

Generate a Regression Equation

Excel Step-by-Step How-to for Windows and Excel for Mac 2016 (v.16) or later

Excel for Mac 2011 (v.14) Instructions on [page 12](#)

Instructions: Use this guide to generate a regression equation using Excel. Three methods are described: creating a scatter plot, using the built-in Excel functions, and running a regression report using statistics tools.

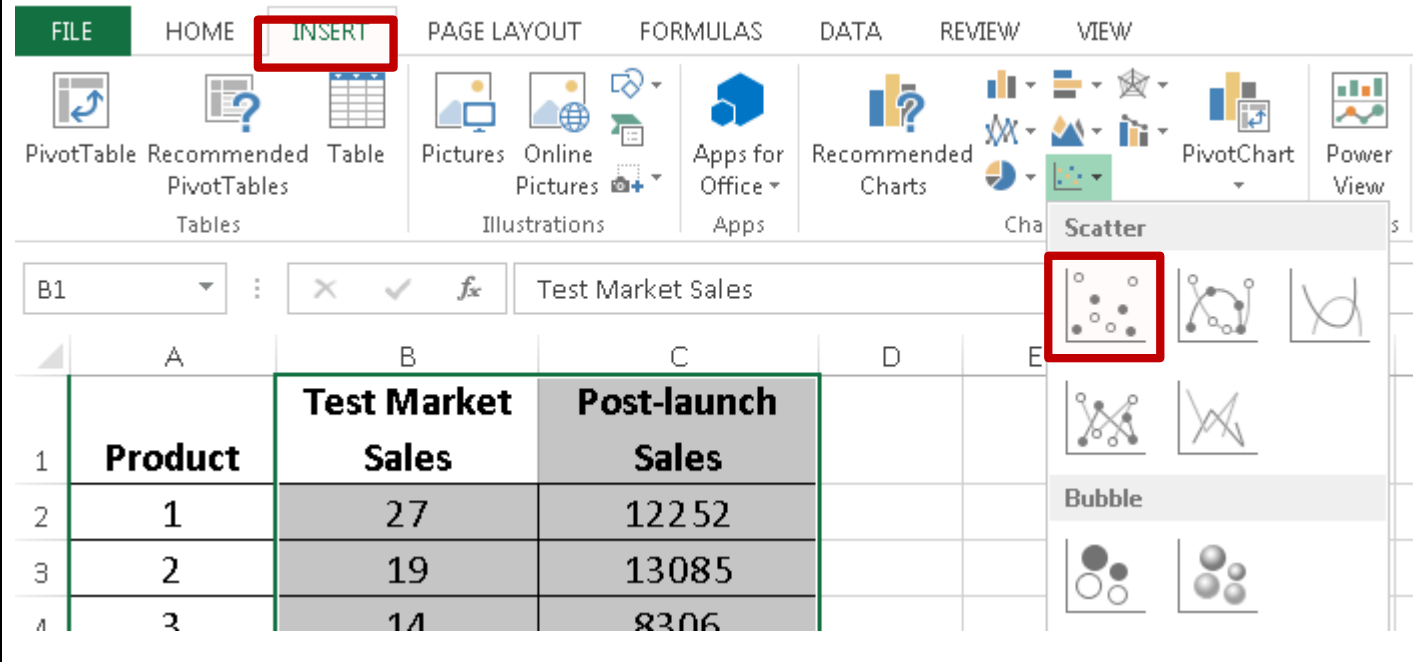
Data requirement: Two variables (one independent, one dependent), quantitative data

Sample data: Test sales and post-launch sales

Step	Windows Instructions + Screen Shot																								
Scattler Plot Method																									
1. Arrange the data into two adjacent columns so that the independent variable is in the left column and the dependent variable is in the right column.	<table><tr><td></td><td>A</td><td>B</td><td>C</td></tr><tr><td></td><td>Product</td><td>Test Market Sales</td><td>Post-launch Sales</td></tr><tr><td>1</td><td>1</td><td>27</td><td>12252</td></tr><tr><td>2</td><td>2</td><td>19</td><td>13085</td></tr><tr><td>3</td><td>3</td><td>14</td><td>8306</td></tr><tr><td>4</td><td>4</td><td>7</td><td>4081</td></tr></table> <p>In this example, the data are in columns B and C. The values in column B correspond to the X-axis (independent variable) and the column C data correspond to the Y-axis (dependent variable).</p> <p><i>Note:</i> Data for your dependent variable should always be the in right column.</p>		A	B	C		Product	Test Market Sales	Post-launch Sales	1	1	27	12252	2	2	19	13085	3	3	14	8306	4	4	7	4081
	A	B	C																						
	Product	Test Market Sales	Post-launch Sales																						
1	1	27	12252																						
2	2	19	13085																						
3	3	14	8306																						
4	4	7	4081																						

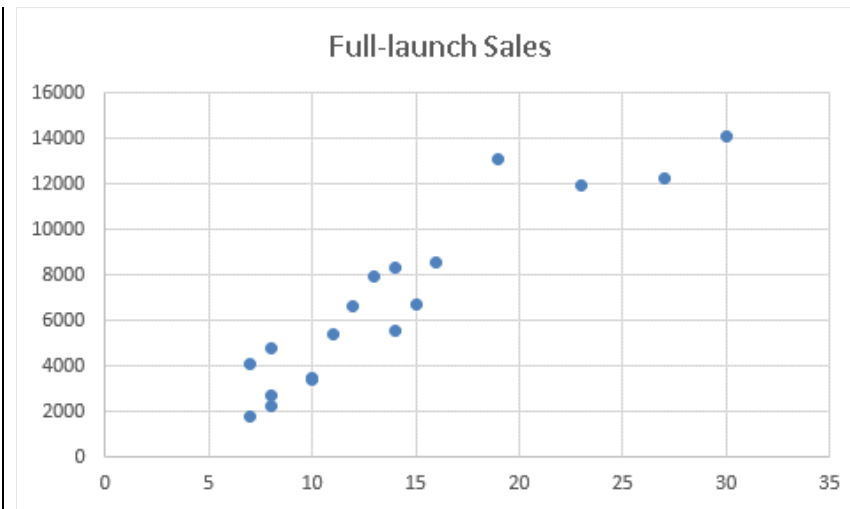
2. Select your data and insert a Scatter (X,Y) chart.

After selecting the data in both columns, including the header rows, select the Insert tab. Navigate to the Scatter (X,Y) chart icon and select it. Excel always uses the first column for the x-axis and the second column for the y-axis. If your dependent variable data is in the right column, it will always correspond to the y-axis.



