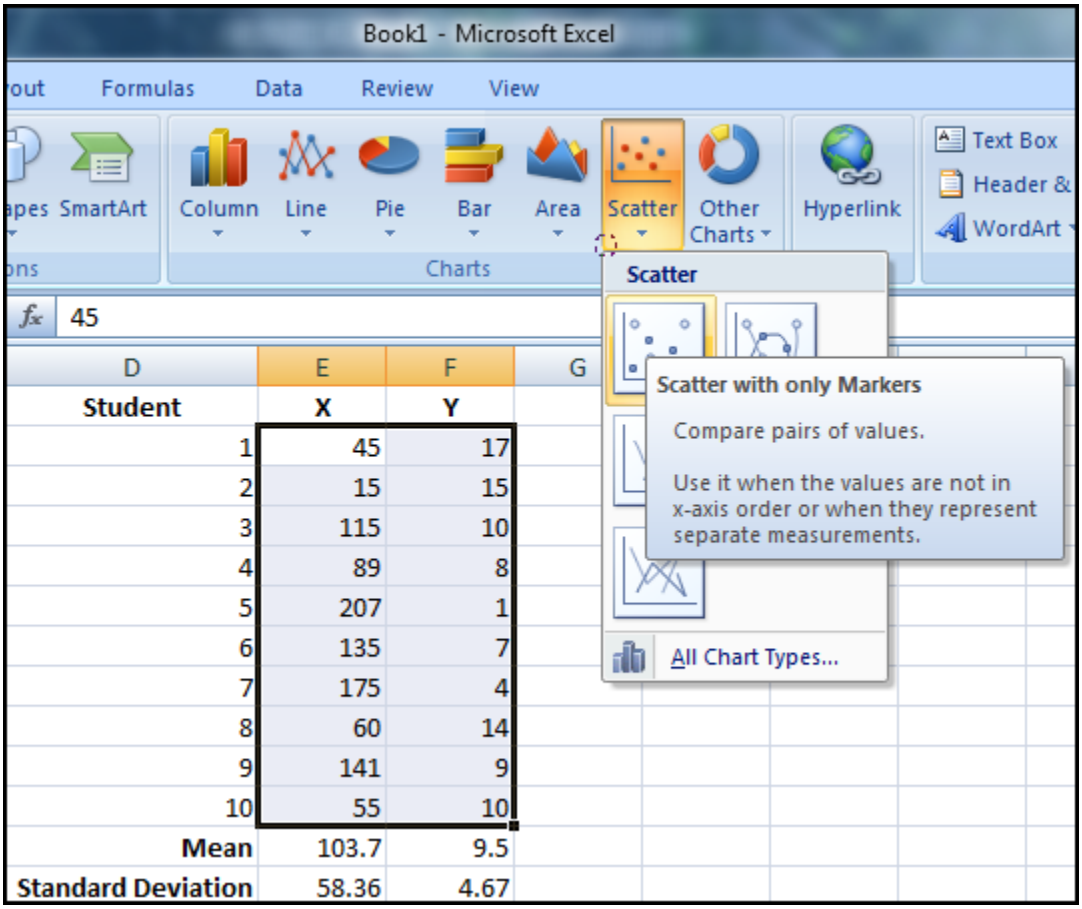
This Tip Sheet provides an overview of how to create a scatterplot for bivariate data and use it to find the equation for the linear regression line based on the least-squares criterion. We will use this equation to compute the standard error of the estimate of y from x ($S_{y \cdot x}$). To begin, input the data below (from Tip Sheet #7) and select the "Insert" tab from the options along the top of Excel.



	A	B	C	D	E	F
1				**Student**	**X**	**Y**
2				1	45	17
3				2	15	15
4				3	115	10
5				4	89	8
6				5	207	1
7				6	135	7
8				7	175	4
9				8	60	14
10				9	141	9
11				10	55	10
12				**Mean**	103.7	9.5
13				**Standard Deviation**	58.36	4.67

