

Appendix B – Using Other Technologies

Section 2: Finding Regressions Using Excel (Mac)

When given a set of data, you will frequently need to find a regression model using that data. In Appendix A, you saw that you can find many regression models using a graphing calculator. One limitation of the calculator is that, when you are finished creating your model, you have no way to save the work you have done so that you can include it in documents, such as Word documents.

The spreadsheet software Excel can fill this need. Using steps that are similar to what you used with your graphing calculator, you can create scatterplots and find regression models and values for r^2 or R^2 . In this section, you can work through the process needed to create graphs that can be saved and/or imported into other documents.

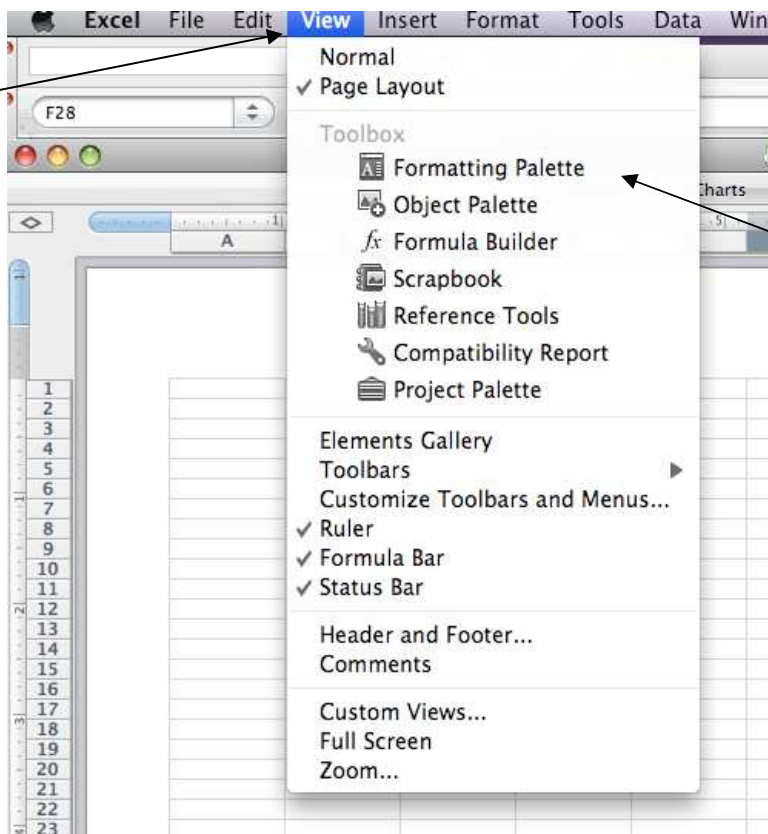
In this section, all screenshots will show the Mac version of Excel 2008. See Appendix B, Section 1 for examples worked using Excel 2010 for a PC.

Operations in Excel are performed on data that is entered into rectangles, called cells. Each cell in a spreadsheet has a unique name consisting of the letter corresponding to the column it is in and the number corresponding to the row. So cell D8 is in column D, row 8.

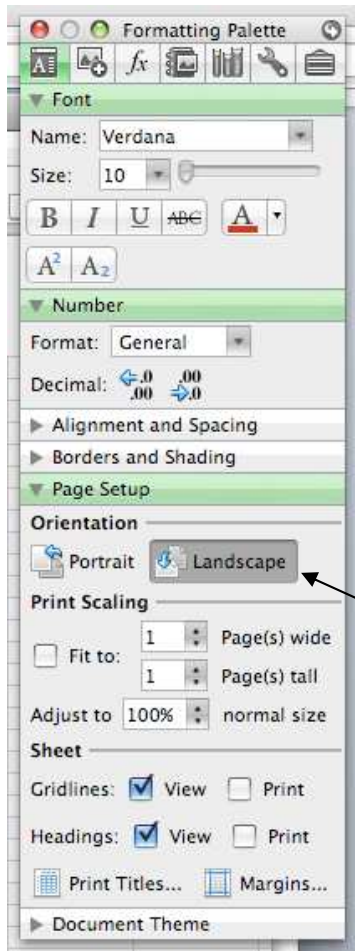
One powerful feature of Excel is its ability to perform repeated computations with just a few keystrokes. We can use this to create a table of values in Excel.

Example 1: Create a table of values that gives the function values for $f(x) = 0.27x^2 + 6.13x - 17.84$ for $x = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9$.

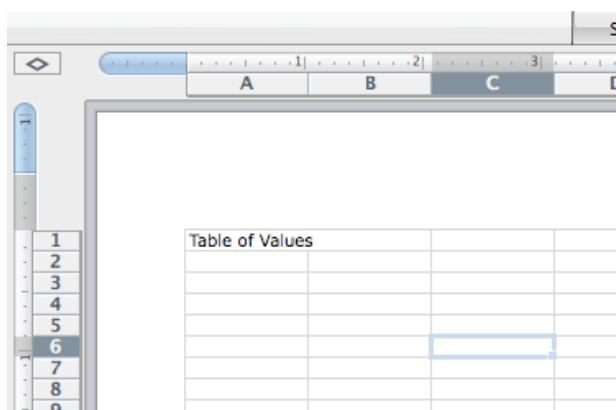
Solution: Start by opening Excel. Click on **View** and select **Formatting Palette** to open the Formatting Palette, if it is not already open.



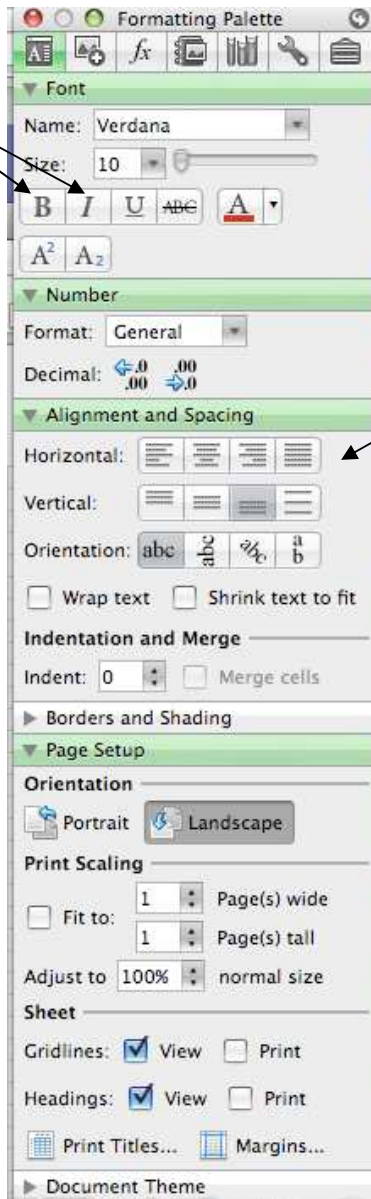
Choose **Landscape** orientation.



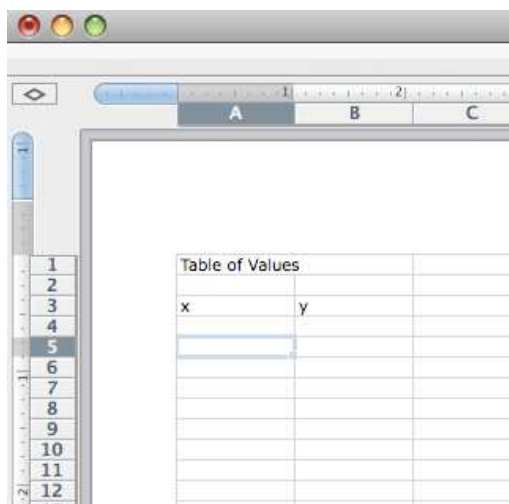
In cell A1, title the spreadsheet as “Table of Values” by putting your cursor on the cell and then typing.



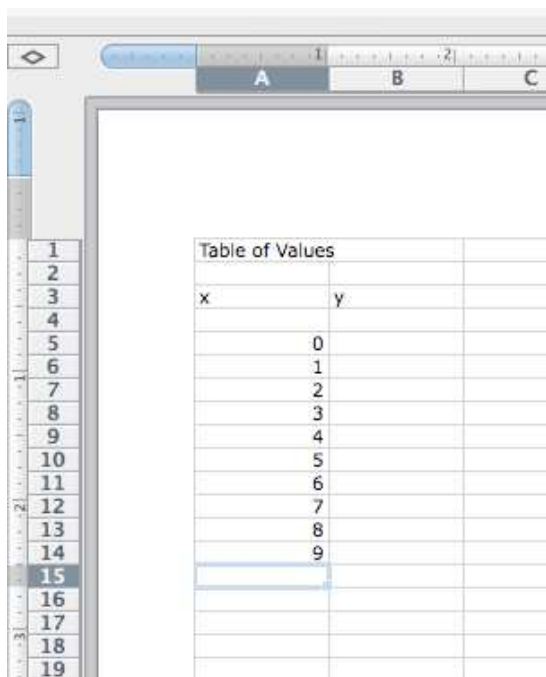
You can change the font, convert to bold font, italic font or change the alignment of the text by highlighting the cell and then clicking the icons in the **Formatting Palette**.



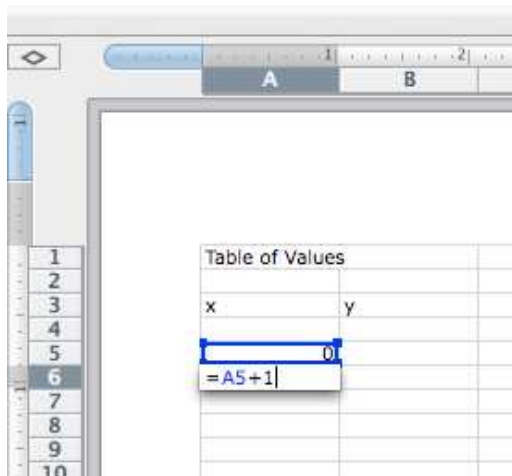
Move the cursor down to cell A3, and type x , then move the cursor over to cell B3 and type y .



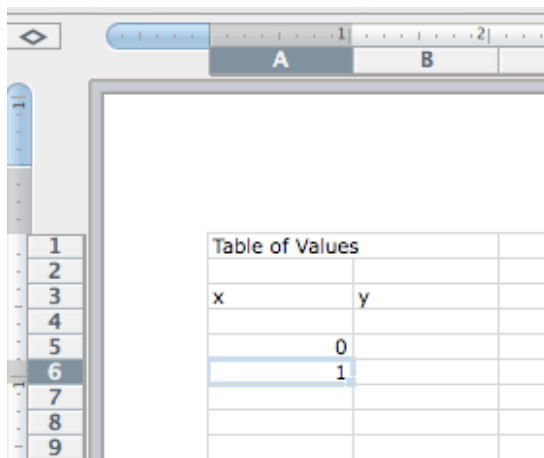
Next, you'll want to enter the values for x . Since there are just a few values, you can put the cursor on cell A5, type **0**, **Enter**. The cursor moves down automatically to cell A6. You can type **1**, **Enter** to input the second x value. Repeat the process until all of the x values listed in the problem are in the A column.



You can also enter these values by formula. This would be very helpful if you needed to enter a large number of values. To do this, type **0**, **Enter** in cell A5. Now cell A6 is highlighted. Type **=**, then move the cursor to cell A5. Then type **+1**, **Enter**. This creates a formula, which tells the software to take the value of cell A5 and add 1 to it. See this in cell A6 of the screenshot below.



Press **Enter**.



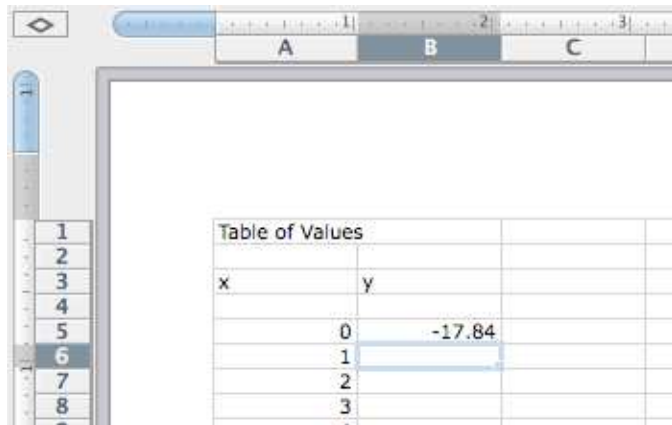
Now you want to replicate this formula. Notice the small blue box at the lower right corner of cell A6. Move your cursor close to that box. The cursor will turn into a + sign. Click and drag straight down column A. The cells will be highlighted, and when you let go of the mouse, the spreadsheet will fill in the values.

	A	B
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		

Now you can fill in the function values. You'll enter the function once, then copy it by clicking and dragging, as you did for the x values. Put the cursor on cell B5 and type **=0.27*A5^2+6.13*A5-17.84**. Notice that you must use an asterisk (*) to indicate multiplication. Excel does not understand juxtaposition to mean multiplication. All formulas in Excel must begin with an equal sign.

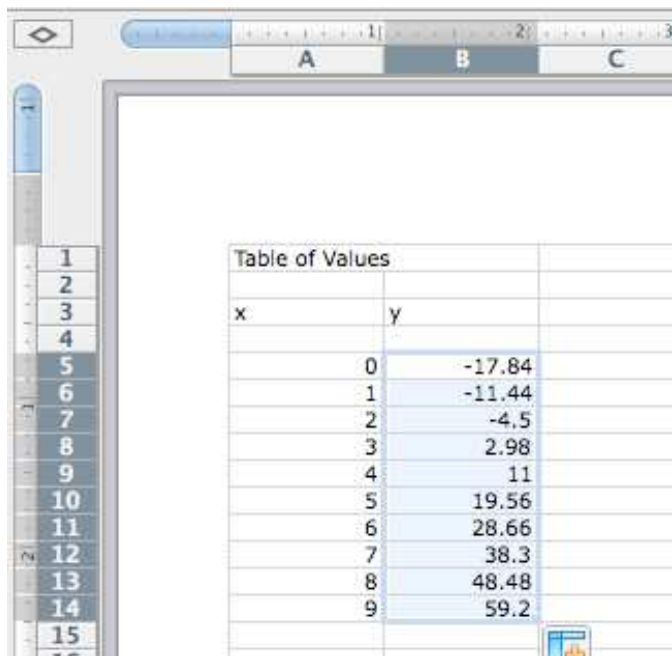
	A	B	C	D	E
1					
2					
3					
4					
5		=0.27*A5^2+6.13*A5-17.84			
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					

Now press **Enter**. The function value will appear in cell B5.



x	y
0	-17.84
1	
2	
3	
4	
5	
6	
7	
8	
9	

Now you can replicate the formula. Make sure the cell containing the first function value (B5) is highlighted. Then put the cursor close to the blue box on the lower right corner of cell B5. Click and drag until the cells in column B corresponding to the values in column A are all highlighted. Let go, and all of the values will appear in column B.



x	y
0	-17.84
1	-11.44
2	-4.5
3	2.98
4	11
5	19.56
6	28.66
7	38.3
8	48.48
9	59.2

Solution: Now that you have a table of values, you can plot the points. Click on the **Insert** tab at the top of the screen.