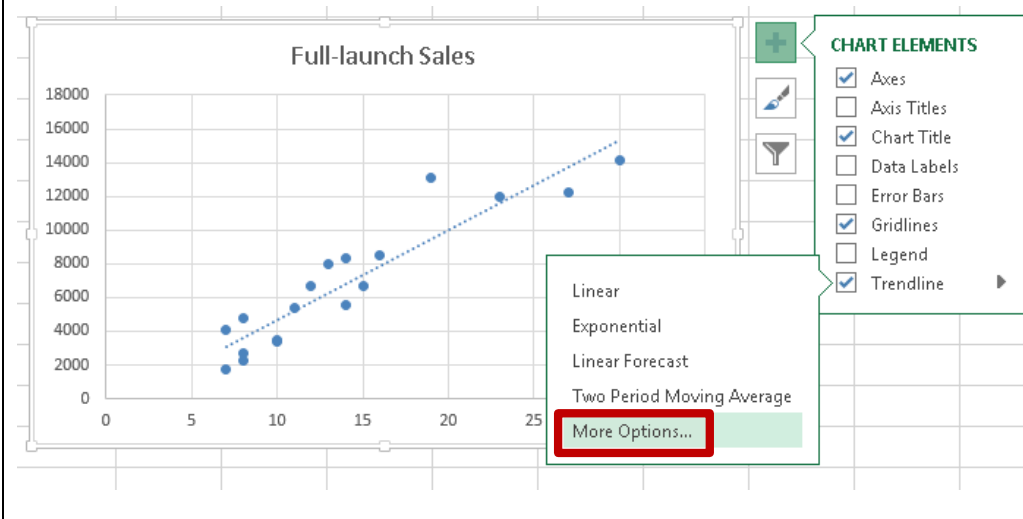
The screenshot shows the Microsoft Excel interface. The 'INSERT' tab is selected and highlighted with a red box. In the 'Charts' group, the 'Scatter' chart icon is highlighted with a red box. Below the ribbon, the formula bar shows 'Test Market Sales'. The worksheet contains the following data:

	A	B	C	D	E
1	Product	Test Market Sales	Post-launch Sales		
2	1	27	12252		
3	2	19	13085		
4	3	14	8306		



3. Display the best fit line on the scatter plot.

Select (right click) any point in the scatter plot to bring up the Chart Elements menu. At the bottom of this menu, select Trendline. After adding the trendline, select More Options.






4. Display the equation and R-squared value of the trendline on the chart.


Navigate to the Format Trendline pane, and select Display Equation on chart and Display R-squared value on chart.


Format Trendline ✕


TRENDLINE OPTIONS ▾


  


▲ TRENDLINE OPTIONS


 ☐ Exponential

 ☒ Linear

 ☐ Logarithmic

 ☐ Polynomial Order

 ☐ Power

 ☐ Moving Average Period

Trendline Name

☒ Automatic Linear (Post-launch Sales)

☐ Custom

Forecast

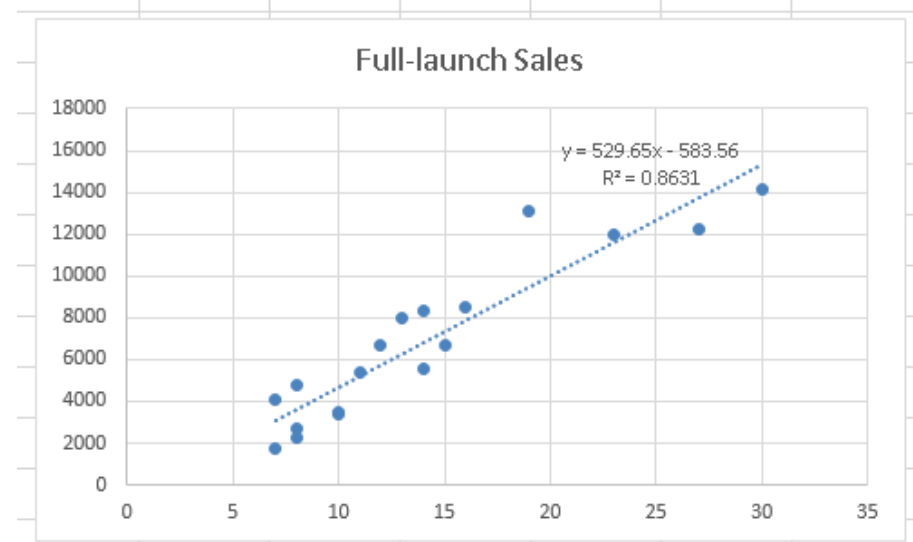
Forward periods

Backward periods

☐ Set Intercept

☒ Display Equation on chart

☒ Display R-squared value on chart



Excel Function Method

1. Label area to calculate elements of regression equation.

Label cells for the slope, intercept, and the R-squared value in the sheet.

Test Market Sales	Post-launch Sales			
27	12252			
19	13085		SLOPE	
14	8306		INTERCEPT	
7	4081		RSQ	
8	4807			

2. Use the SLOPE function to calculate the slope.

=SLOPE(known y's, known x's)

Next to the cell labeled SLOPE, enter the SLOPE function into the function box. Select the column of dependent data for the first parameter of the function and select the column of independent data for the second parameter. For this example, the known y's are populated in the cell range C2:C19, so that is the range for that parameter of the formula.

<div> ✕ ✓ <i>f_x</i> <input type="text" value="=SLOPE(C2:C19,B2:B19)"/> </div>				
B	C	D	E	F
Test Market Sales	Post-launch Sales			
27	12252			
19	13085		SLOPE	529.651
14	8306		INTERCEPT	
7	4081		RSQ	
8	4807			



TIP: To select a column of data




- Use the INTERCEPT formula to calculate the intercept.

TIP: When identifying a data range, to select a block of populated cells in a column of data you can first select the top row in that block. The cell will be surrounded by a highlight. Then double-click in the solid square in the lower right corner of that highlight.

Another method to efficiently collect a column of data is to highlight the top cell of the column that you want and press CTRL + Shift + ↓. This will select everything between your highlighted cell and that first blank row.

=INTERCEPT(*known y's, known x's*)

Type the INTERCEPT formula into the formula box. Select the column of dependent data for the first parameter of the formula and select the column of independent data for the second parameter the same way as for the SLOPE formula.

