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-x}). To begin, input the data below (from <u>Tip Sheet</u> <u>#7</u>) and select "Chart" from the Insert menu as shown.

	Microsoft Exc	el - Book1			- 🗆 🗵	
	Eile Edit Vie	ew Insert Format To	ols <u>D</u> ata <u>W</u> indow <u>H</u> elp		-8×	The data
	D12	✓ Cells				(fictitious) in this
The descriptive	A	<u>R</u> ows	D	E	F	(includus) in uns
	1	Columns	Student	X	Y T	example
standard	2	Chart	1	45	17	represent the
deviation is	3		2	15	15	number of
used here	4	Page Break	3	115	10	college credits a
(STDEVP).	5	f. Function	4	89	8	student has
For more	6	···	5	207	1	earned (\mathbf{X}) and
information on	7	<u> </u>	6	135	7	
Information on	8		7	175	4	the student s
standard	9		8	60	14	score on a metric
deviations, see	10		9	141	9	of test anxiety
Tip Sheet #6.	11		10	55	10	(Y) administered
	12		Mean	103.7 🔺	9.5	before the final
	13		Standard Deviation	58.36	4.67	exam period.
			21) PIÈ	period.
	Ready					

In the Chart Wizard window shown below left select "XY (Scatter)" in the Chart type field. Next click on the Chart sub-type that has no lines connecting the data points. Click Next > and highlight the Data range field as shown on the lower right, making sure the Data Range tab (circled) is selected. Either select the cells containing your data (to which the arrow points) or manually type in the cell references (Tip Sheet #2). Make sure the radio button next to "Columns" (circled) is selected. Click Next > to continue.

Chart Wizard - Step 1 of 4 - Cha	art Type 🔹 👔 👔	Source Data	<u>?×</u>
Standard Types Chart type: Column Bar Column Bar Column Pie Pie XV (Scatter) Area Doughnut Radar Surface Bubble Stock V	Chart sub-type: Chart sub-type: Chart sub-type: Chart sub-type: Chart sub-type: Scatter. Compares pairs of values. Press and Hold to Yiew Sample	Data Range Series 18 • 16 • 12 • 10 • 11 • 12 • 10 • 6 • 2 • 0 50 100 150 Data range: Sheet1 !\$E\$2:\$F\$11 Series in: • • • • • • • • •	+ Series 1 + 200 250
Cancel	< Back Next Einish	Cancel < <u>B</u> ack	Next > Einish

SAMLab Tip Sheet #8

Give names to your axes as shown on the left below and click Next >. Click the radio button next to "As object in," select "Sheet 1" or wherever you want your chart, and click Finish.

Chart Wizard - Step 3 of 4 - Chart	Options		? ×
Titles Axes Gridlines Le Chart title:	Drefter from the program of the prog	vata Labels	◆ Series1
Second value (Y) axis:	0	50 100 150 200 250 College Credits Earned	
	ncel	< Back Next >	Einish

Chart Wizard	l - Step 4 of 4 - Cha	art Location	? ×
Place chart: -	C As new sheet:	Chart1	
	• As <u>o</u> bject in:	Sheet1	·
2	Cancel	< <u>B</u> ack Next > Einit	sh _R

You should now have a graph like the one on the left below. Right click on one of the data points as shown on the right and select "Add Trendline..."



When you see the screen below left, select "Linear" for Trend/Regression type. Next, click the "Options" tab and in the window shown below right check the box next to "Display equation on chart." Click OK.

Add Trendline	Add Trendline ? 🗙
Type Options Trend/Regression type Image: Comparison type Image: Logarithmic Image: Comparison type Image: Logarithmic Image: Comparison type Image: Logarithmic Image: Comparison type Image: Comparison type Image: Comparison type	Type Options Trendline name • Automatic: Linear (Series1) • Qustom: • Qustom: • Gustom: • Display equation on chart • Spelay R-squared value on chart
OK Cancel	OK Cancel

SAMLab Tip Sheet #8

Calculating the Standard Error of the Estimate of Y from X

You should now see a regression equation on the your graph. If you cannot read it easily, click and drag it to a better place. 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.

副 File Edit View Insert Format Iools Data Window E	<u>l</u> elp <u>M</u> acro	<u>F</u> ilter		_8
G2 = =-0.0731*E2+17.076]				
B C D Formul	a Bar	F	G	H
1 Student	X	Y	Y est.	
2 y = -0.0731x + 17.076	45	17	13.8	
3]	15	15	16.0	
4 Series1	115	10	8.7	
5	89	8	10.6	
	207	1	1.9	
7 (Series1)	135	7	7.2	
	175	4	4.3	
	60	14	12.7	
0 0 100 200 300	141	9	6.8	
1 College Credits Earned	55	10	13.1	
2	103.7	9.5		
3 Standard Deviation	58.36	4.67		
4				
Sheet1 / Sheet2 / Sheet3 /			1	E
A Sheet1 / Sheet2 / Sheet3 /				

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-x} .

ne equation for the andard error of the timate is:

$$y \bullet x = \sqrt{\frac{\sum (Y - Y')}{N}}$$

D	E Formu	ula Bar	G	H	Î
1 Student	x	Y	Y est.	Y - Y est.	(Y - Y est.) ²
2 y = -0.0731x +	45	17	13.8	3.2	10.33
3 17. Series1	15	15	16.0	-1.0	0.98
4	115	10	8.7	1.3	1.77
5	89	8	10.6	-2.6	6.61
5 Linear	207	1	1.9	-0.9	0.89
7 (Series1)	135	7	7.2	-0.2	0.04
3	175	4	4.3	-0.3	0.08
3 400	60	14	12.7	1.3	1.72
0 its	141	9	6.8	2.2	4.98
1	55	10	13.1	-3.1	9.34
2	103.7	9.5		Sum	36.71
3 Standard Deviation	58.36	4.67		Sum/N	3.67
4				S	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.