The screenshot shows the Microsoft Excel 2008 for Mac interface. The menu bar at the top includes Apple logo, Excel, File, Edit, View, Insert, Format, Tools, Data, Window, and Help. The 'Insert' menu is open, showing options like Text, Table, Chart, and others. The 'Chart' option is highlighted. The 'Charts' task pane is open on the right, showing various chart types: All, Area, Bar, Bubble, Column, Doughnut, Line, Pie, Radar, Stock, Surface, and XY (Scatter). The 'Line' chart type is selected. The spreadsheet shows a table of values with columns x and y.

x	y
0	-17.84
1	-11.44
2	-4.5
3	2.98

Make sure both columns of numbers in your table of values are highlighted.

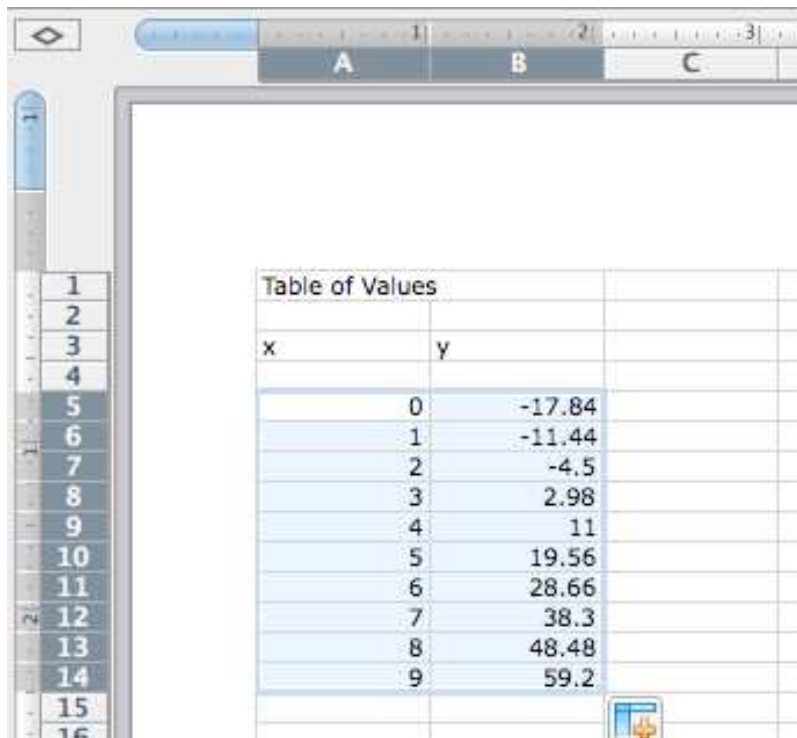
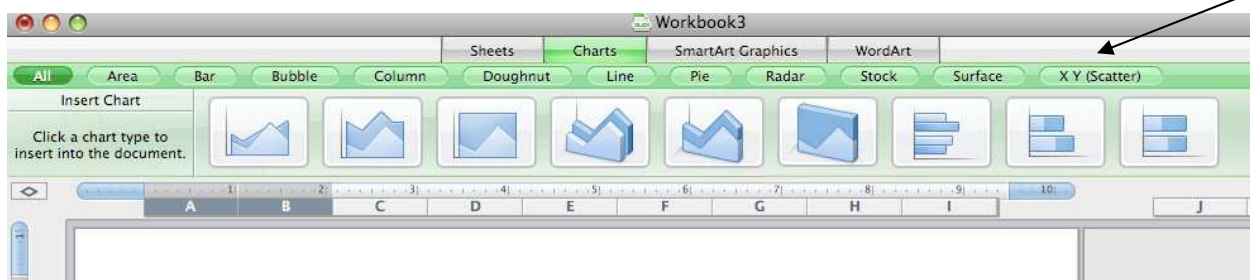
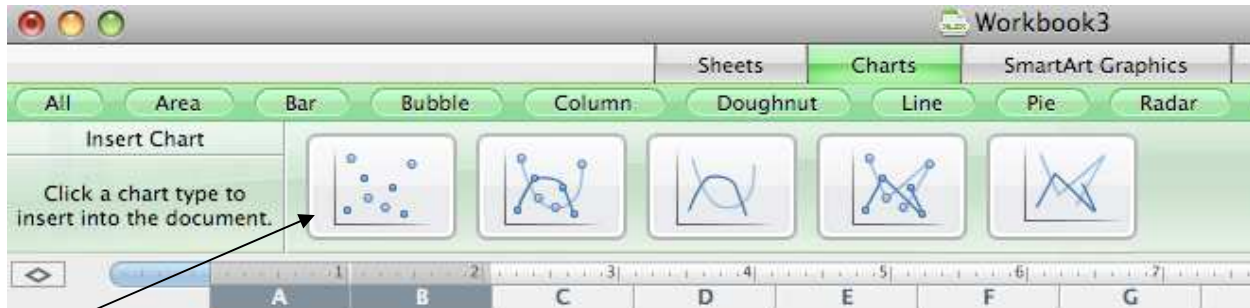


Table of Values	
x	y
0	-17.84
1	-11.44
2	-4.5
3	2.98
4	11
5	19.56
6	28.66
7	38.3
8	48.48
9	59.2

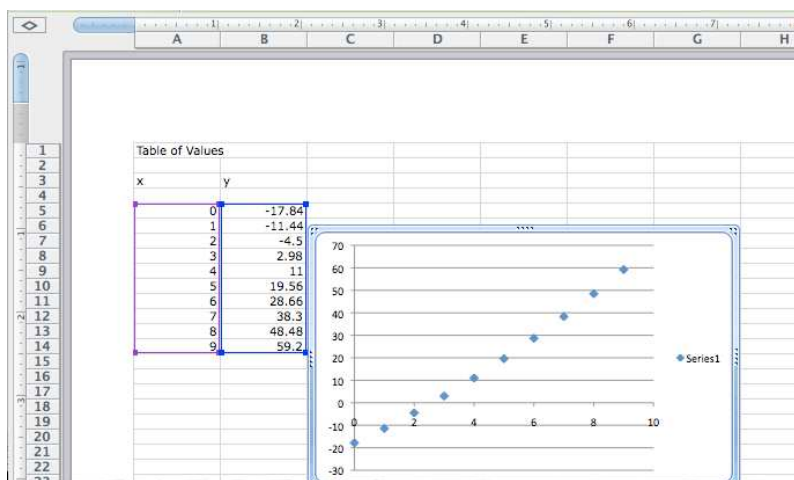
You want to create a scatterplot, so find the tab in the green bar that reads **X Y (Scatter)**.



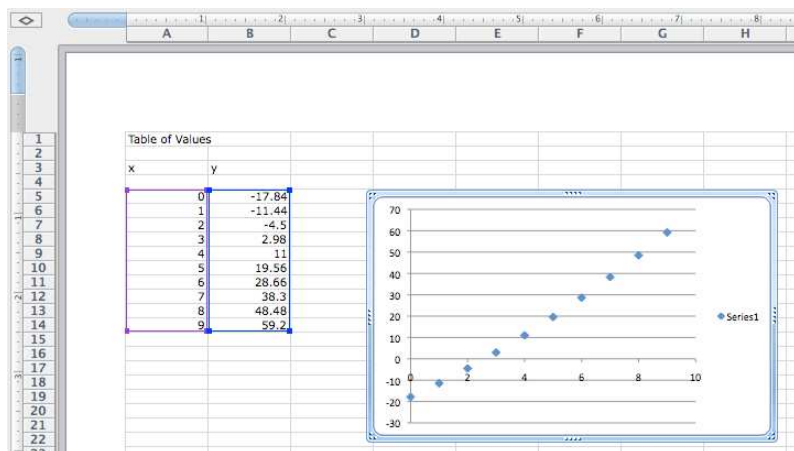
Click on **X Y Scatter**. Here's the result:



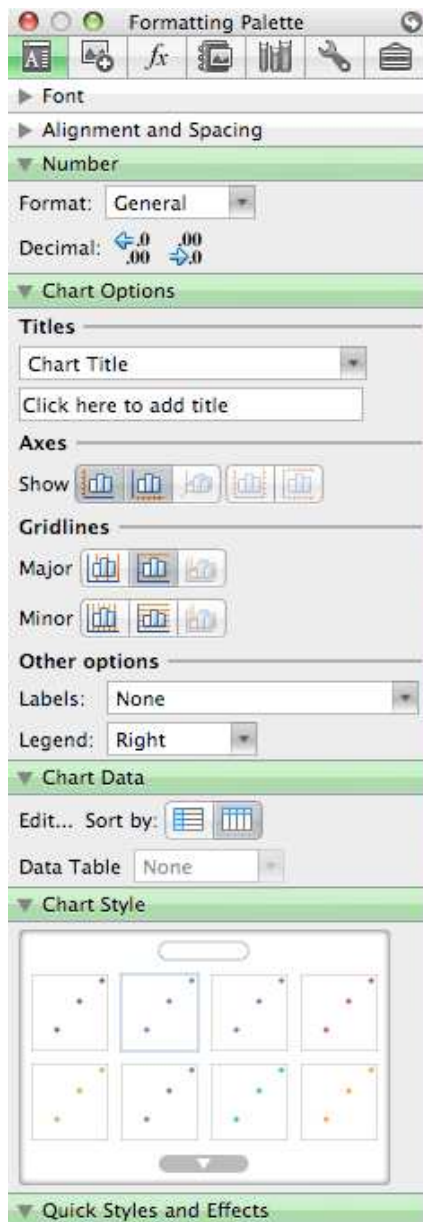
Now click on the scatterplot that contains only the points. Here's the result:



This shows both the table of values and the scatterplot. You can click and drag the scatterplot if you need to move it so that you can see both the plot and the table of values.

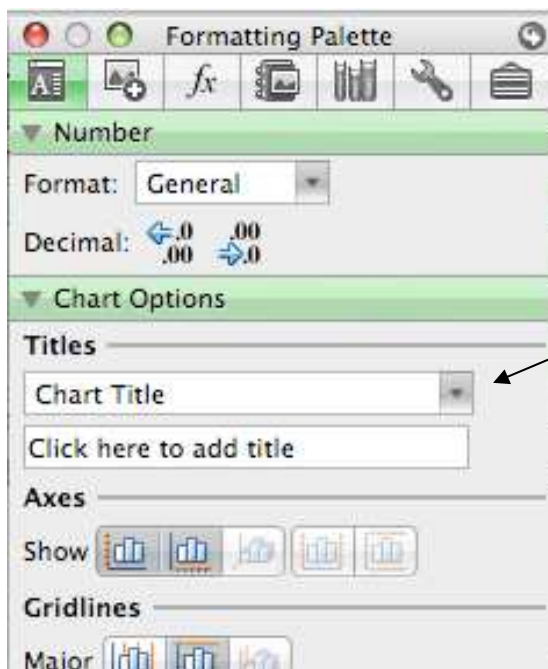


You can customize the look of the graph by inserting a title, inserting axis labels, changing the color of the markers, changing or deleting the legend, and changing or removing gridlines. You can also insert the regression equation and value for r^2 or R^2 . Most of these modifications are made using the **Formatting Palette**. Make sure it is open. Then click inside the border of the scatterplot, and notice the change in the **Formatting Palette**.



The palette now contains items that are specific to formatting a chart.

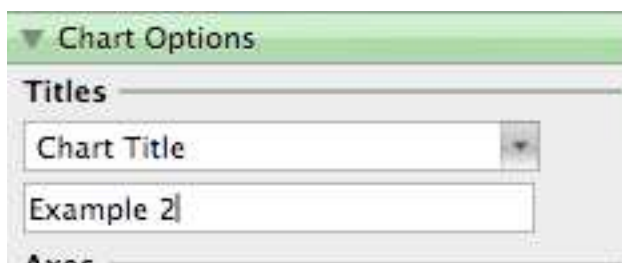
To insert a title: Note the first topic under **Chart Options** on the **Formatting Palette**. You can create titles for your chart.



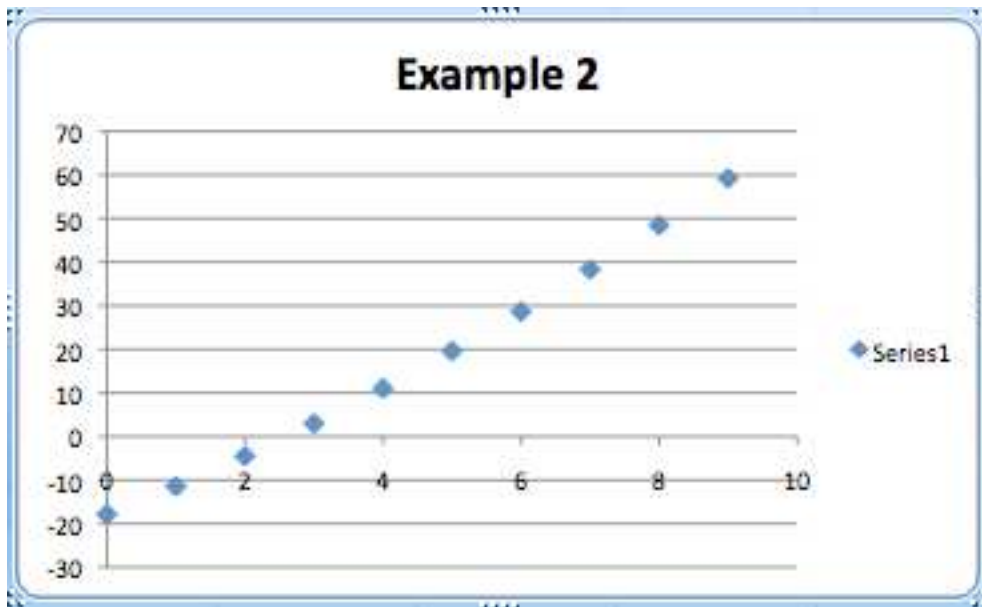
Put your cursor on the line that reads “Click here to add title”. The line will be highlighted in blue. You can then type in your chart title.