  		=INTERCEPT(C2:C19,B2:B19)			
B	C	D	E	F	G
Test Market Sales	Post-launch Sales				
27	12252				
19	13085				
14	8306		SLOPE	529.651	
7	4081		INTERCEPT	-583.559	
8	4807		RSQ		

4. Use the RSQ formula to calculate the R-squared value.

=RSQ(*known y's, known x's*)

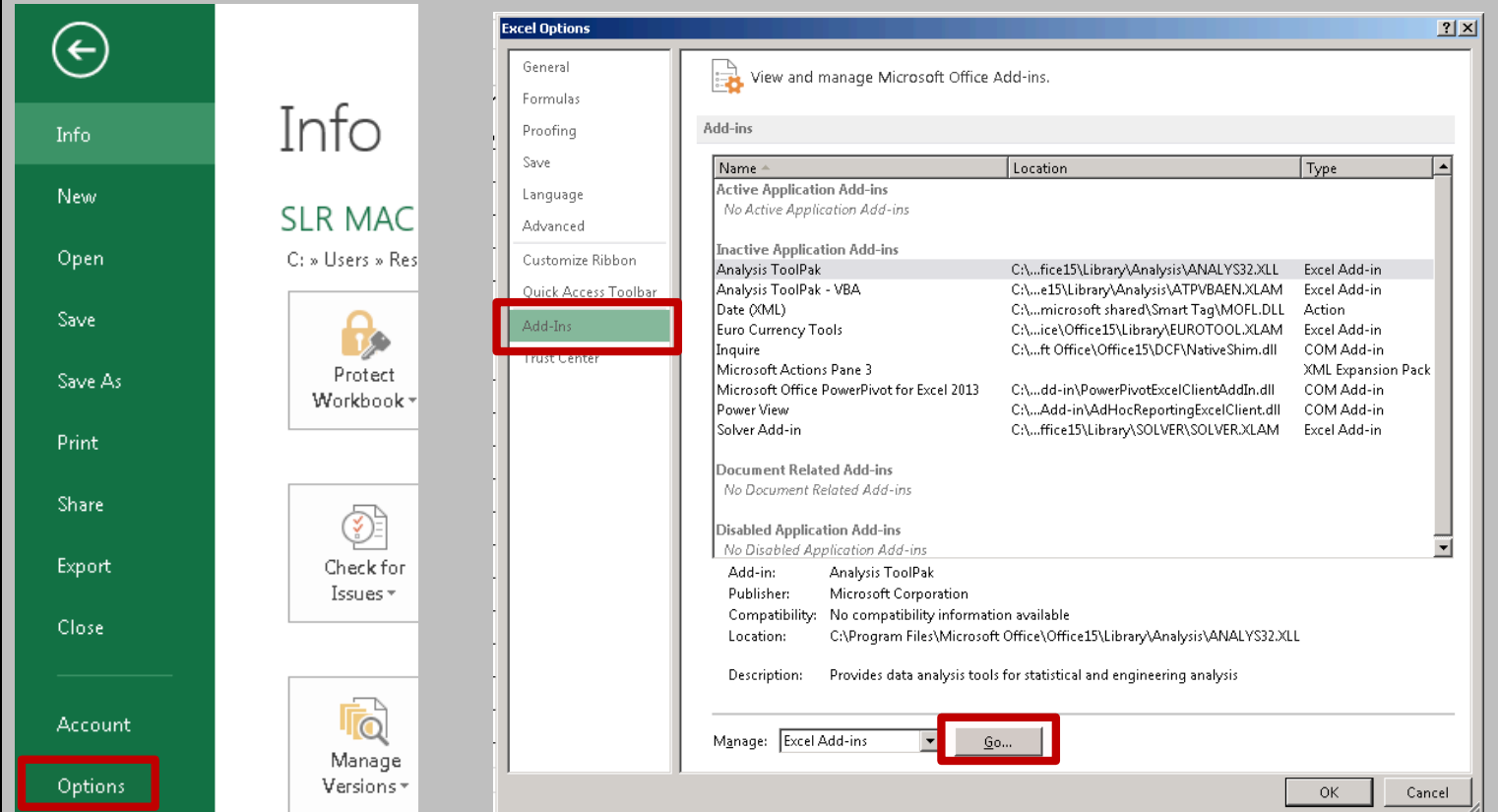
Type the RSQ formula into the formula box. Select the column of dependent data for the first parameter of the formula and select the column of independent data for the second parameter the same way as for the SLOPE and INTERCEPT formula.

✕ ✓ <i>f_x</i>		=RSQ(C2:C19,B2:B19)			
B	C	D	E	F	G
Test Market Sales	Post-launch Sales				
27	12252				
19	13085		SLOPE	529.651	
14	8306		INTERCEPT	-583.559	
7	4081		RSQ	0.863092	
8	4807				

Data Analysis Method

1. Set up a third-party data analysis tool

Navigate to File and select Options. In the Excel Options pop-up window, navigate to Add-Ins and click Go.

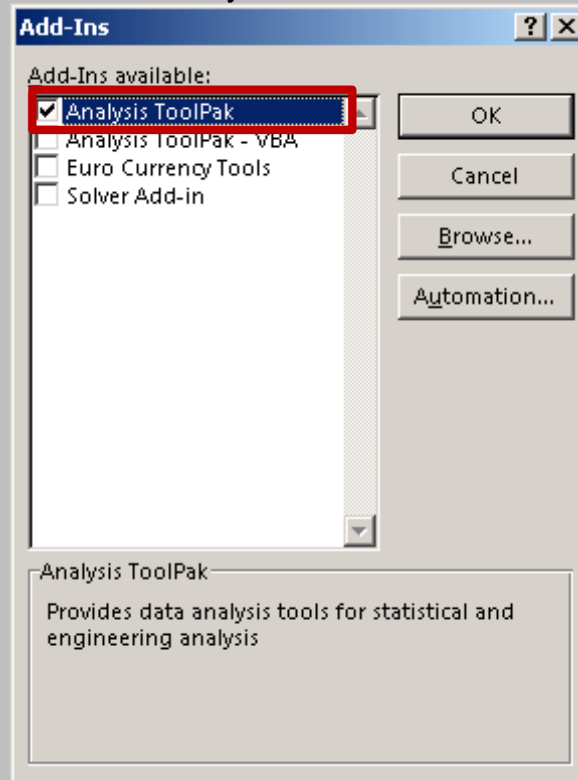


The screenshot shows the Excel Options dialog box with the 'Add-Ins' tab selected. The 'Go...' button at the bottom of the Add-Ins section is highlighted with a red box. The 'Options' button in the Excel ribbon is highlighted with a red box.