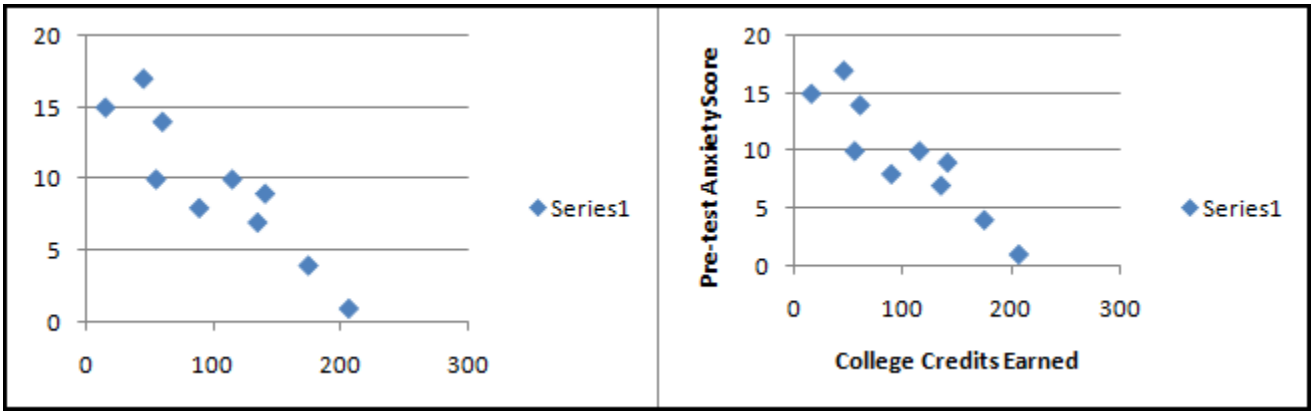
Now, use your cursor to select your data for both X and Y (you should be selecting a total of 20 cells). With the data selected, click on the "Scatter" option located in the "Charts" area. Then, select "Scatter with only Markers." This is all shown in the picture on the top of the next page.

SAMLab Tip Sheet #8

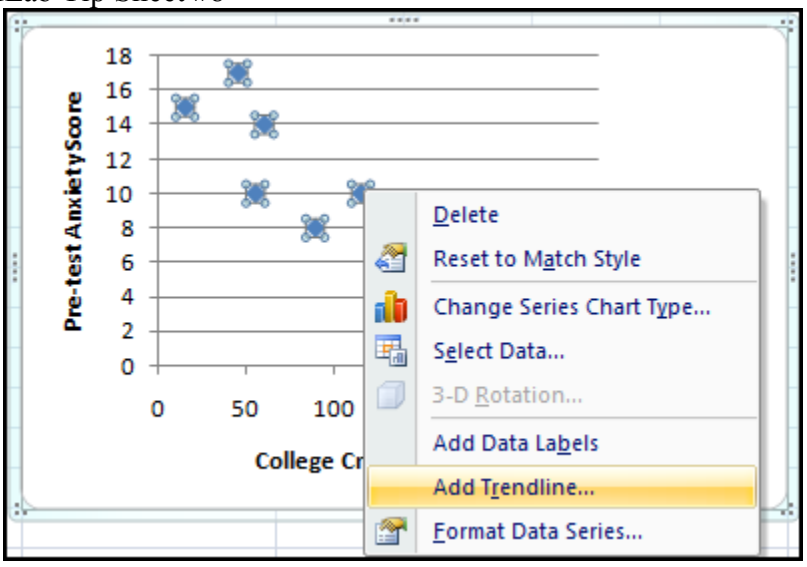


Student	X	Y
1	45	17
2	15	15
3	115	10
4	89	8
5	207	1
6	135	7
7	175	4
8	60	14
9	141	9
10	55	10
Mean	103.7	9.5
Standard Deviation	58.36	4.67