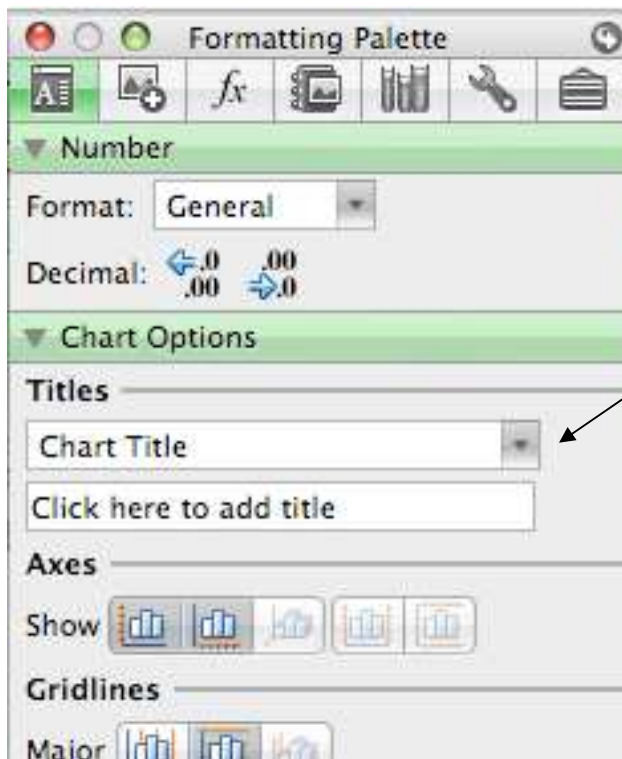
Call this scatterplot “Example 2”.



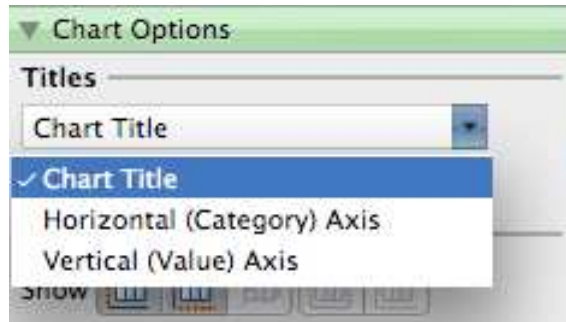
As you type, the title appears above the scatterplot. The scatterplot will automatically re-size as you type.



To label the axes: You'll use a similar method to label the axes. Start by labelling the horizontal (x) axis. There is a pull-down menu under **Chart Options**. You currently see "Chart Title".



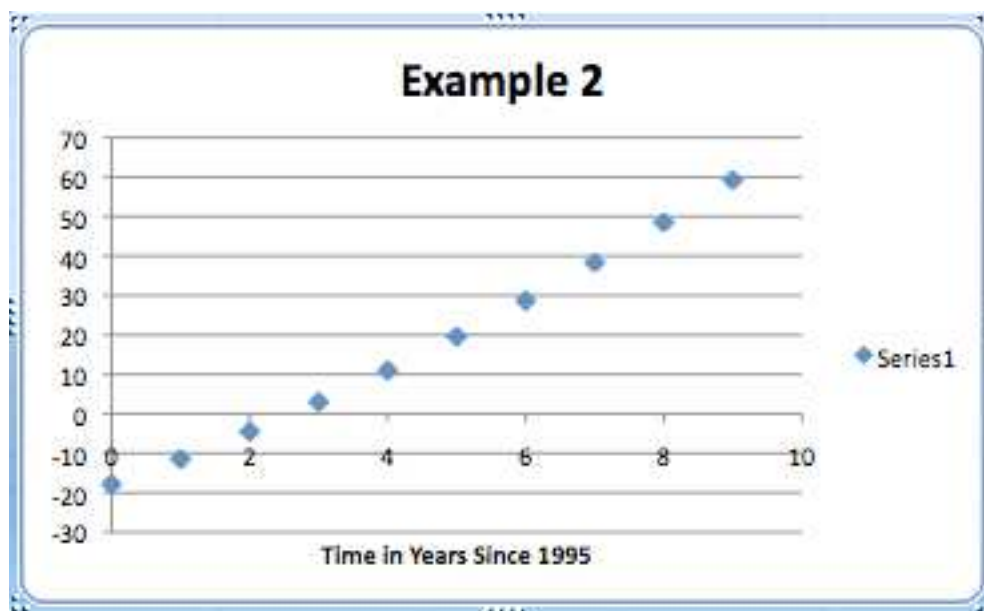
Click on the down arrow and you'll see some other options.



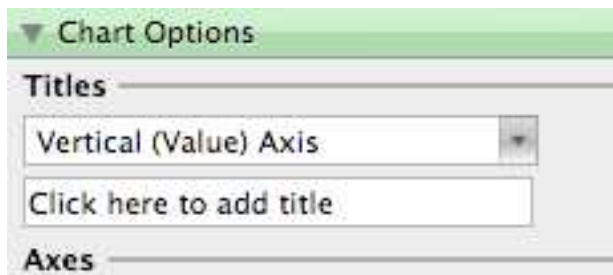
Select **Horizontal (Category) Axis** to rename the horizontal (x) axis. You can type in your label in the lower box. This problem did not give us a context for these numbers, so you can just make something up for this example. Call it “Time in Years Since 1995.”



Here's the result. Note that as you type in the title line, the title appears below the scatterplot.



You can also enter a vertical axis title, by repeating the steps but selecting **Vertical (Value) Axis**.



▼ Chart Options


Titles

Vertical (Value) Axis

Click here to add title

Axes

You can type the title in the second box. Just as for the horizontal axis, just make something up for the vertical axis. You can use “Profits in Millions of Dollars.”



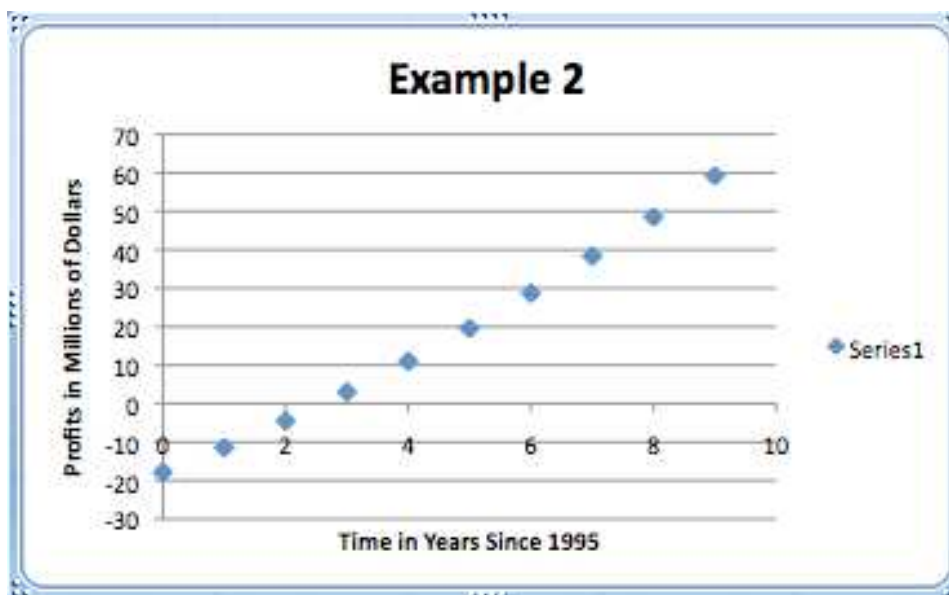
▼ Chart Options

Titles

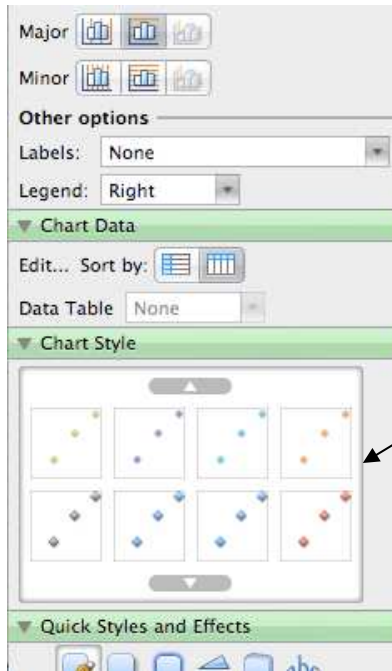
Vertical (Value) Axis

Profits in Millions of Dollars

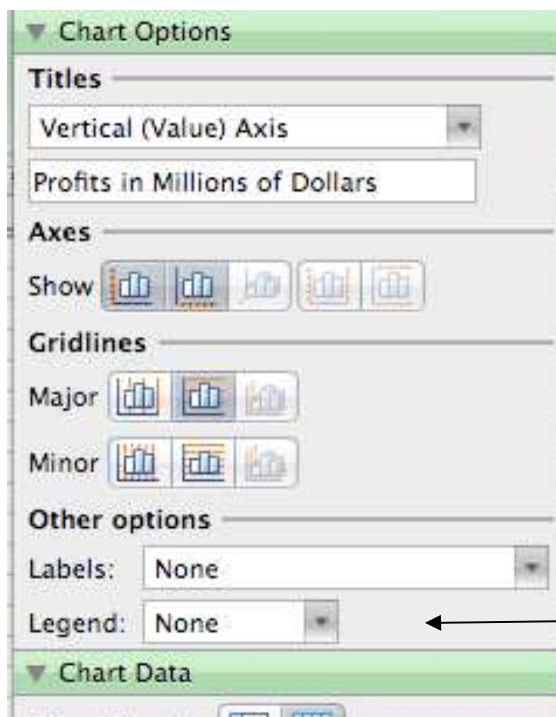
Note that as you type in the title line, the title appears rotated and to the left of the scatterplot. Here's the result:



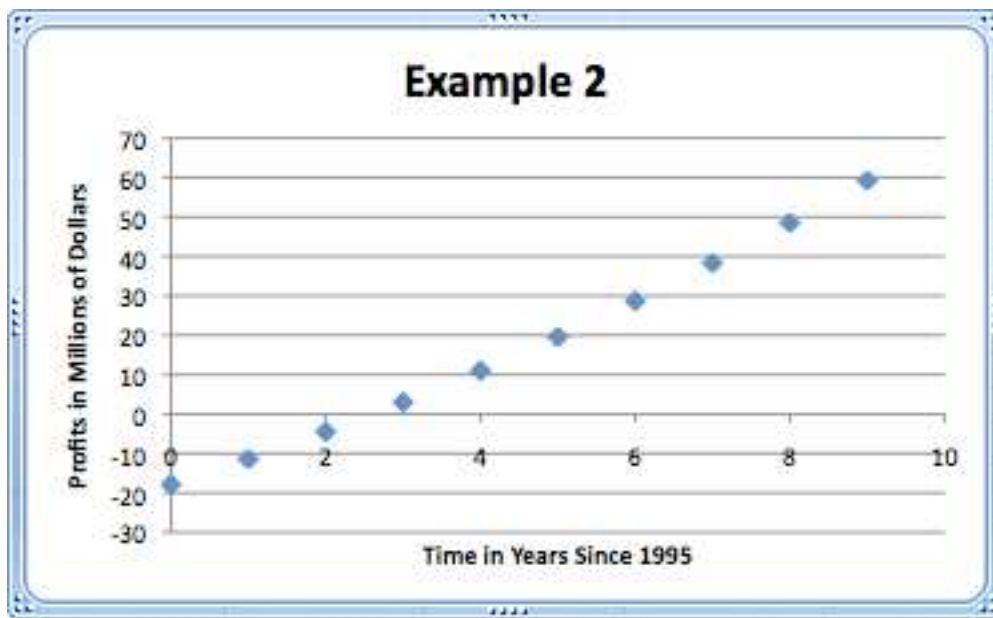
To change the color of the markers: If you'd like to change the color of the markers, select the **Chart Style** on the **Formatting Palette** and select another color. You can use the up and down arrows to scroll through the options.



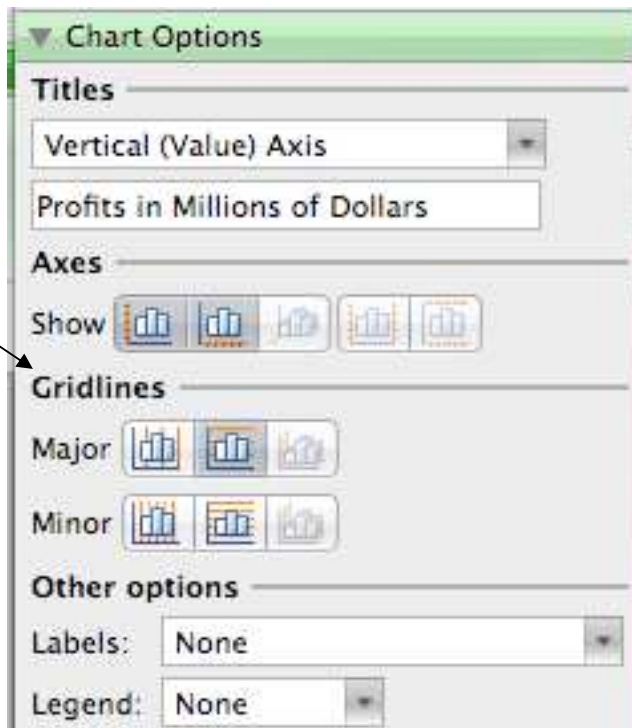
To change or delete the legend: The legend is the little diamond with the notation “Series1” next to it. Legends are helpful if you have two (or more) different sets of data shown on the same graph. This feature will be used in the next example. If you only have one set of data, the legend is not really needed. You can delete it by selecting **Legend** on the **Formatting Palette**, clicking on the down arrow and selecting **None**. The scatterplot will automatically resize.