Name	Location	Type
Active Application Add-ins		
No Active Application Add-ins		
Inactive Application Add-ins		
Analysis ToolPak	C:\...fice15\Library\Analysis\ANALYS32.XLL	Excel Add-in
Analysis ToolPak - VBA	C:\...e15\Library\Analysis\ATPVBAEN.XLAM	Excel Add-in
Date (XML)	C:\...icrosoft shared\Smart Tag\MOFL.DLL	Action
Euro Currency Tools	C:\...ice\Office15\Library\EUROTOOL.XLAM	Excel Add-in
Inquire	C:\...ft Office\Office15\DCF\NativeShim.dll	COM Add-in
Microsoft Actions Pane 3		XML Expansion Pack
Microsoft Office PowerPivot for Excel 2013	C:\...dd-in\PowerPivotExcelClientAddIn.dll	COM Add-in
Power View	C:\...Add-in\AdHocReportingExcelClient.dll	COM Add-in
Solver Add-in	C:\...fice15\Library\SOLVER\SOLVER.XLAM	Excel Add-in
Document Related Add-ins		
No Document Related Add-ins		
Disabled Application Add-ins		
No Disabled Application Add-ins		
Add-in:	Analysis ToolPak	
Publisher:	Microsoft Corporation	
Compatibility:	No compatibility information available	
Location:	C:\Program Files\Microsoft Office\Office15\Library\Analysis\ANALYS32.XLL	
Description:	Provides data analysis tools for statistical and engineering analysis	

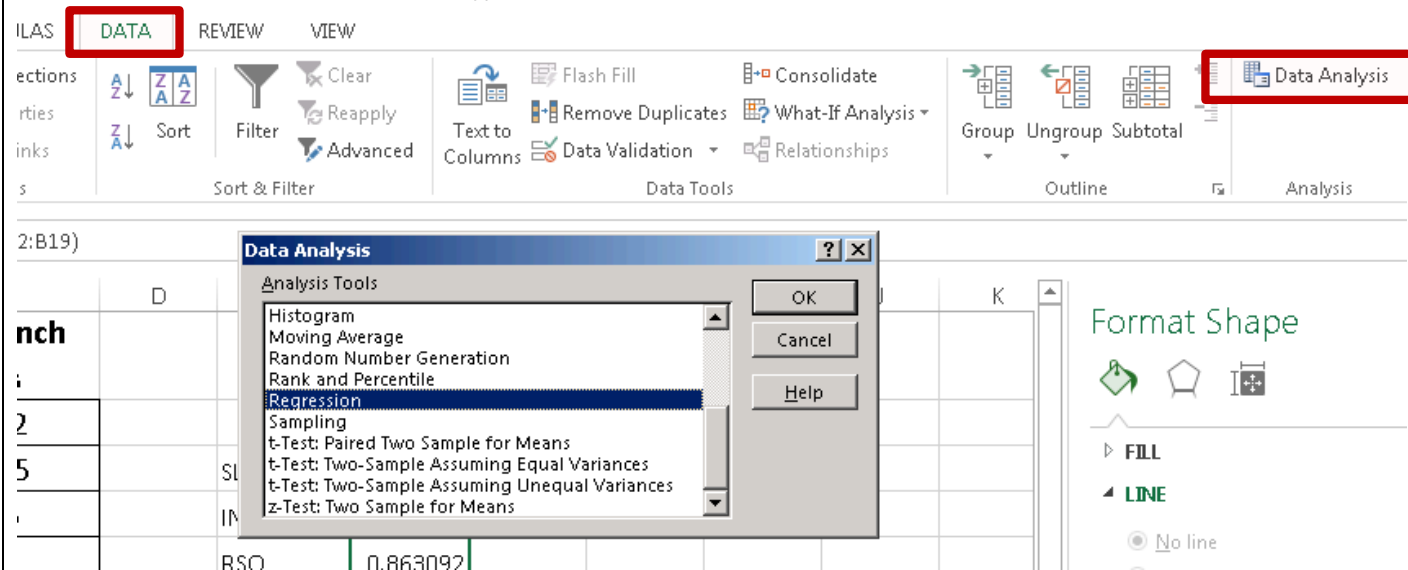
Manage: Excel Add-ins **Go...**

Select the Analysis ToolPak and click OK.



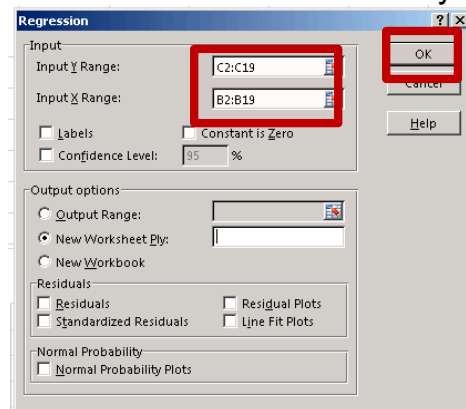
1. Open the Data Analysis tool.

Navigate to the Data ribbon and select the Data Analysis icon on the right. Select Regression and click OK.



2. Enter data.

Enter the range that holds the dependent data in the Y field and enter the range that hold the independent data in the X field. Click OK. By default, Excel creates a new worksheet when a regression is run.



3. View the regression table.

The R-squared value is found under Regression Statistics. The intercept and slope can be found in the bottom table.

SUMMARY OUTPUT																			
Regression Statistics																			
Multiple R	0.929028																		
R Square	0.863092																		
Adjusted R Square	0.854536																		
Standard Error	1480.396																		
Observations	18																		
ANOVA																			
	df	SS	MS	F	Significance F														
Regression	1	2.21E+08	2.21E+08	100.8672	2.59E-08														
Residual	16	35065176	2191574																
Total	17	2.56E+08																	
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%											
Intercept	-583.559	816.6188	-0.7146	0.485151	-2314.71	1147.596	-2314.71	1147.596											
X Variable 1	529.651	52.73694	10.04326	2.59E-08	417.8537	641.4483	417.8537	641.4483											

Create a Regression Equation

Excel Step-by-Step How-to for Mac Excel 2011 (v.14) or earlier

Excel for Windows and Excel for Mac 2016 (v.16) Instructions on [page 1](#)

Instructions: Use this guide to generate a regression equation using Excel. Three methods are described: creating a scatter plot, use the built-in Excel functions, and by running a regression report using statistics tools.