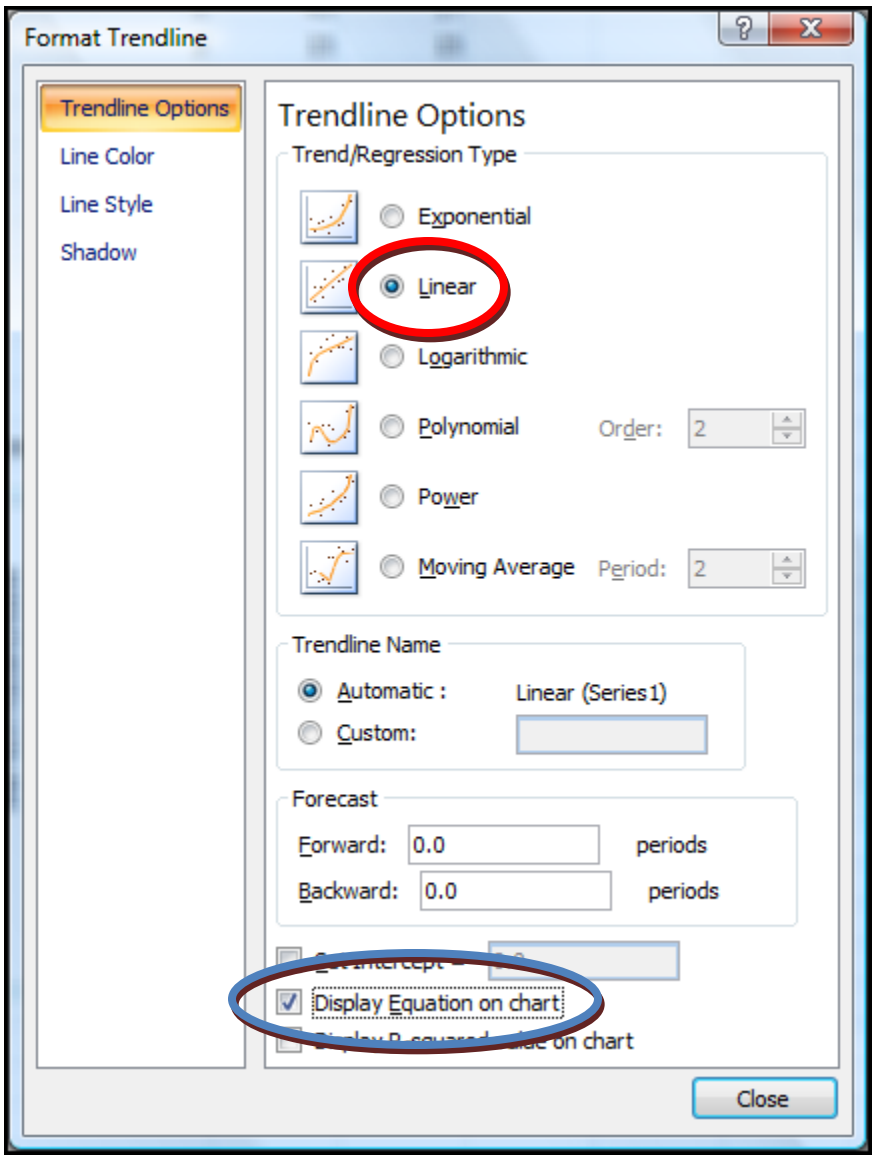
You should now have a graph that looks like the graph below on the left. To add axis titles, click on your graph and select the "Layout" tab from the options along the top of the Excel window. You can then follow the same directions outlined in Tip Sheet #5 for adding titles. The resulting graph should look like the one below on the right.