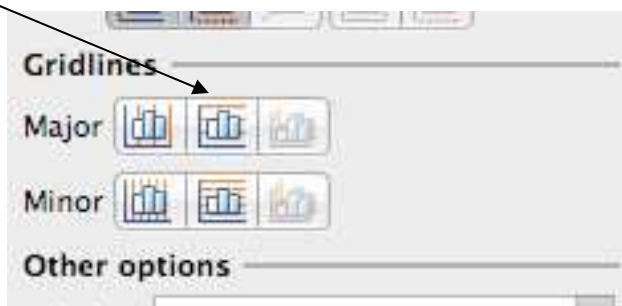
Here's the chart with no legend:



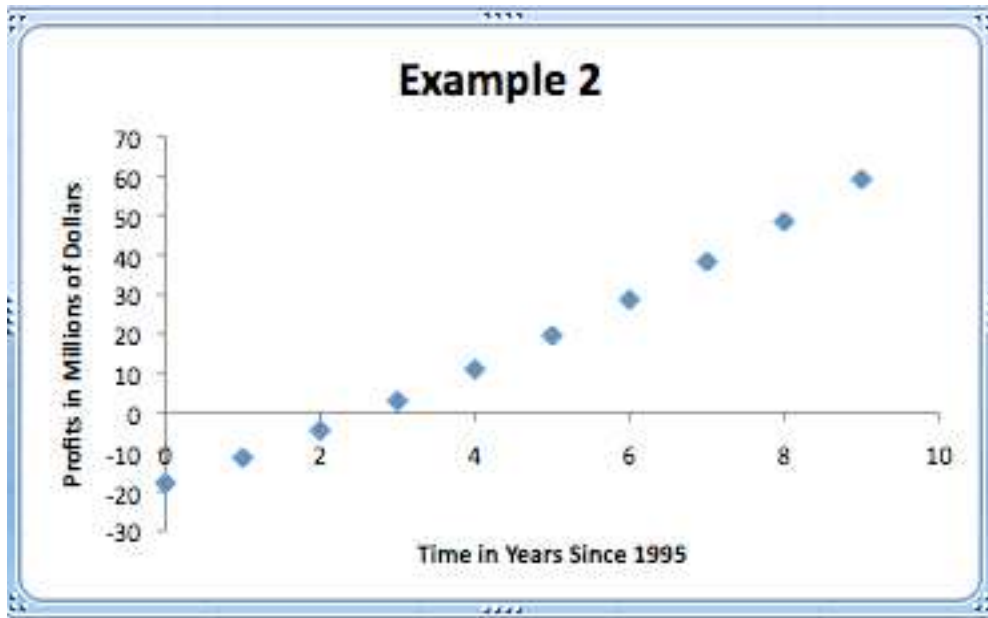
To delete gridlines: The graph currently has horizontal gridlines showing at intervals of ten units. These can help you approximate the y values for each of the points, but you can delete the gridlines if you don't want them. From the **Formatting Palette**, find **Gridlines** under **Chart Options**.



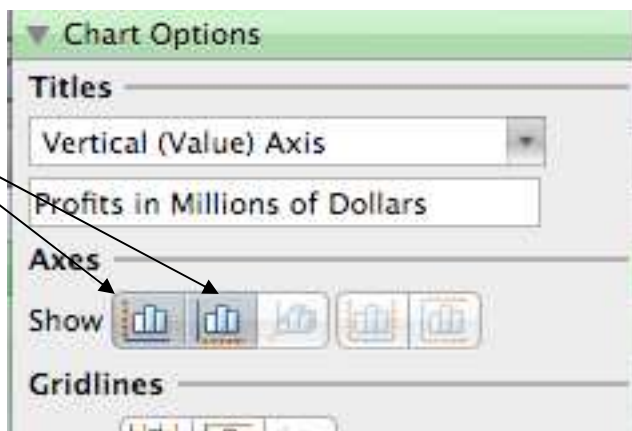
Gridlines can be major (e.g., every 10 units) or minor (e.g., every 5 units), and they can be vertical or horizontal. Note the red lines on the icons. You can activate major and minor axes by clicking on the icons. When activated, the icon will have a darker background. If you do not want the gridlines activated, make sure the icon has a lighter background. If you look at the screenshot above, only major horizontal gridlines are activated, and that's what you've seen on the scatterplot. Deactivate major horizontal gridlines by clicking on that icon, and the gridlines will disappear.



Here's the result:

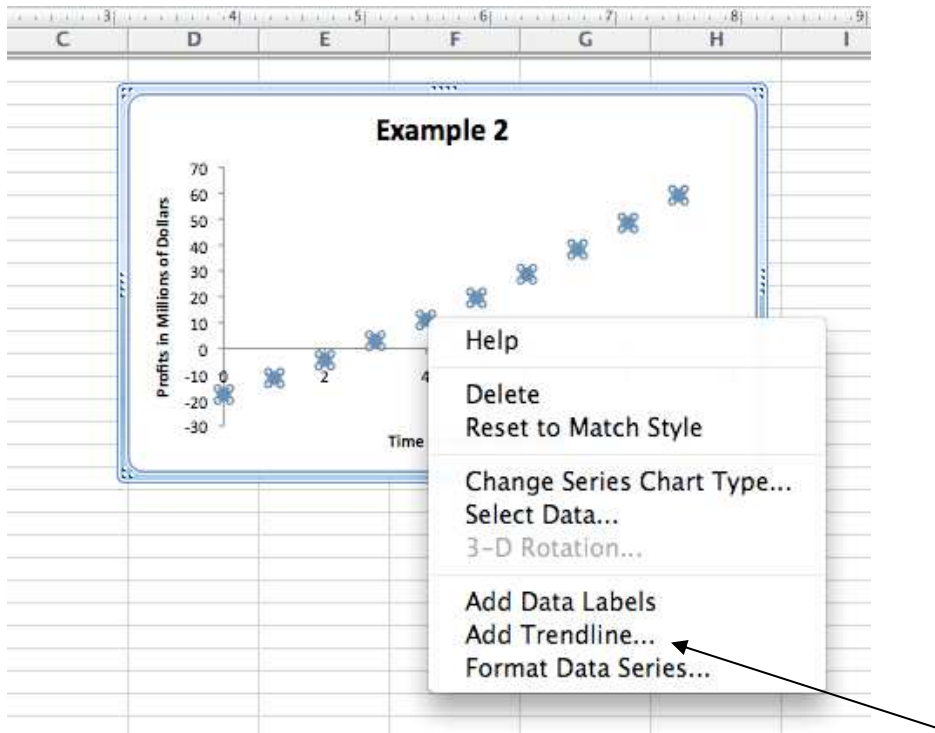


On a PC, you can customize the axes. This is not an option on a Mac. You can, however, enable or disable the axes. The screenshot below shows that both axes should appear on the scatterplot. These icons are located in the **Formatting Palette** under **Chart Options**.

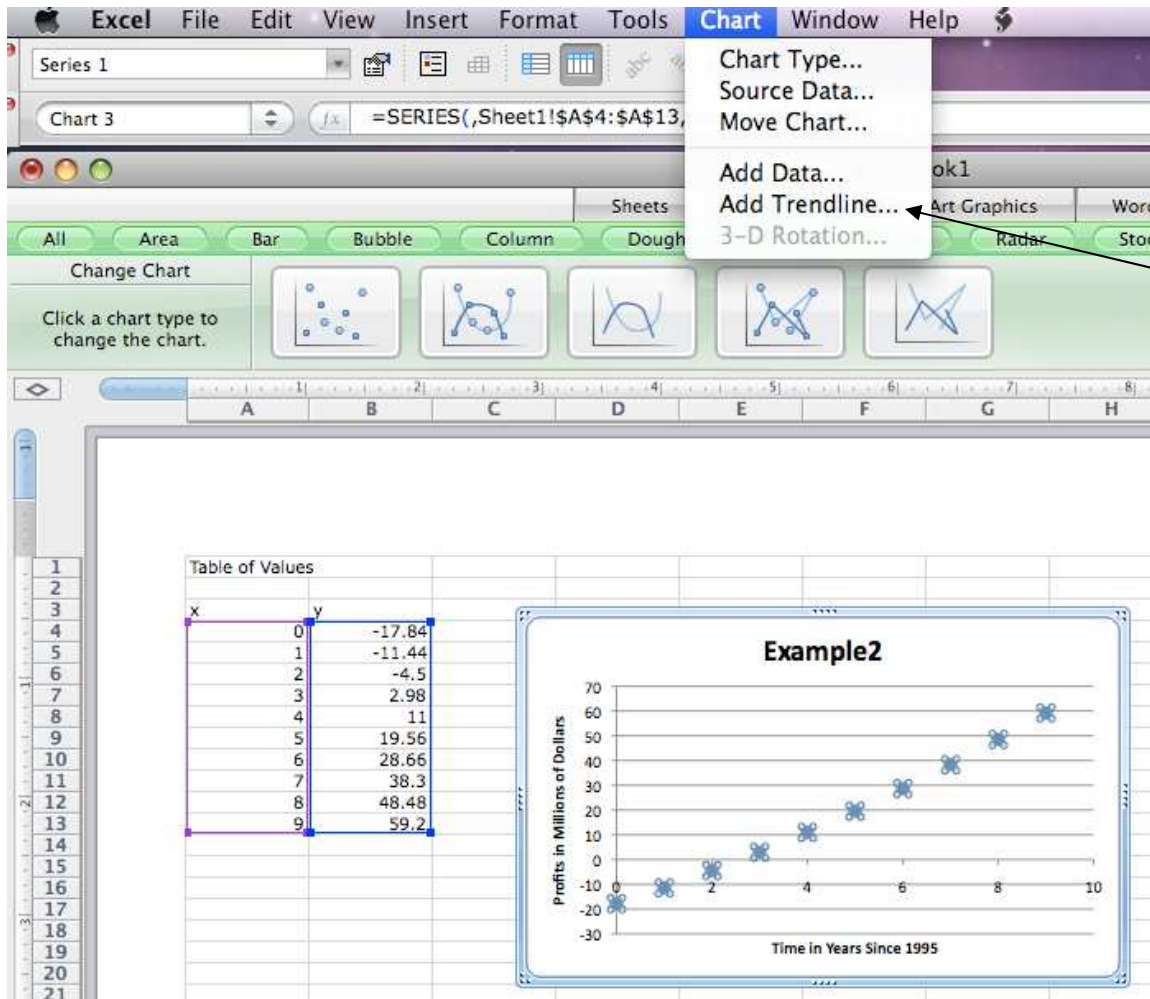


If you do not want one or both axes to appear on your scatterplot, disable the appropriate axis by clicking on the icon. For this example, leave both axes activated.

To display a regression equation and the value for r^2 or R^2 : Regression equations are called “trendlines” in Excel. There are two ways to access the trendline menus. Option 1: put your cursor on any data point in the scatterplot until the data points change. Then press **Control** while clicking the mouse. Select **Add Trendline**.

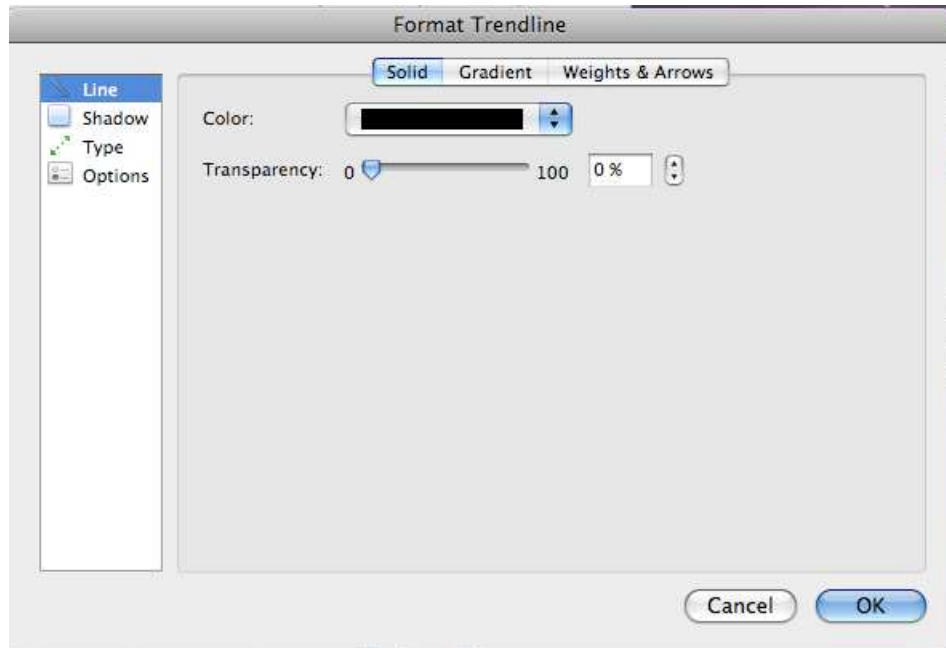


Option 2: put your cursor on any data point in the scatterplot until the data points change. Then click on **Chart** at the top of the page. You'll be able to click on the **Add Trendline** option.

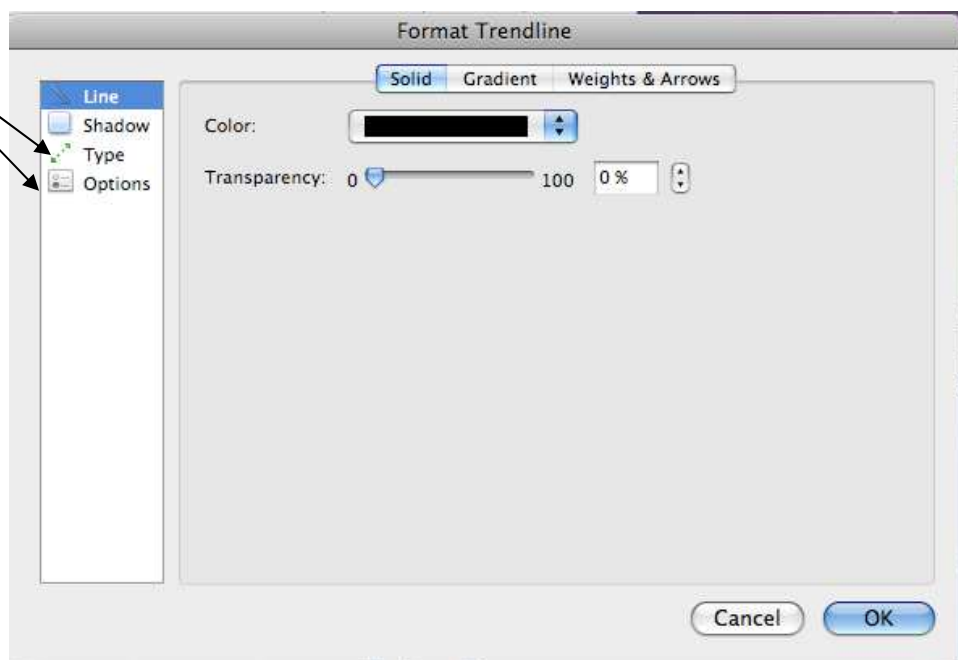


The change to the data points is temporary.