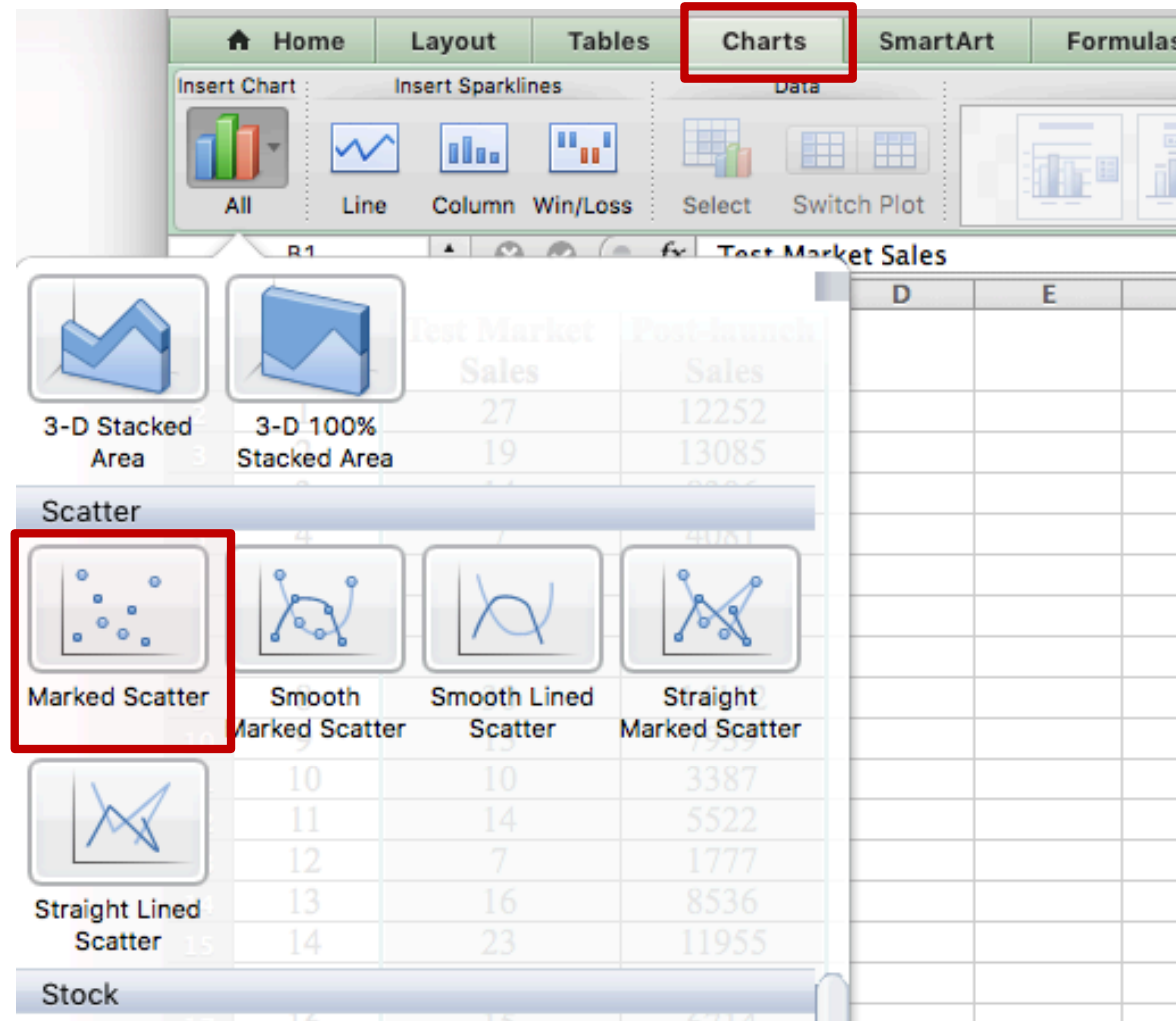
Data requirement: Two variables, quantitative data

Sample Data: Test sales and post-launch sales

Step	Windows Instructions + Screen Shot			
Scattler Plot Method				
1. Arrange the data into two adjacent columns so that the independent variable is in the left column and the dependent variable is in the right column.		A	B	C
			Test Market	Post-launch
	1	Product	Sales	Sales
	2	1	27	12252
	3	2	19	13085
	4	3	14	8306
	5	4	7	4081
	6	5	8	4807
	7	6	12	6666
	8	7	11	5374
	9	8	30	14112
In this example, the data are in columns B and C. The values in column B correspond to the X-axis (independent variable) and the column C data correspond to the Y-axis (dependent variable). <i>Note:</i> Data for your dependent variable should always be the in right column.				

2. Select your data and insert a Marked Scatter chart.

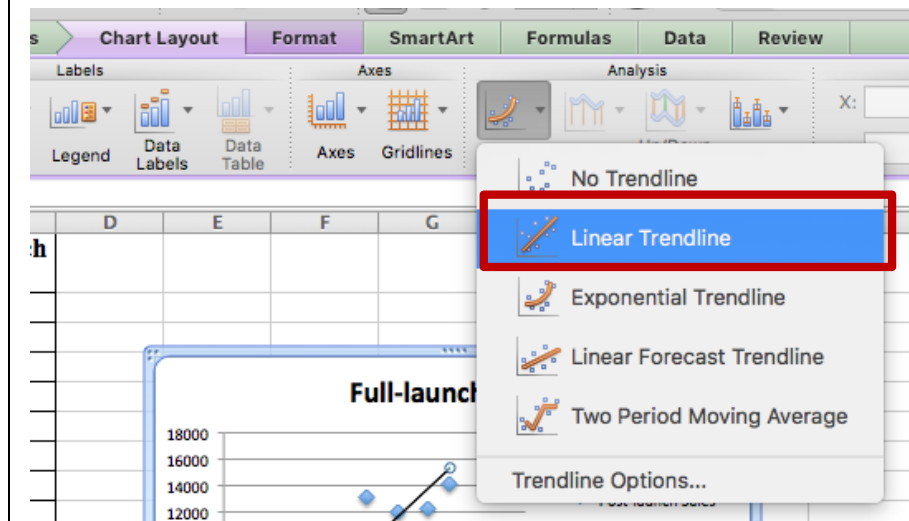
After highlighting the data in both columns, including the header rows, select the insert tab. Navigate to the Charts and select a Marked Scatter chart.





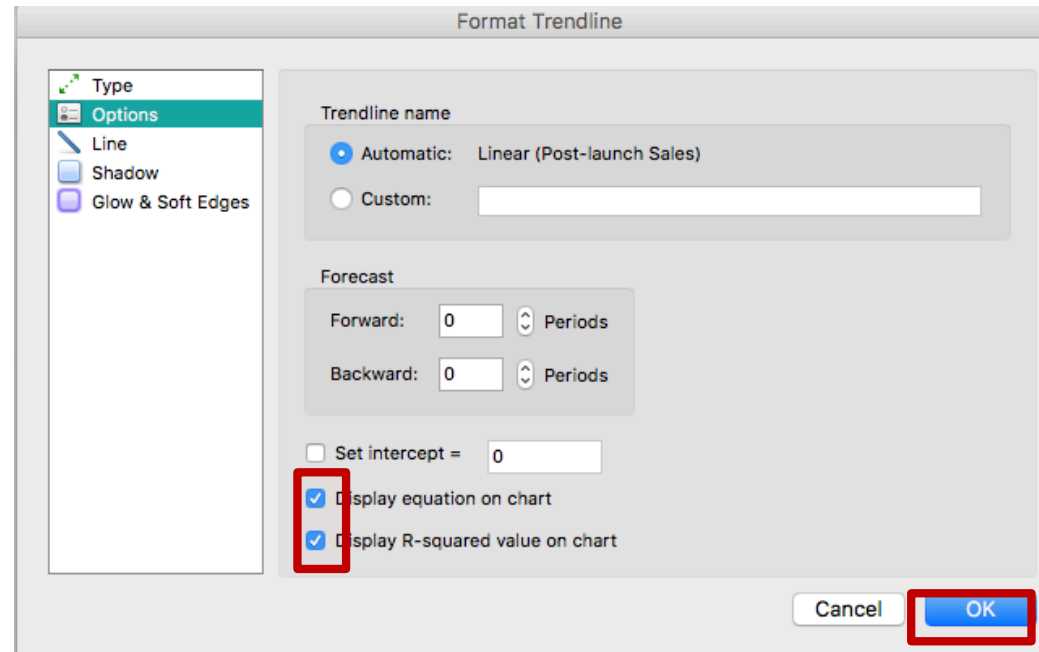
3. Display the best fit line on the scatter plot.