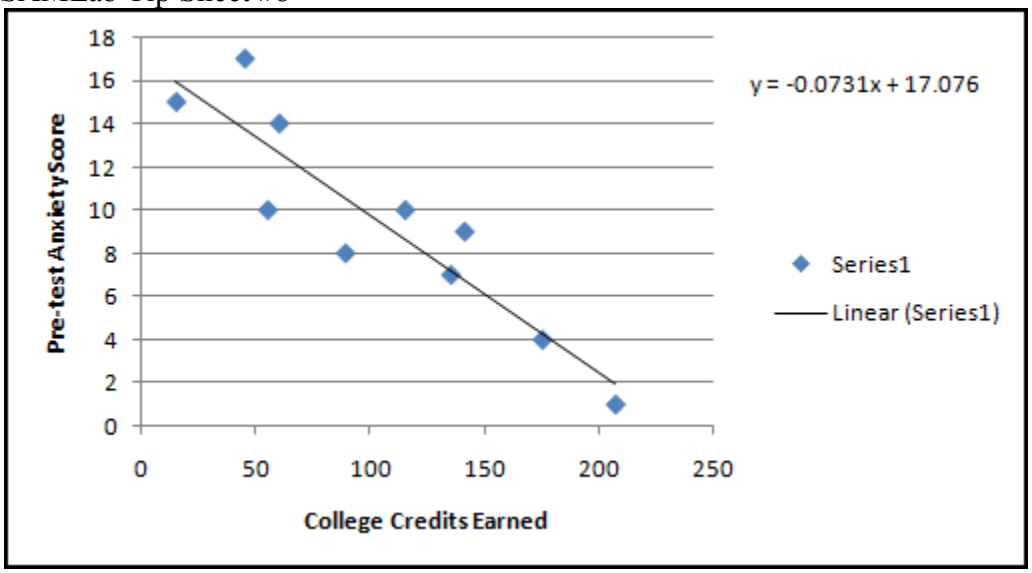
SAMLab Tip Sheet #8



Right click on one of the data points on your graph and select "Add Trendline."



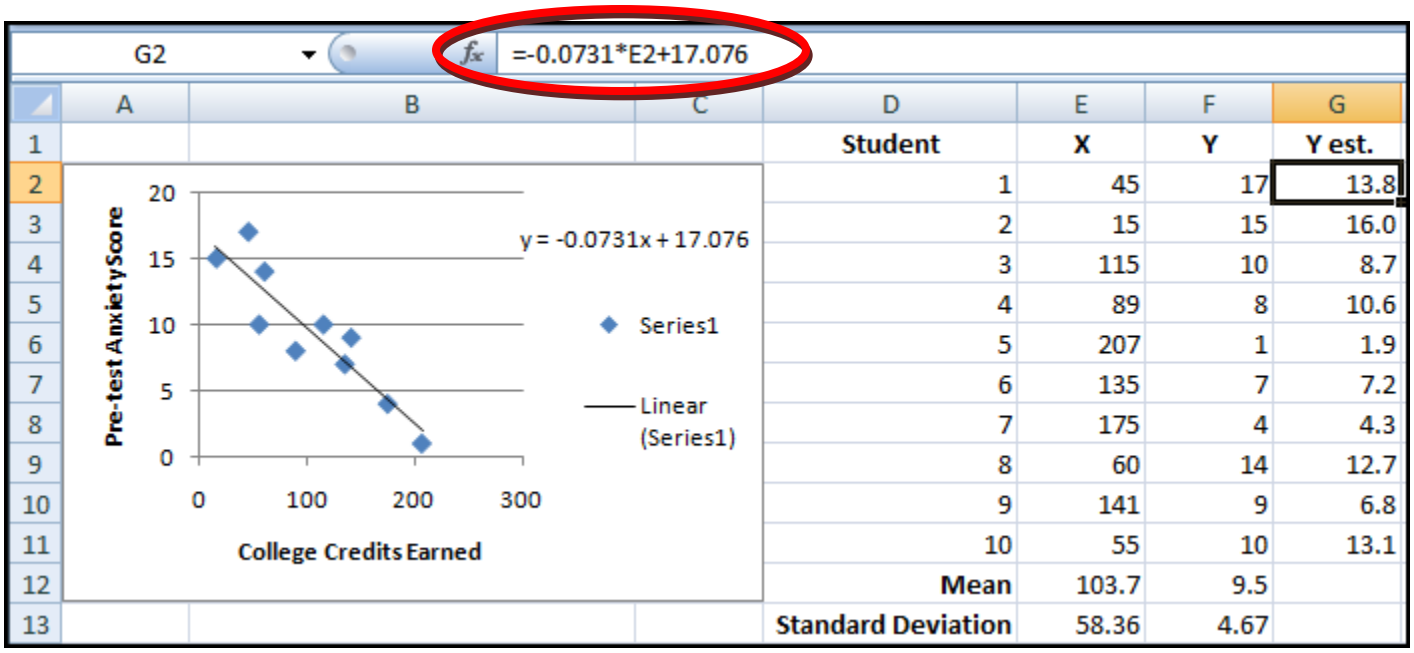
The "Format Trendline" window will open. Select "Linear" from the "Trend/Regression Type" section of the window (circled in red). Then, towards the bottom of the window, check "Display Equation on chart." You can now click on "Close" (circled in blue). The regression equation and the trendline will now appear on your graph. If the regression equation interferes with seeing the chart properly, you can click on the equation and drag it to an area that is out of the way of the graph.



Your graph should now resemble the one on the left. Next we are going to calculate the standard error of the estimate of Y from X.