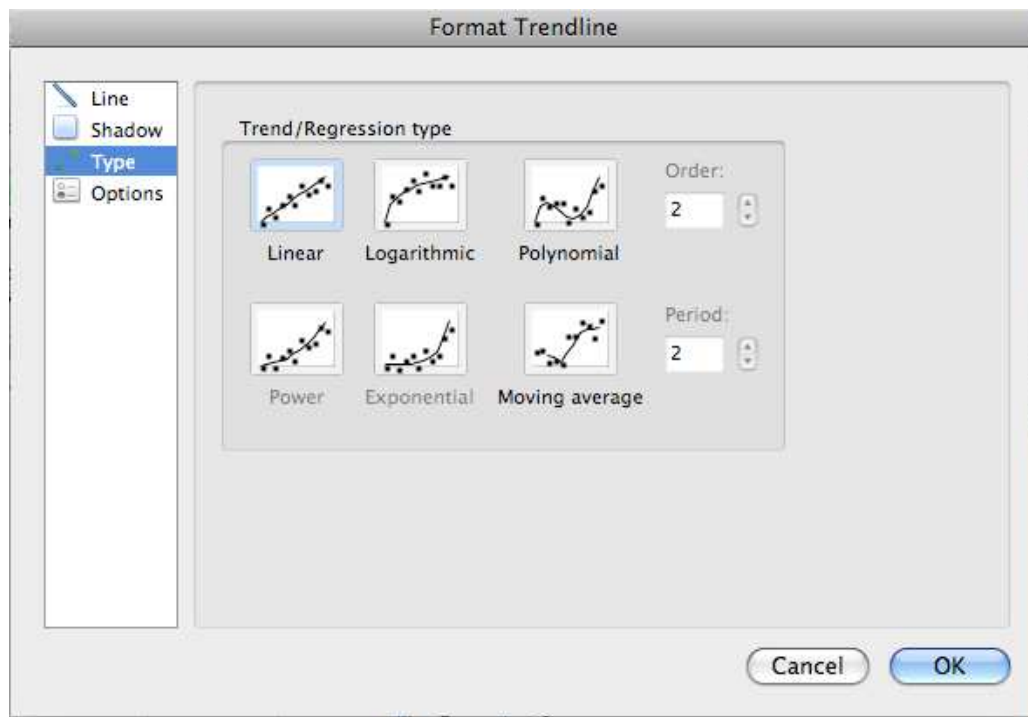
Here's the result when you click on **Add Trendline** using either option:



You'll need to use two of the items on the left, **Type** and **Options**.

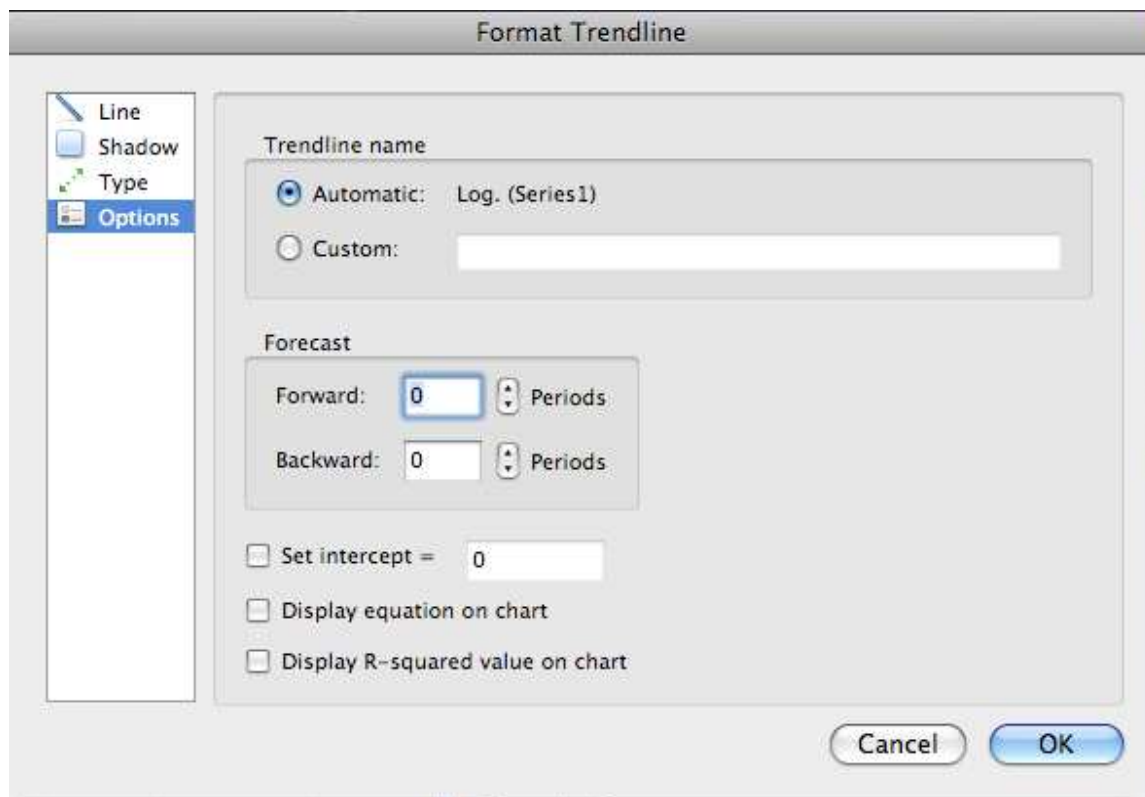


First, click on **Type**. Here's the result:

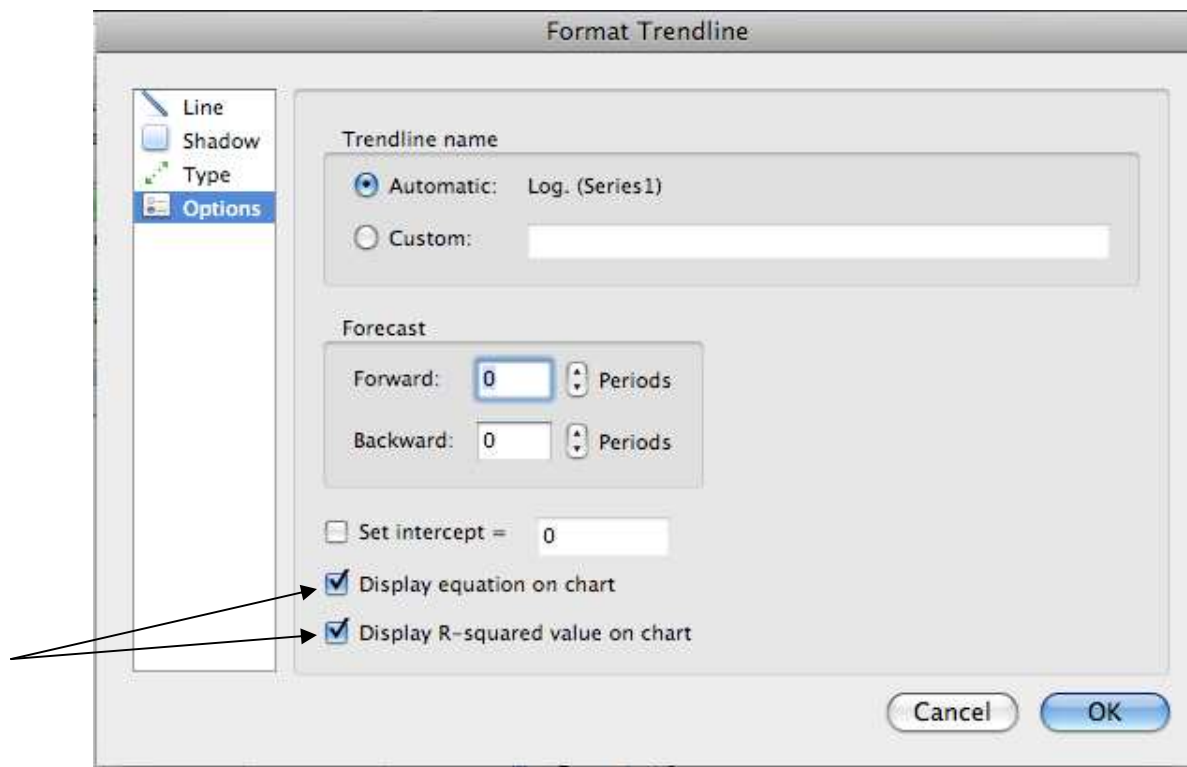


You can select the type of regression line you'd like to display: linear, logarithmic, polynomial from order 2 (quadratic) to order 6 (6th degree polynomial), power and exponential. If you want a quadratic, cubic or quartic function, choose **Polynomial** and then select the order (2 for quadratic, 3 for cubic and 4 for quartic, etc.). For this example, we'll find a linear regression model. Click on the linear plot and then click on **OK**.

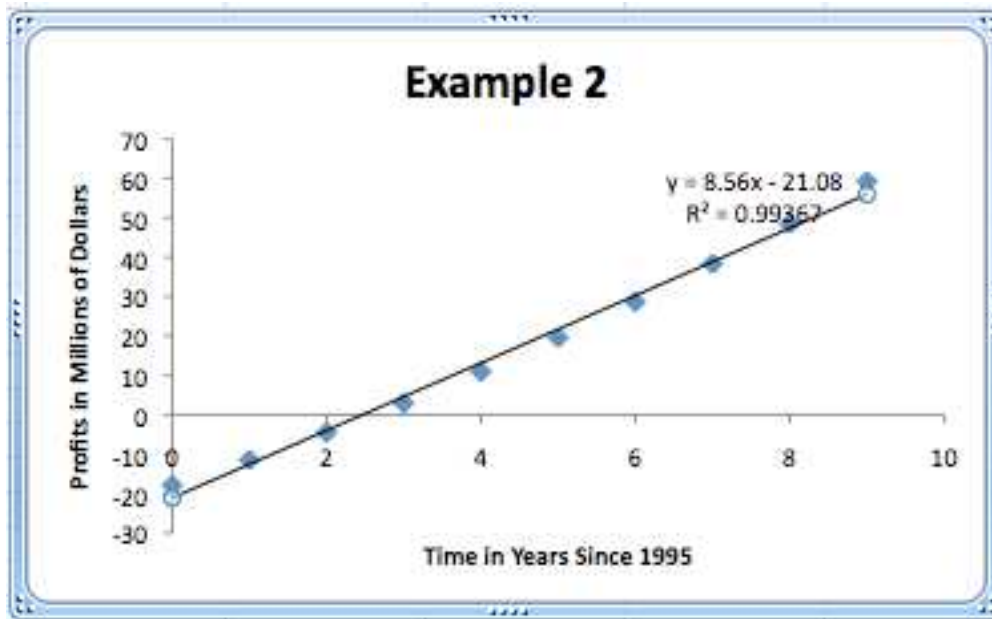
We also want to display the regression equation and the value for R^2 . To do this, select **Options** from the list at the left.



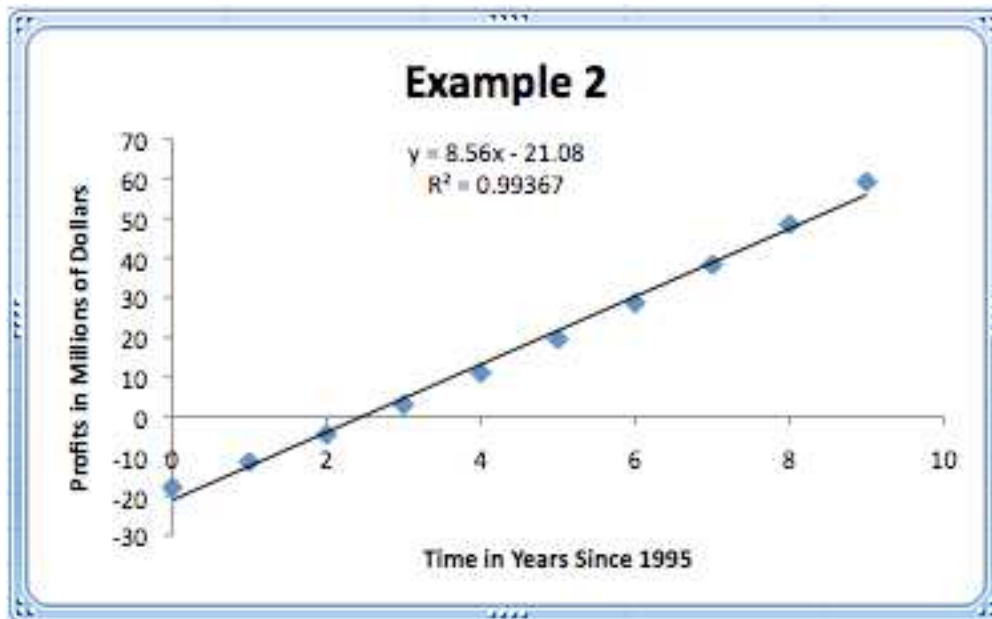
Click on the two boxes at the bottom of the screen; then click on **OK**.



The regression equation and the value for R^2 now appear on the scatterplot.



The equation and the R^2 are obscured by the graph. They are in a text box, and you can move it by clicking on the text box and dragging it to where you want it.



The chart is now completely formatted.

Note that the coefficient of x and the constant are given to just two decimal places. You may want or need to display and use equations with more decimal places. Excel can do this for you. Double click on the text box containing the equation and R^2 value. This opens the **Format Trend Label** menu. Choose **Number** from the list of options on the left; then choose **Number** from the list of categories. Now locate the **Decimal places** box. Use the up/down arrows until the number of decimal places you want displayed is in the box, or just type in the number of decimal places you want in the **Decimal places** box. Then click on **OK**, and the displayed equation will adjust to the requested number of decimal places.

Example 3: Suppose you want to compare the prices of roundtrip airfares for two different airlines between Houston and Newark over a one year period. You find a website that gives the average airfare for Airline A and for Airline B for each month, so you have 12 data points for each airline. This information is given in the table shown below.