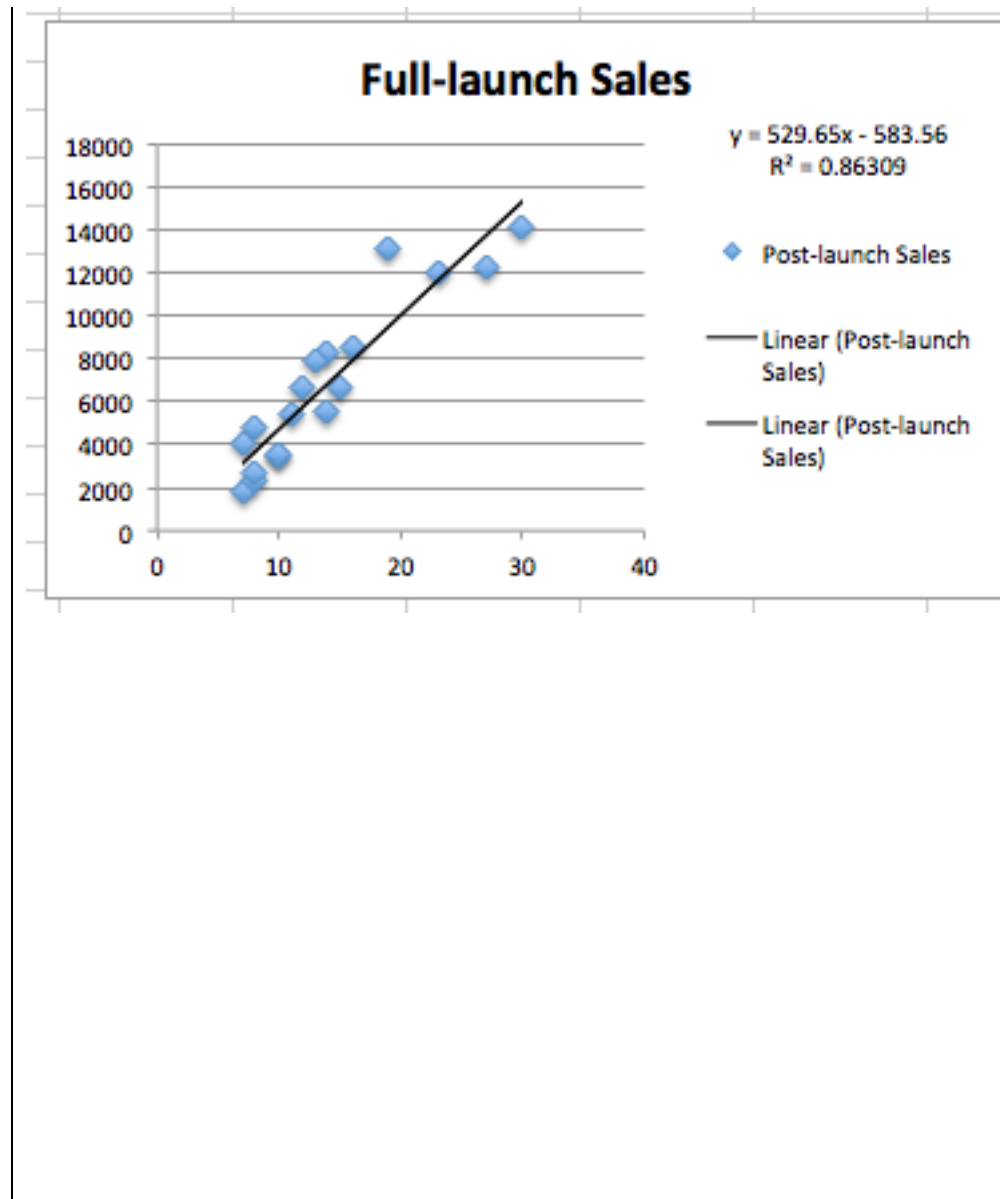
Navigate to the Chart Layout menu and add a Linear Trendline. After adding the trendline select More Options at the bottom of the menu.



4. Display the equation and R-squared value of the trendline on the chart.

Navigate to the Options tab. Select Display equation on chart and Display R-squared value on chart. Click OK.





Excel Function Method

1. Label area to calculate elements of regression equation.

Labels cells for the slope, intercept, and the R-squared value in the sheet.

	A	B	C	D	E	F
1	Product	Test Market Sales	Post-launch Sales			
2	1	27	12252			
3	2	19	13085		SLOPE	
4	3	14	8306		INTERCEPT	
5	4	7	4081		RSQ	
6	5	8	4807			
7	6	12	6666			
8	7	11	5374			
9	8	30	14112			
10	9	13	7939			
11	10	10	2287			

2. Use the SLOPE function to calculate the slope.

=SLOPE(known y's, known x's)

Type the SLOPE function into the function box. Select the column of dependent data for the first parameter of the function and select the column of independent data for the second parameter. For this example, the known y's are populated in the cell range C2:C19, so that is the range for that parameter of the formula.

F2		fx		=SLOPE(\$C\$2:\$C\$19,\$B\$2:\$B\$19)		
	A	B	C	D	E	F
1		Test Sales	Post-launch Sales			
2		27	12252		SLOPE	529.651
3		19	13085		INTERCEPT	
4		14	8306		RSQ	
5		7	4081			
6		8	4807			
7		12	6666			
8		11	5374			
9		30	14112			



TIP: To select a column of data

3. Use the INTERCEPT formula to calculate the intercept.

TIP: To select a block of populated cells in a column of data, select the top row in that block. The cell will be surrounded by a highlight. Now double-click in the solid square in the lower right corner of that highlight.

Another method to efficiently collect a column of data is to highlight the top cell of the column that you want and press CTRL + Shift + ↓. This will select everything between your highlighted cell and that first blank row.

=INTERCEPT(*known y's, known x's*)

Type the INTERCEPT formula into the formula box. Select the column of dependent data for the first parameter of the formula and select the column of independent data for the second parameter the same way as for the SLOPE formula.