Calculating the Standard Error of the Estimate of Y from X

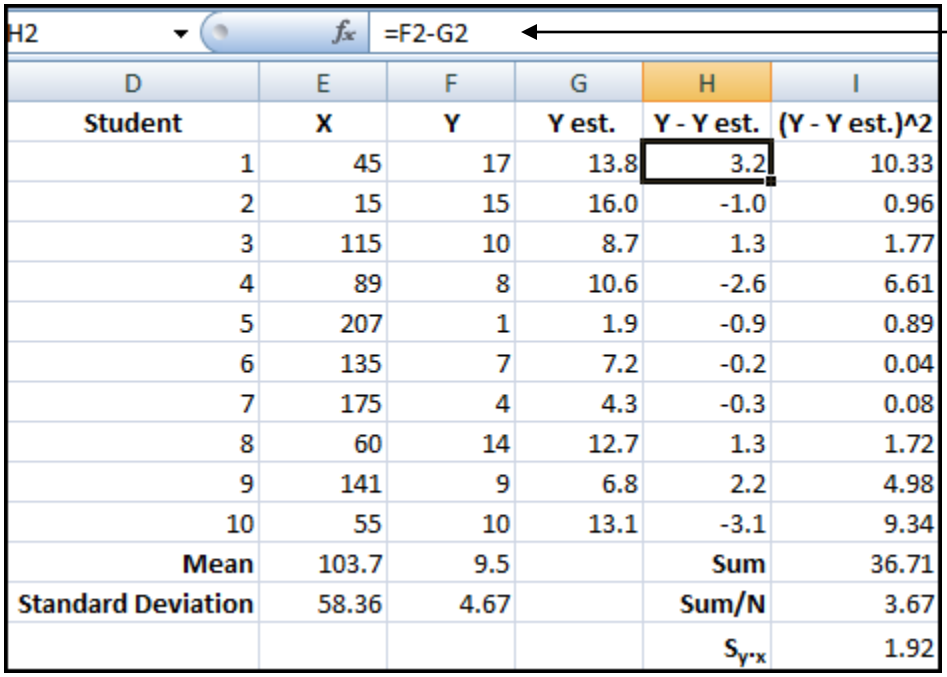
You should now see a regression equation on your graph. Label the column next to your "**Y**" scores "**Y est.**" and use the regression equation from your graph to compute the estimated y-values from the x-values. The equation is shown in the formula bar (circled).



	A	B	C	D	E	F	G
1				Student	X	Y	Y est.
2				1	45	17	13.8
3				2	15	15	16.0
4				3	115	10	8.7
5				4	89	8	10.6
6				5	207	1	1.9
7				6	135	7	7.2
8				7	175	4	4.3
9				8	60	14	12.7
10				9	141	9	6.8
11				10	55	10	13.1
12				Mean	103.7	9.5	
13				Standard Deviation	58.36	4.67	

Next, create a column for difference scores, subtracting estimated y-values from actual y-values. Create another column containing the squares of these difference scores and find their sum. Divide the sum by N (10 in this case) and take the square root to obtain $S_{y \cdot x}$. This can all be seen in the picture at the top of the next page.

SAMLab Tip Sheet #8



H2 | fx | =F2-G2

D	E	F	G	H	I
Student	**X**	**Y**	**Y est.**	**Y - Y est.**	**(Y - Y est.)^2**
1	45	17	13.8	3.2	10.33
2	15	15	16.0	-1.0	0.96
3	115	10	8.7	1.3	1.77
4	89	8	10.6	-2.6	6.61
5	207	1	1.9	-0.9	0.89
6	135	7	7.2	-0.2	0.04
7	175	4	4.3	-0.3	0.08
8	60	14	12.7	1.3	1.72
9	141	9	6.8	2.2	4.98
10	55	10	13.1	-3.1	9.34
Mean	103.7	9.5		**Sum**	36.71
Standard Deviation	58.36	4.67		**Sum/N**	3.67
				$S_{y \cdot x}$	1.92

Note that the difference scores here do not require a static cell reference. Also, Excel does have a paste function for $S_{y \cdot x,}$ but it is an inferential version, which divides the sum of squares by N-2.

The equation for the standard error of the estimate is:

$$S_{y \cdot x} = \sqrt{\frac{\sum (Y - Y')^2}{N}}$$