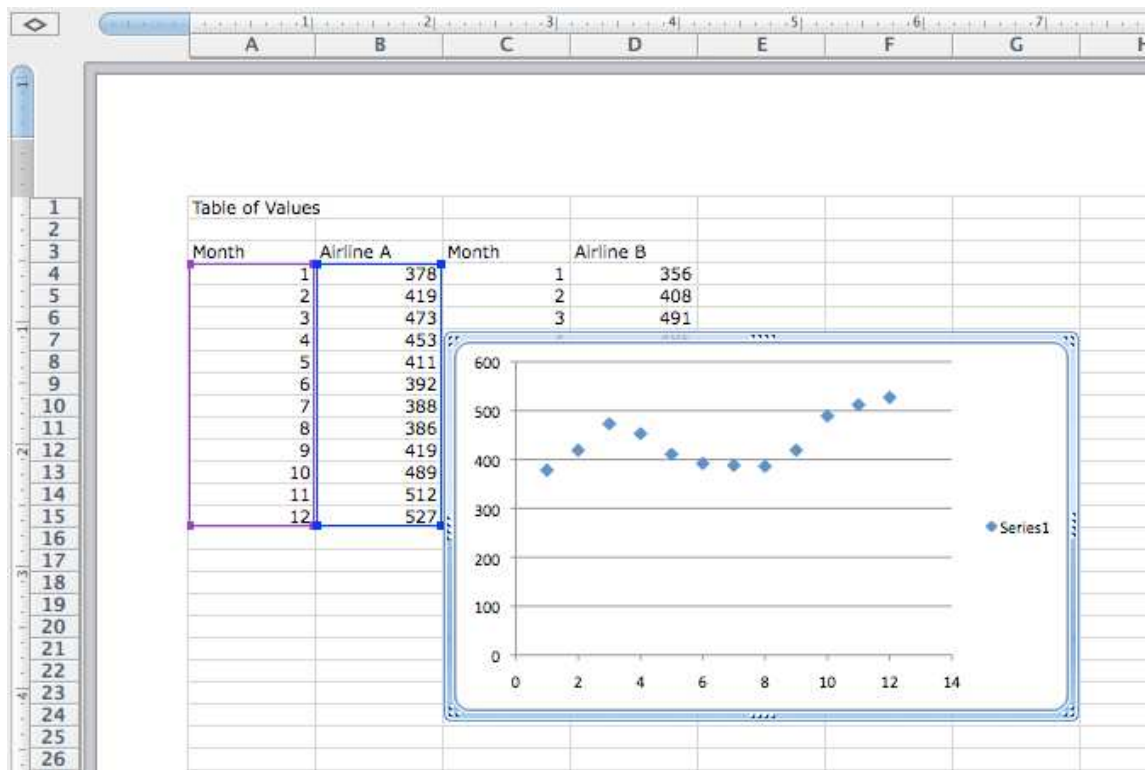
Roundtrip Airfares: Houston - Newark												
Month	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Airline A	378	419	473	453	411	392	386	388	419	489	512	527
Airline B	356	408	491	485	461	421	432	424	461	409	395	384
Fares are stated in US dollars												

Create one scatterplot that shows both sets of data. Include a legend. Find quartic regression models (4^{th} order polynomials) and display each on the scatterplot. Also display the value for R^2 for each regression model.

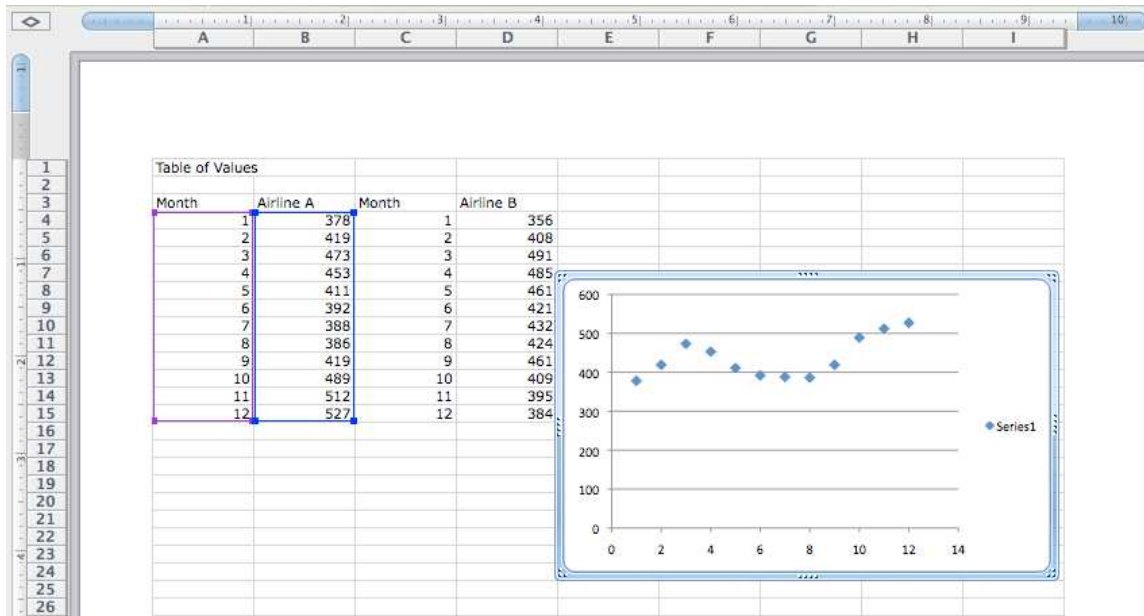
Solution: Start by creating two tables of values, one for each airline. Use x values of 1, 2, 3, ..., 12 to correspond to the months of the year. No formula is given for the airfares, so you will need to enter the y values manually. You can enter the x values manually or by using a formula as shown in Example 1.

	A	B	C	D	E	F
1	Table of Values					
2	Month	Airline A	Month	Airline B		
3	1	378	1	356		
4	2	419	2	408		
5	3	473	3	491		
6	4	453	4	485		
7	5	411	5	461		
8	6	392	6	421		
9	7	388	7	432		
10	8	386	8	424		
11	9	419	9	461		
12	10	489	10	409		
13	11	512	11	395		
14	12	527	12	384		
15						
16						
17						

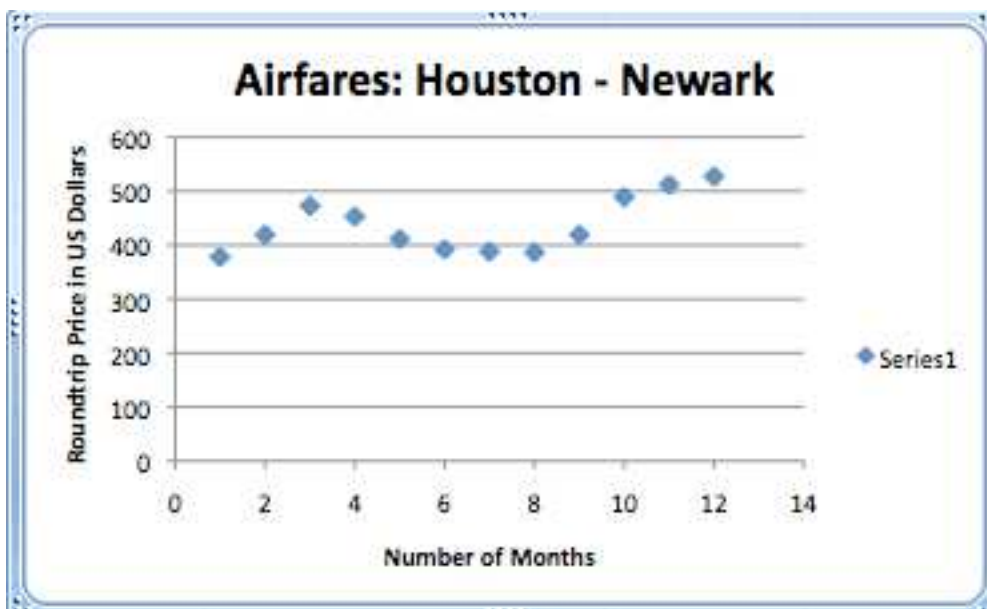
Now highlight the Airline A data and use it to create a scatterplot.



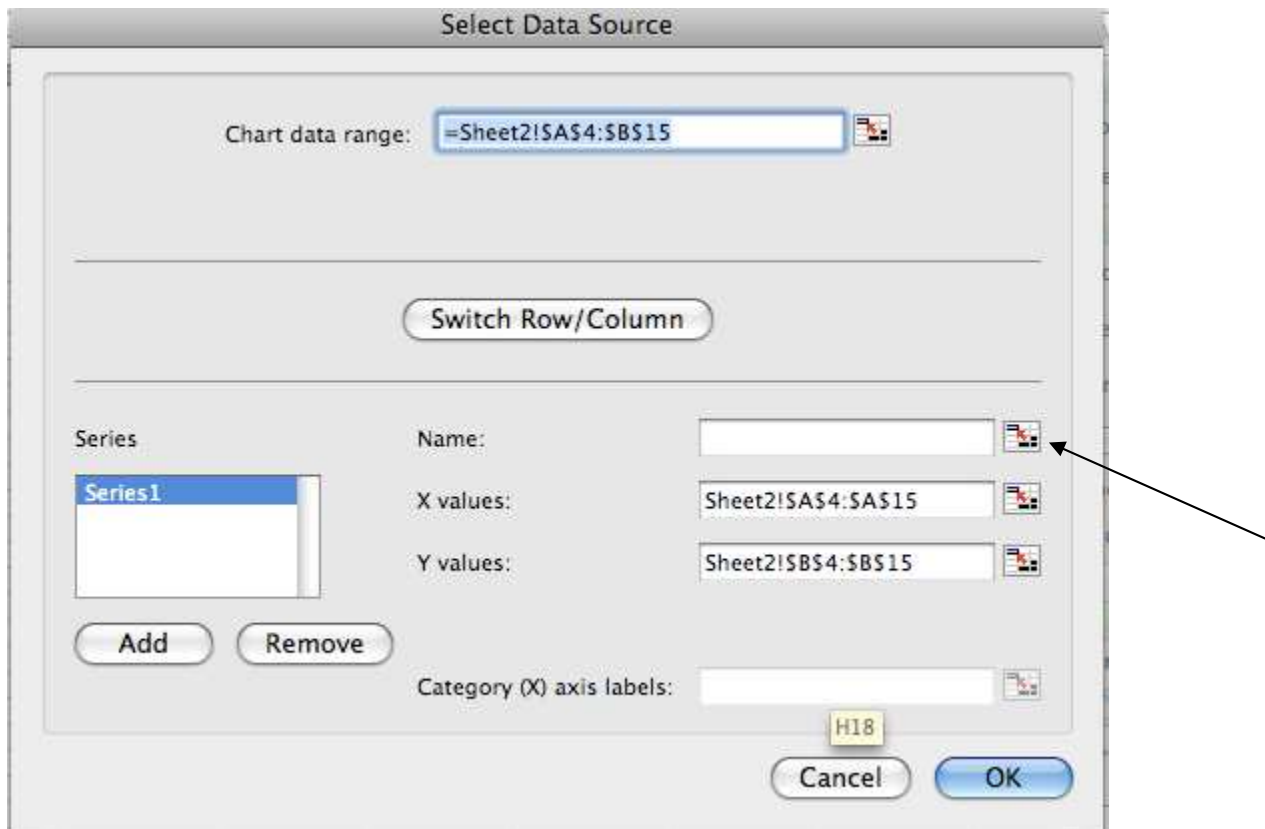
You can drag the scatterplot to the right so that you can see everything on the spreadsheet.



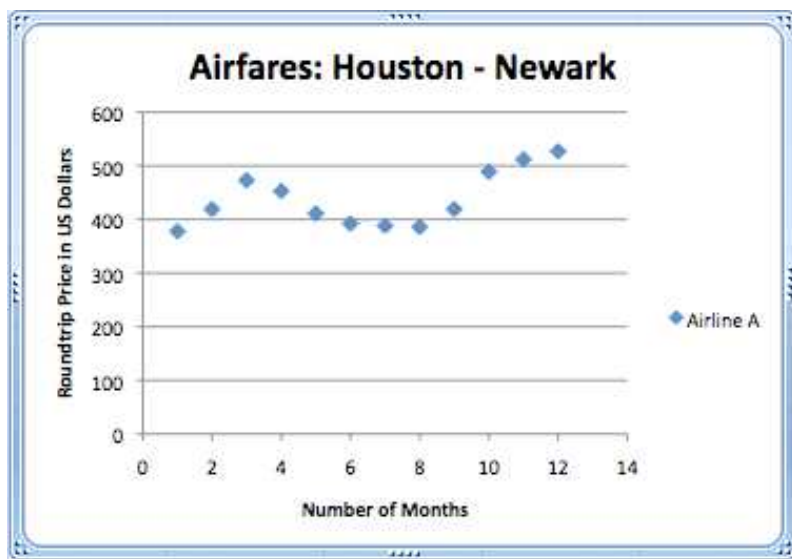
Customize the scatterplot to match the information given in this problem. Give the scatterplot a title, and label the axes. Since the data you are given has a context, you will need to write labels that make sense given the problem. In this case, a good chart title might be “Airfares: Houston – Newark”, a good label for the horizontal axis might be “Number of Months” and a good label for the vertical axis might be “Roundtrip Price in US Dollars”. Note that these could be worded differently – these titles just need to clearly describe what is represented by the chart and by each axis.



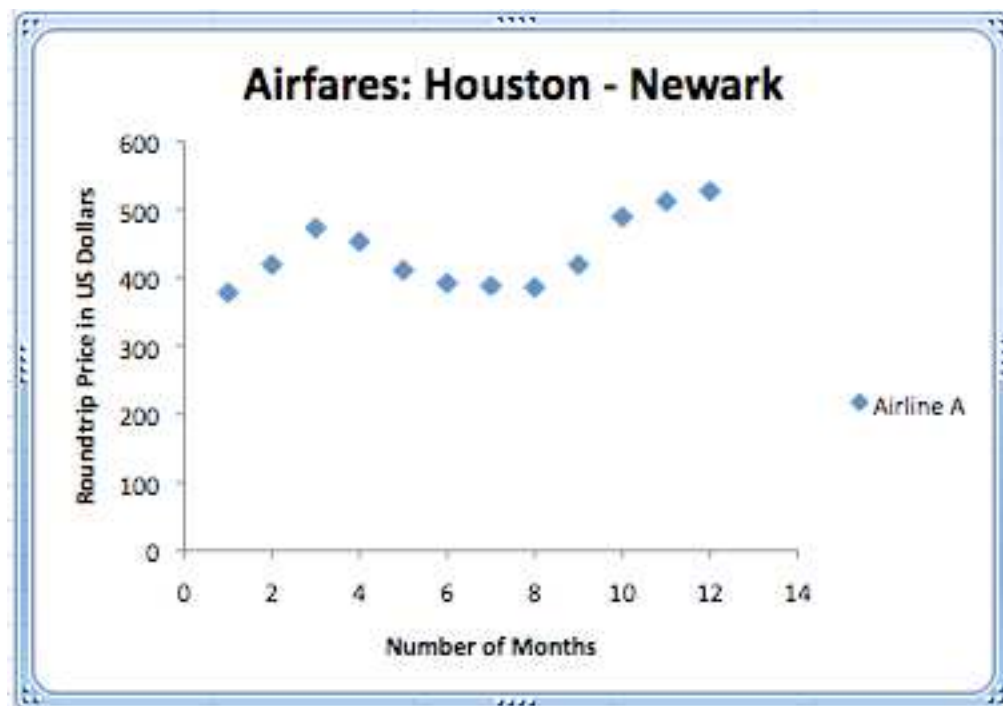
You'll need to edit the legend so that it shows "Airline A" instead of "Series1". To do this, put the cursor on one of the data points and press **Control** while clicking. Then select **Select Data Source**. (A second method for accessing this menu is to click on a data point until it changes appearance and then click **Chart, Source Data**.) Here's the resulting screen.



You want to enter the legend name. Type "Airline A" in the **Name:** box and click **OK**. Here's the result. Note that the legend has changed.