F3		fx		=INTERCEPT(\$C\$2:\$C\$19,\$B\$2:\$B\$19)			
	A	B	C	D	E	F	G
1		Test Sales	Post-launch Sales				
2		27	12252		SLOPE	529.651	
3		19	13085		INTERCEPT	-583.5591	
4		14	8306		RSQ		
5		7	4081				
6		8	4807				
7		12	6666				
8		11	5374				
9		30	14112				

4. Use the RSQ formula to calculate the R-squared value.

=RSQ(*known y's, known x's*)

Type the RSQ formula into the formula box. Select the column of dependent data for the first parameter of the formula and select the column of independent data for the second parameter the same way as for the SLOPE and INTERCEPT formula.

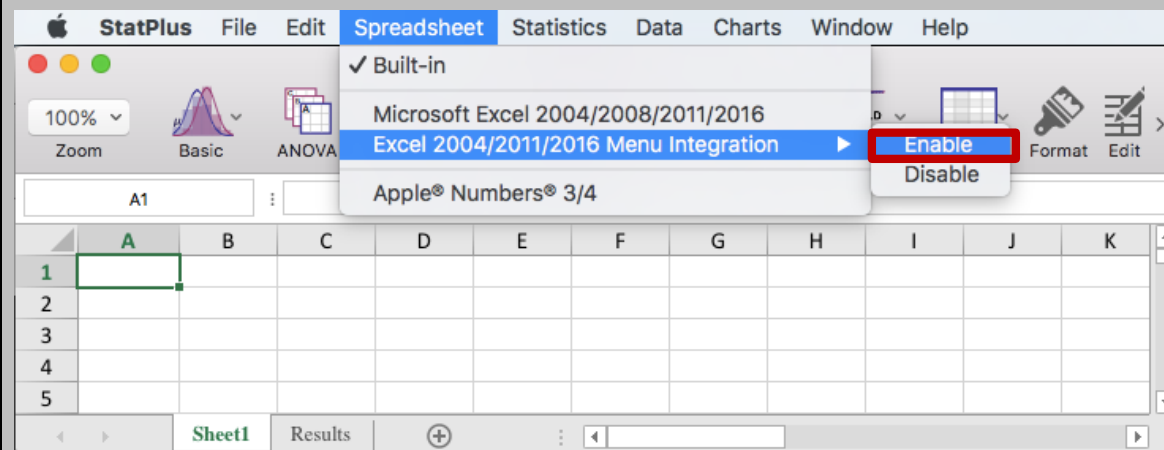
F4		=RSQ(\$C\$2:\$C\$19,\$B\$2:\$B\$19)					
	A	B	C	D	E	F	G
1		Test Sales	Post-launch Sales				
2		27	12252		SLOPE	529.651	
3		19	13085		INTERCEPT	-583.559	
4		14	8306		RSQ	0.863092	
5		7	4081				
6		8	4807				
7		12	6666				
8		11	5374				
9		30	14112				

Data Analysis Method

Data Analysis Tool
Setup: Download
StatPlus:mac to find the
regression using data
analysis.

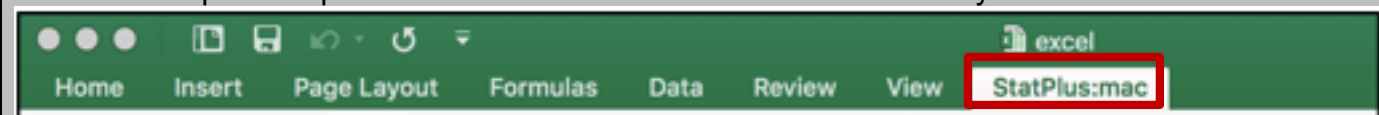
Excel for Mac 2011 and earlier does not have a built-in tool pack like Windows and later Mac versions, so you'll have to download an alternative statistics package to use the data analysis method.

One common free package is StatPlus:mac which can be found at <https://www.analystsoft.com/en/products/statplasmacle/> Use the free download button on this webpage to begin installing the tool. Once installed, you will want to integrate the tool into you Excel menu. Open StatsPlus, and in the **Spreadsheet** menu, select Excel Menu Integration / Enable,



Note: This integration requires Administrator access, and you may be prompted for Administrator username and password.

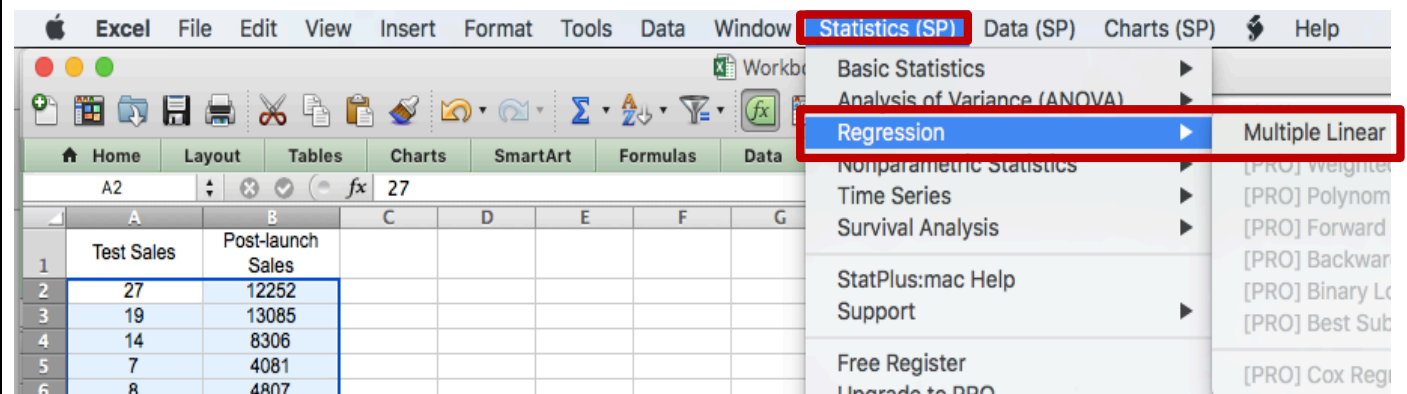
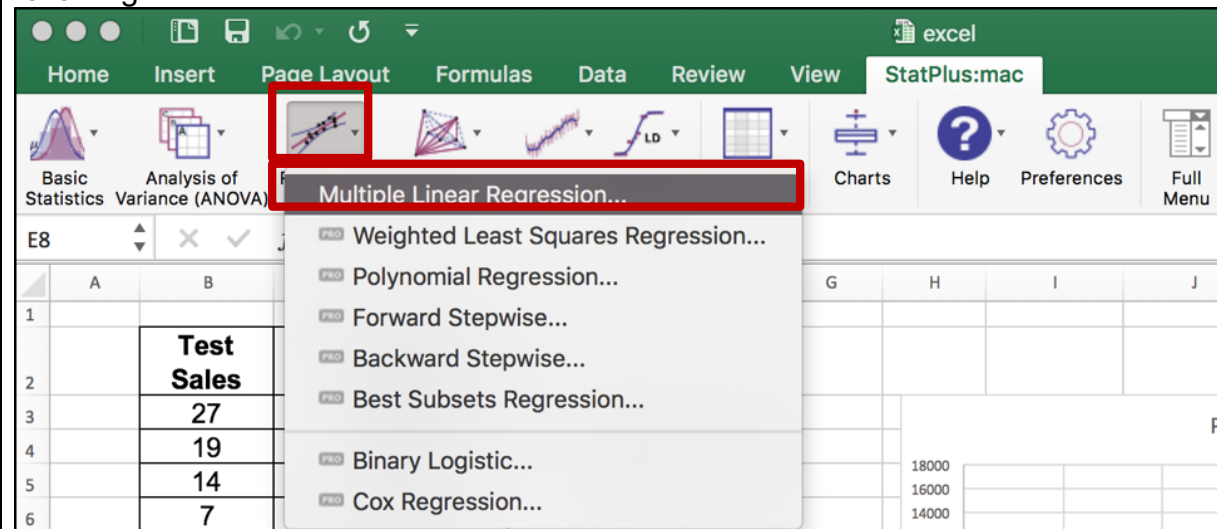
Once this step is complete a new StatPlus:mac tab will be added to your menu bar.