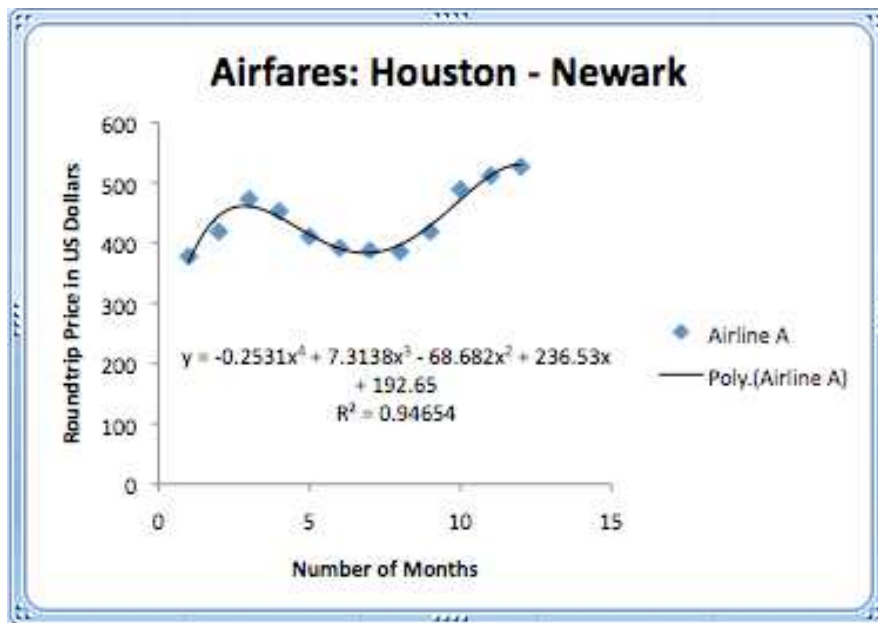


Delete the gridlines.



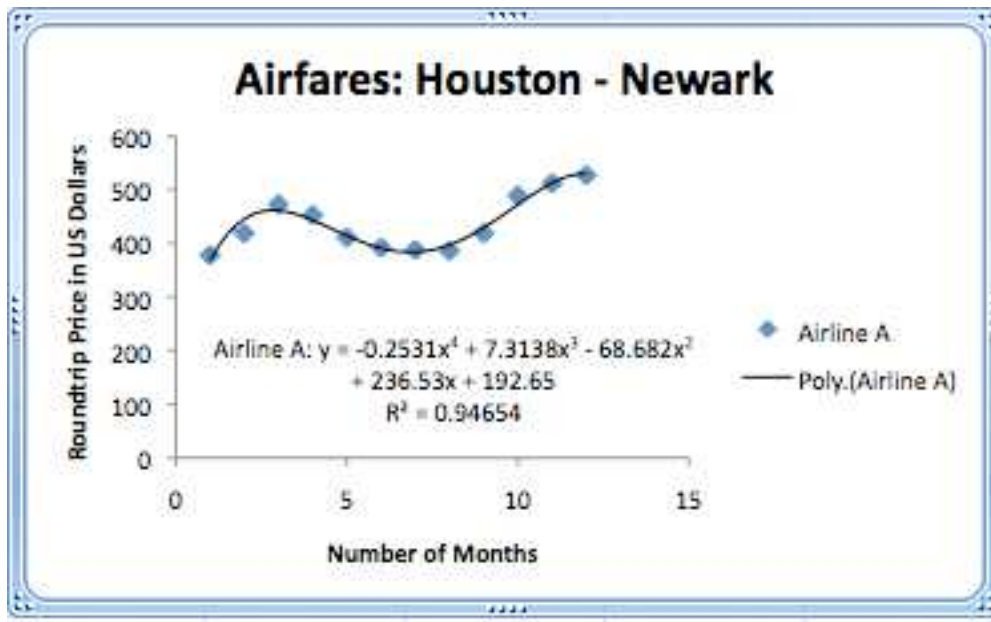
Now display the quartic regression model, along with its R^2 value. After making these selections on the **Add Trendline** menu, click **Close**.

Here's the result. You can drag the equation to a different location on the scatterplot.

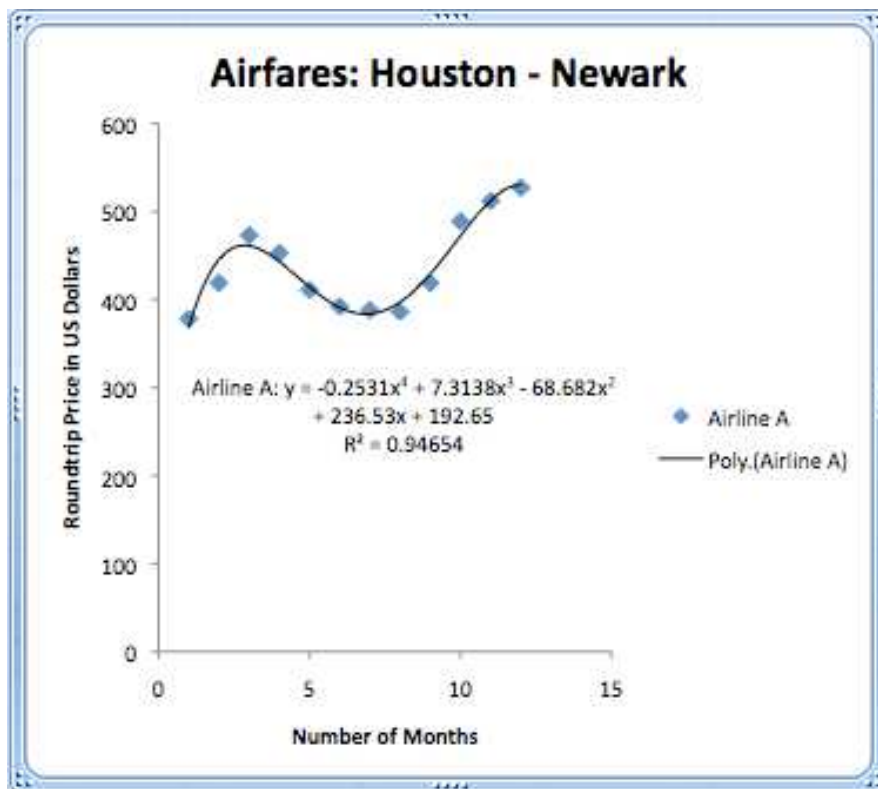


Note that the result also shows the mark for the Airline A polynomial in the legend.

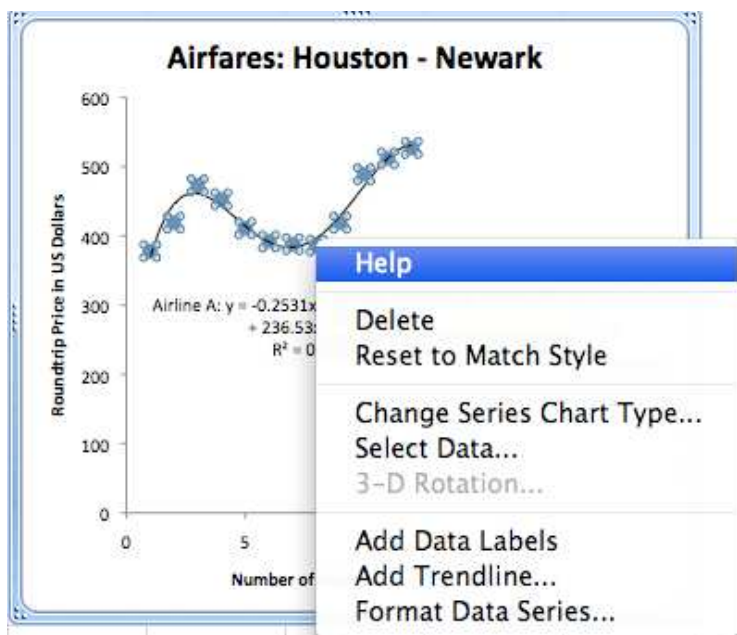
You will be inserting a second equation shortly, and it will be helpful to identify each equation. The quartic equation is given in a text box, so you can edit the text. Click inside the textbox and use the arrow keys to move the cursor to the desired location. Double click on the text box containing the equation. Then Type in "Airline A". Here's the result:



The graph area is getting a little busy, so you can resize it. Click on the little dots in the middle of the bottom border and pull down. This will give a larger graph area. Here's the result.



Now we'll insert a graph for Airline B. The steps are different for the second data set. Start by putting the cursor on any data point on the scatterplot and press **Control** while clicking. Here's the result:



Choose **Select Data**.

We want to add a series, so click on **Add**. This is the resulting screen.

Select Data Source

Chart data range:

The Chart Data Range is too complex to be displayed. If a new Data Range is selected, it will replace all of the series on the Series Panel.

Switch Row/Column

Series

Airline A
Series2

Add Remove

Name:

X values:

Y values: {1}

Category (X) axis labels:

Cancel OK

You want to call this series “Airline B,” so type that in the **Name:** box. Now you need to specify the x and y values for this series. Put your cursor in the **X values:** box and click. Then move the cursor back to the spreadsheet and click and drag to highlight the numbers from 1 – 12 in your table of values for Airline B. Let go, and you’ll see that some corresponding cell information appears in the **X values:** box. You’ll repeat this process for the y values, but first delete the characters that are in the **Y values:** box (if there are any). Then put the cursor in the **Y values:** box. Highlight the y values in the spreadsheet. Let go and information containing cell locations for the y values appears in the box.

Select Data Source

Chart data range:

The Chart Data Range is too complex to be displayed. If a new Data Range is selected, it will replace all of the series on the Series Panel.

Switch Row/Column

Series

Airline A
Airline B

Name: Airline B

X values: =Sheet2!\$C\$4:\$C\$15

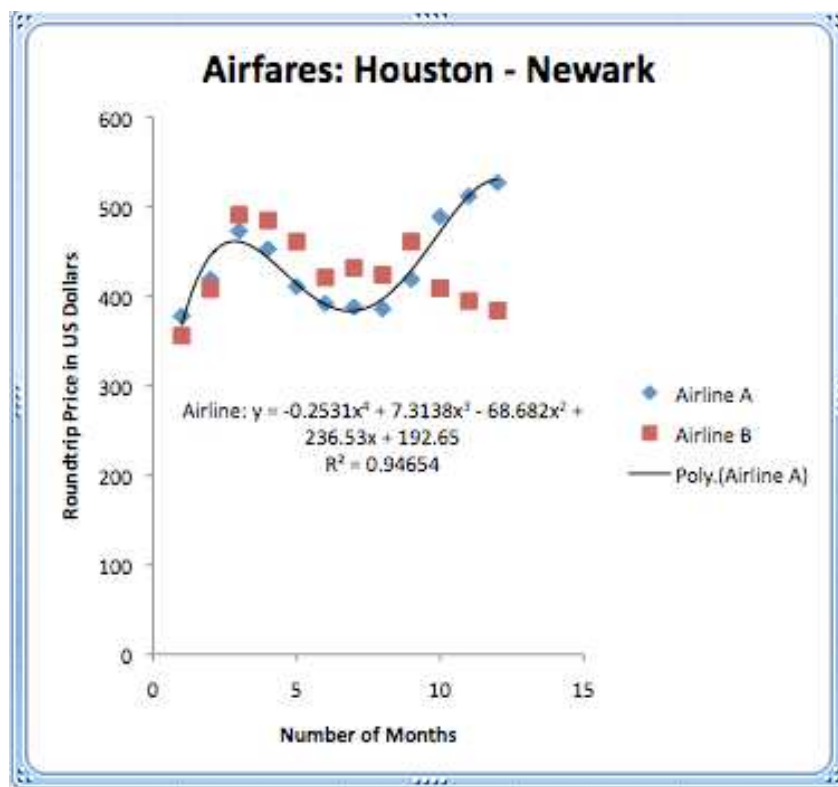
Y values: =Sheet2!\$D\$4:\$D\$15

Add Remove

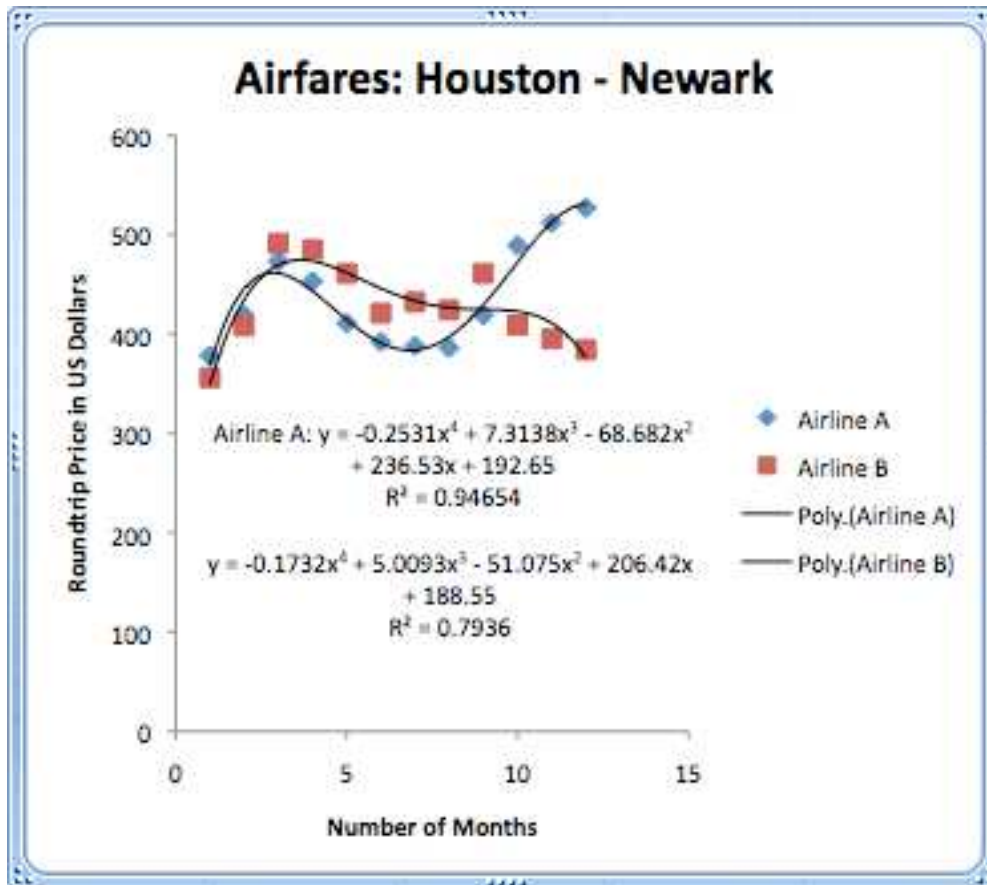
Category (X) axis labels:

Cancel OK

Click **OK**. The result is a chart that shows both sets of data. Note that Excel chose a different mark for the second series. The Legend already reflects the correct title for the second data set.