1. Use the Data Analysis tool to insert a multiple linear regression.

Open the data analysis tool by accessing the **StatsPlus:mac** menu and selecting Regression/Multiple Linear Regression.

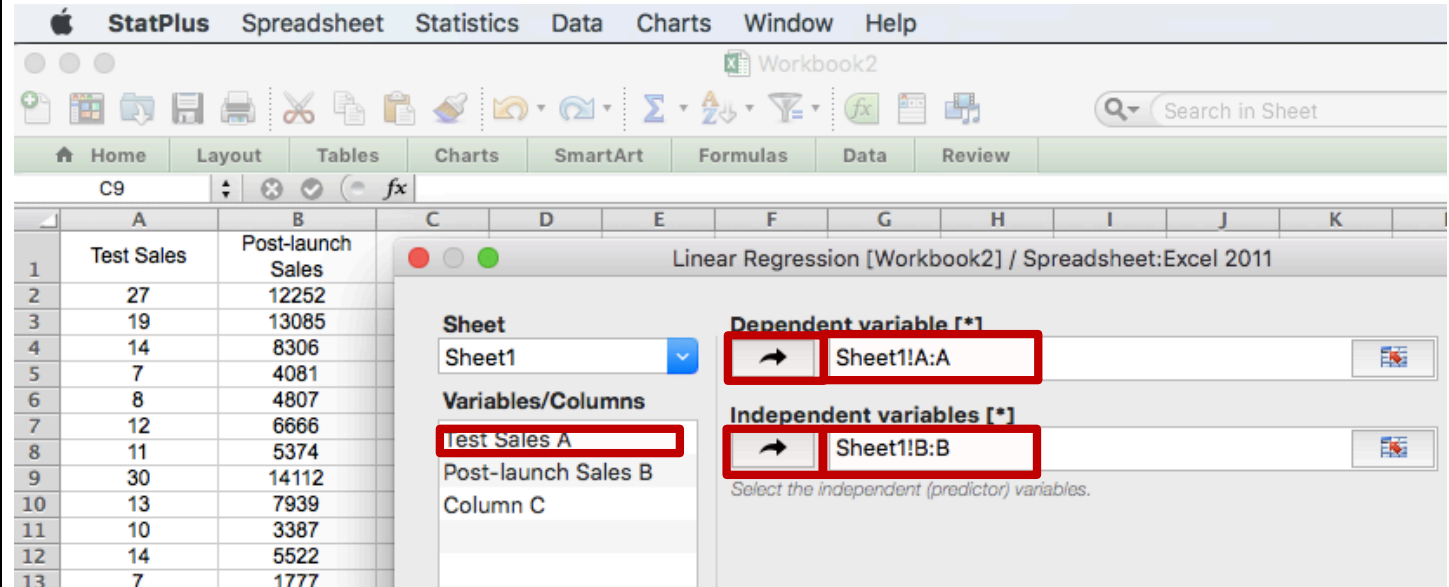
Note: Depending on details of your StatsPlus integration, you may see one or the other of the following:



2. Enter data.

Select the dependent variable in the Variables/Columns box, then click the arrow to paste the data range in the Dependent variable field. Verify that the range selected corresponds to the dependent variable.

Repeat this process with the independent variable to paste the range in the Independent variables field.



The screenshot shows the StatPlus Spreadsheet interface. A linear regression window is open, titled 'Linear Regression [Workbook2] / Spreadsheet:Excel 2011'. The 'Sheet' dropdown is set to 'Sheet1'. The 'Variables/Columns' list contains 'Test Sales A', 'Post-launch Sales B', and 'Column C'. The 'Dependent variable [*]' field is set to 'Sheet1!A:A' and the 'Independent variables [*]' field is set to 'Sheet1!B:B'. Both fields are highlighted with red boxes. The background spreadsheet shows data for 'Test Sales' and 'Post-launch Sales'.

	A	B
1	Test Sales	Post-launch Sales
2	27	12252
3	19	13085
4	14	8306
5	7	4081
6	8	4807
7	12	6666
8	11	5374
9	30	14112
10	13	7939
11	10	3387
12	14	5522
13	7	1777

3. View the regression table.

The R-squared value is found under Regression Statistics. The intercept and slope can be found in the bottom table.

	A	B	C	D	E	F	G	H	I	J
1	Linear Regression									
2										
3	Regression Statistics									
4										
5	R-Squared		0.86309							
6	Adjusted R-Squared		0.8545							
7	S		2.59667							
8	MSE		107.88317							
9	RMSE		10.38668							
10	PRESS		154.4555							
11	PRESS RMSE		2.92931							
12	Predicted R-Squared		0.80399							
13	N		18							
14	Test Sales = 2.86764 + 0.00163 * Post-launch Sales									
15										
16										
17	ANOVA									
18		d.f.	SS	MS	F	p-value				
19	Regression	1.	680.11683	680.11683	100.86716	2.58663E-8				
20	Residual	16.	107.88317	6.7427						
21	Total	17.	788.							
22										
23		Coefficient	Standard Error	LCL	UCL	t Stat	p-value	H0 (5%)	VIF	TOL
24	Intercept	2.86764	1.26619	0.18344	5.55184	2.26478	0.03777	rejected	**	**
25	Post-launch Sales	0.00163	0.00016	0.00129	0.00197	10.04326	2.58663E-8	rejected	**	**
26	t (5%) 2.11991									
27	LCL - Lower limit of the 95% confidence interval									
28	UCL - Upper limit of the 95% confidence interval									
29	** - Requires Pro version, please upgrade.									