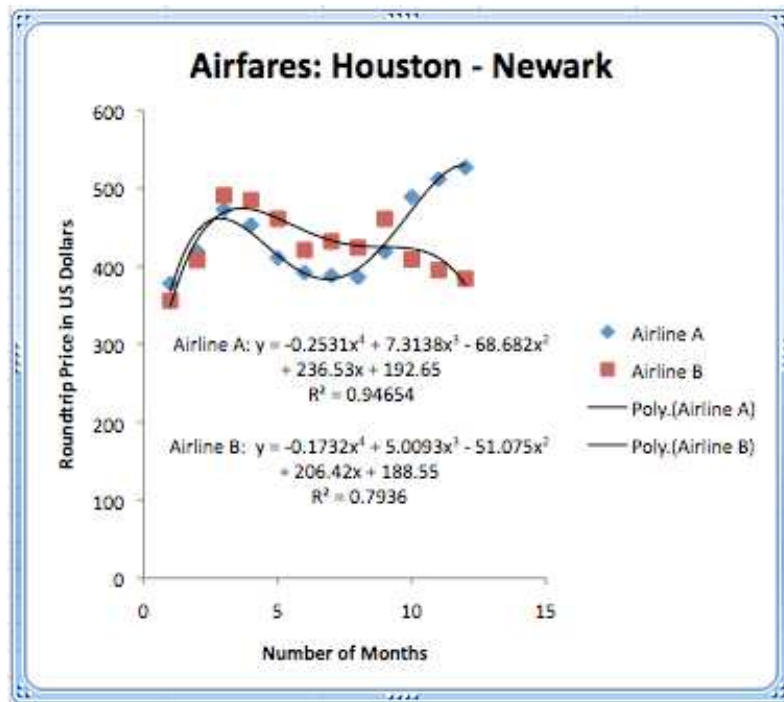


Finally, you need to find a quartic (4th order polynomial) trendline for the second data set. Repeat the steps for finding a trendline, but this time, put the cursor on one of the data points for the second data set (in this case, one of the orange boxes). Select 4th order polynomial from the **Type** menu and check both boxes on the **Options** menu to display the equation and the value for R^2 .



Click on the text box that contains the equation and the R^2 value and drag it so that it is below the first equation. The double click on the text box to edit the textbox containing the second equation so that it is labeled as the equation for Airline B.

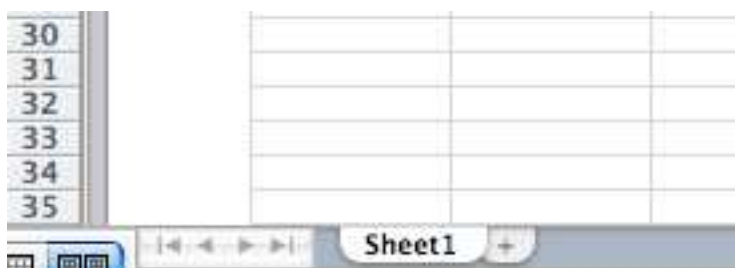
Here's the final product:



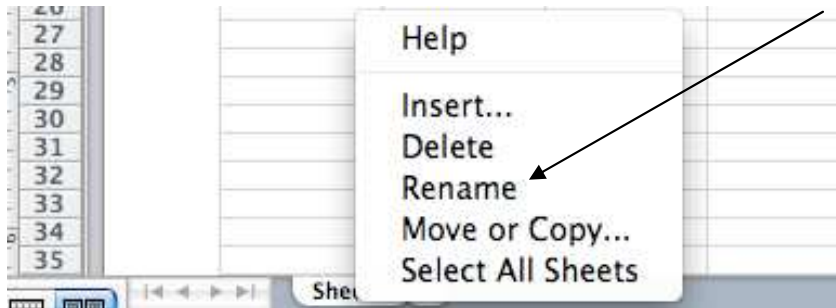
There are numerous videos on YouTube which demonstrate construction of charts in Excel. You may find those useful.

A few final notes about working with Excel.

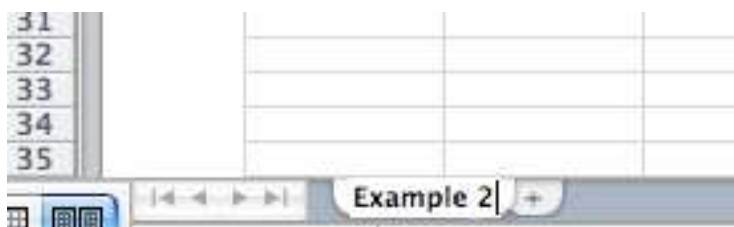
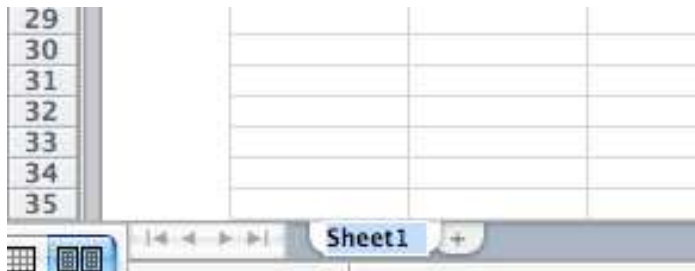
Each workbook in Excel contains or can contain more than one spreadsheet. It is often convenient to put related data in the same workbook on different spreadsheets. Look at the bottom of your Excel screen, and you should see a tab labeled **Sheet1**. There is also a tab with a little plus sign to the right of **Sheet1**.



You can name each sheet of your workbook. Earlier in this section, you worked through Example 2, so you could label the page containing your table of values and scatterplot as Example 2. This work will be one Sheet1 of your workbook. To rename this sheet, put your cursor on the tab that says Sheet1 and press **Control** while clicking. Select **Rename**.



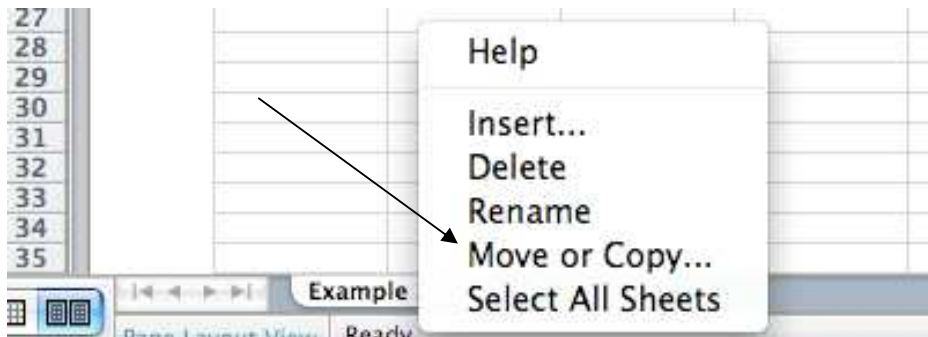
The sheet title is now highlighted, and you can type in a different name for the sheet. Call this sheet “Example 2”.



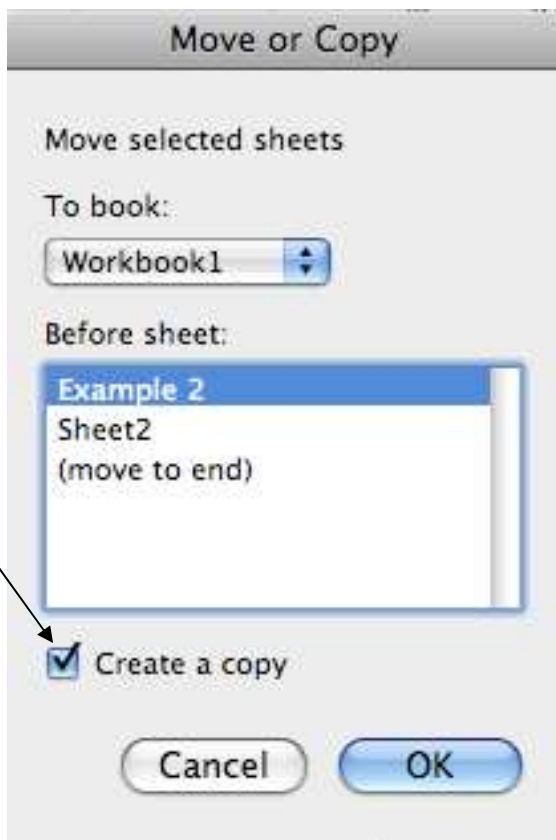
If you need to add an additional blank spreadsheet, click on the plus sign to the right of the tab you just named “Example 2”.

You may need to copy a spreadsheet entirely. For example, you have created a scatterplot and found a linear regression model for Example 2. Now suppose you want to find a different regression model. Instead of recreating all of your work, you can just copy the page containing all of your Example 2 work and then edit it.

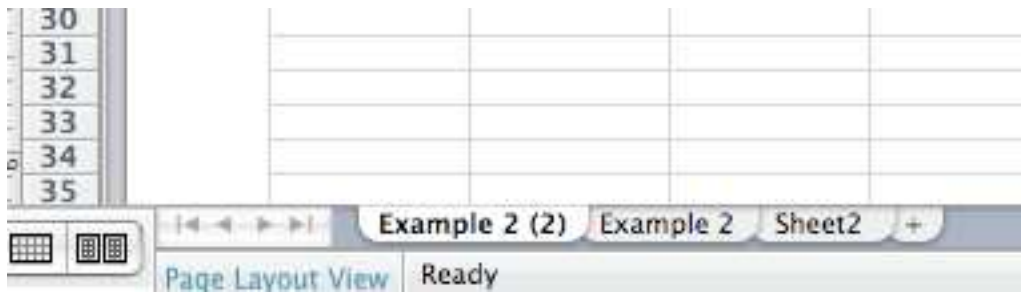
To do this, put your cursor on the tab labeled Example 2. The press **Control** while clicking.



Select **Move or Copy**. This menu allows you to rearrange the spreadsheets or to make a copy. Click the box that says **Create a copy**. Then click **OK**.

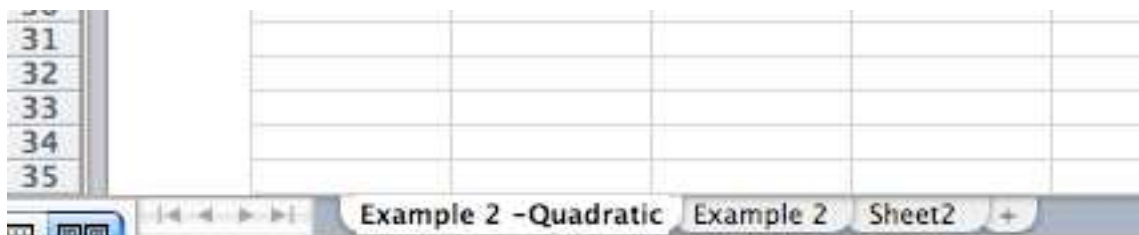
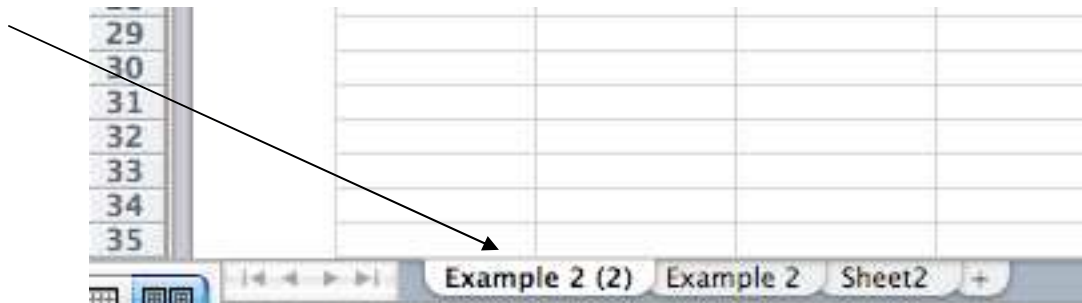


Here's the result:

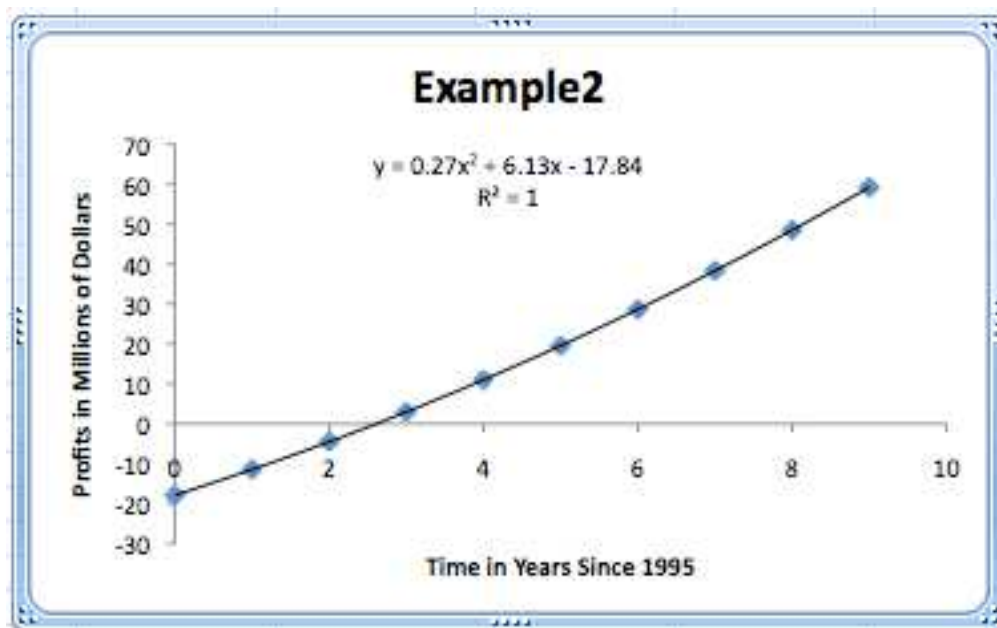


The copy is named Example 2 (2), but if you look at the content, the sheets are identical. You can now re-label the tabs as needed.

Now that you have copied Example 2, you can modify your work on one page. Suppose you also need to find a quadratic regression model for this data. Choose the page you want to edit, and rename the tab.



You can delete the regression line and insert a different regression equation. Put the cursor right next to the regression line (the line that Excel drew through the blue diamonds) and press **Control** while clicking. Select **Delete**, and the line, equation and R^2 should all disappear. Then you can add a new trendline.



You have successfully copied and edited the scatterplot.

You can also save your work. Click on **File** and select **Save As**. Give your work a name and put it in an appropriate folder. If you want to insert your work into a Word document, you can highlight, copy and paste the desired items into Word.