Trendlines and Graphing

By Joseph F. Lomax, Wayne Pearson, ENS Peter Ronald Bleday and John Schultz; all of the Chemistry Department, USNA.

Part III: An Exercise in Trendlines and Graphing

---

1. First, a Warning.

To do this exercise you must be sitting in front of a computer with the application Excel 2007 running. This write-up is not intended to be a keystroke-by-keystroke or mouse-click to mouse-click instruction set. The assumption is that you have gone through the **Getting Started** and the **Further Adventures in Excel 2000** tutorials that are on the **Excel 2000 Tutorial Homepage**. If you have not done these, do so now. Otherwise let us proceed onward.

2. Let's Get Started.

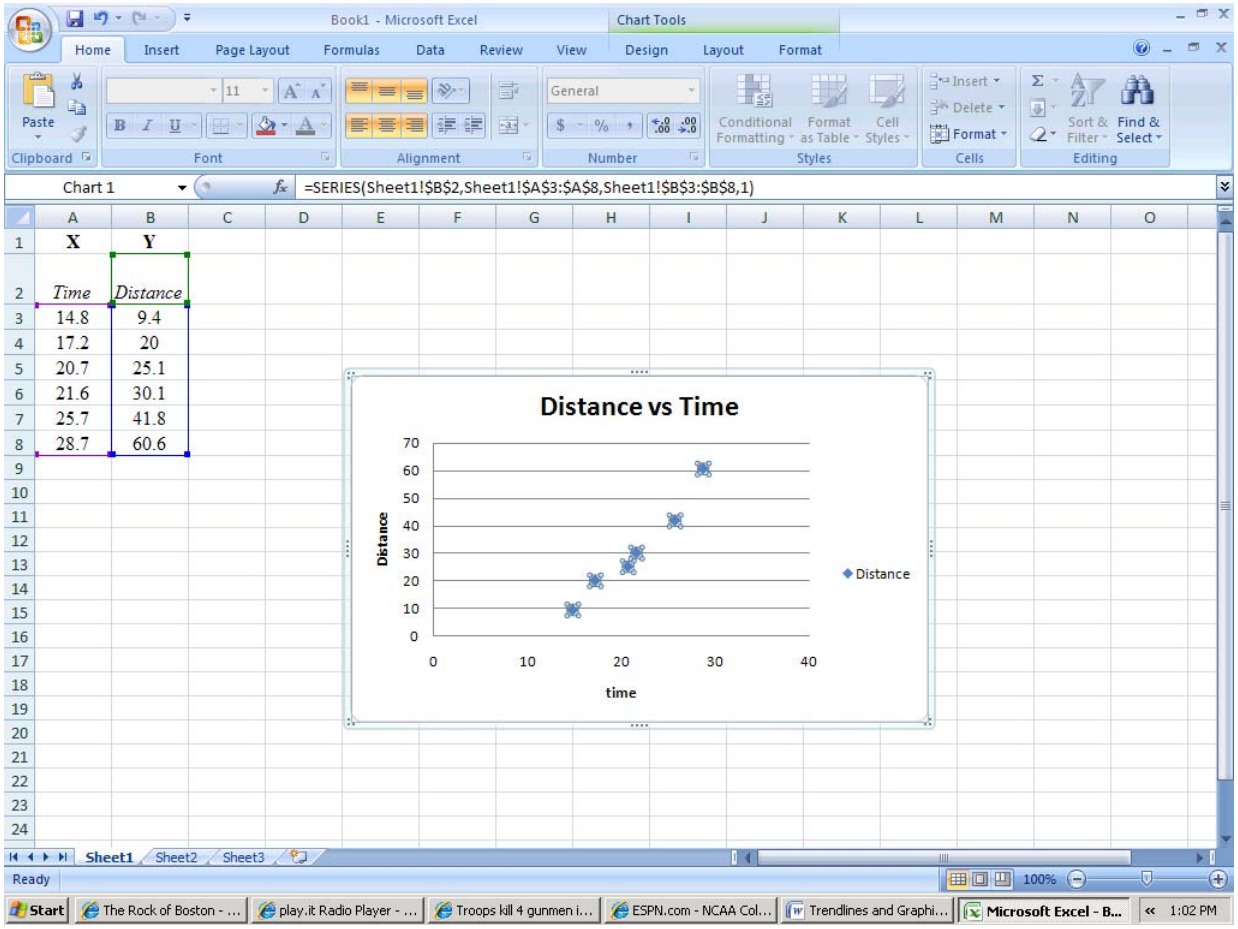
Keep in mind that Excel 2007 is a highly redundant system. There are at least two ways, and maybe more, to do the same thing. Larry Wall, the inventor of the programming language PERL, describes this redundancy as TIMTOWTDI (There Is More Than One Way To Do It). Your way of accomplishing a task need not match exactly what is described below and that is OK. The goal is to discover the physical relationship described by the data, not to get a notebook page and graph that match in every way what are in the images below.

3. Here goes.

Copy into the Excel 2077 sheet 1 the 6-pairs of xy data shown in the table below.

X	**Y**
Time	*Distance*
14.8	9.4
17.2	20.0
20.7	25.1
21.6	30.1
25.7	41.8
28.7	60.6

I will use the *A* and *B* columns for the raw data, you can, too. Then try to generate the graph by highlighting both columns of cells then clicking of the **Scatter/Chart** tab in the **INSERT** menu at the top of the screen. If you need a further reminder on how to make a graph, go back to Exercise 7: Making a Chart (a.k.a. Graph) in Excel from **Excel 2007 Getting Started**. See if you can make your graph look as much like the image below as possible, including the text (obviously, change the name).



If you end up with a strange looking graph, you may need to change the graph type. To do this you need to activate the use the tool bar on the top of the screen. Be sure to choose *XY,Scatter*, and not *Line* graph. These two graph types are not identical! Your markers for the plotted points should be either open squares or open circles, with no lines connecting the points. To change the marker style and remove any connecting lines, place the mouse cursor on one of the plotted points and click on the right mouse button. You should go to the bottom of the pop-up menu to *Format Data Series*. You may want to explore the menus a bit to see what is available.

4. The Significance of the Data?

This is an exercise and we don't really care about where the data came from. But if it makes you feel better, think of it as having been gathered by a mid timing a bug crawling along a meter stick.

5. Problem.

A problem common to experimental data is that it contains some random errors (sometimes called noise). The data you just entered has lots of noise! This results in plots which have a certain amount of scatter. If one is trying to fit the data to a straight line, it is not always obvious how the line should be drawn in. Two different mids, drawing in what they consider to be the best fit straight line through the data, could end up with significantly different results.

6. Solution?

Fortunately, there is a mathematical function that will unambiguously determine the best straight line through a field of data points, no matter how scattered. This function is called *Linear Regression* and will be used in a number of chemistry lab experiments. One does not need to understand how the function works to be able to use it.

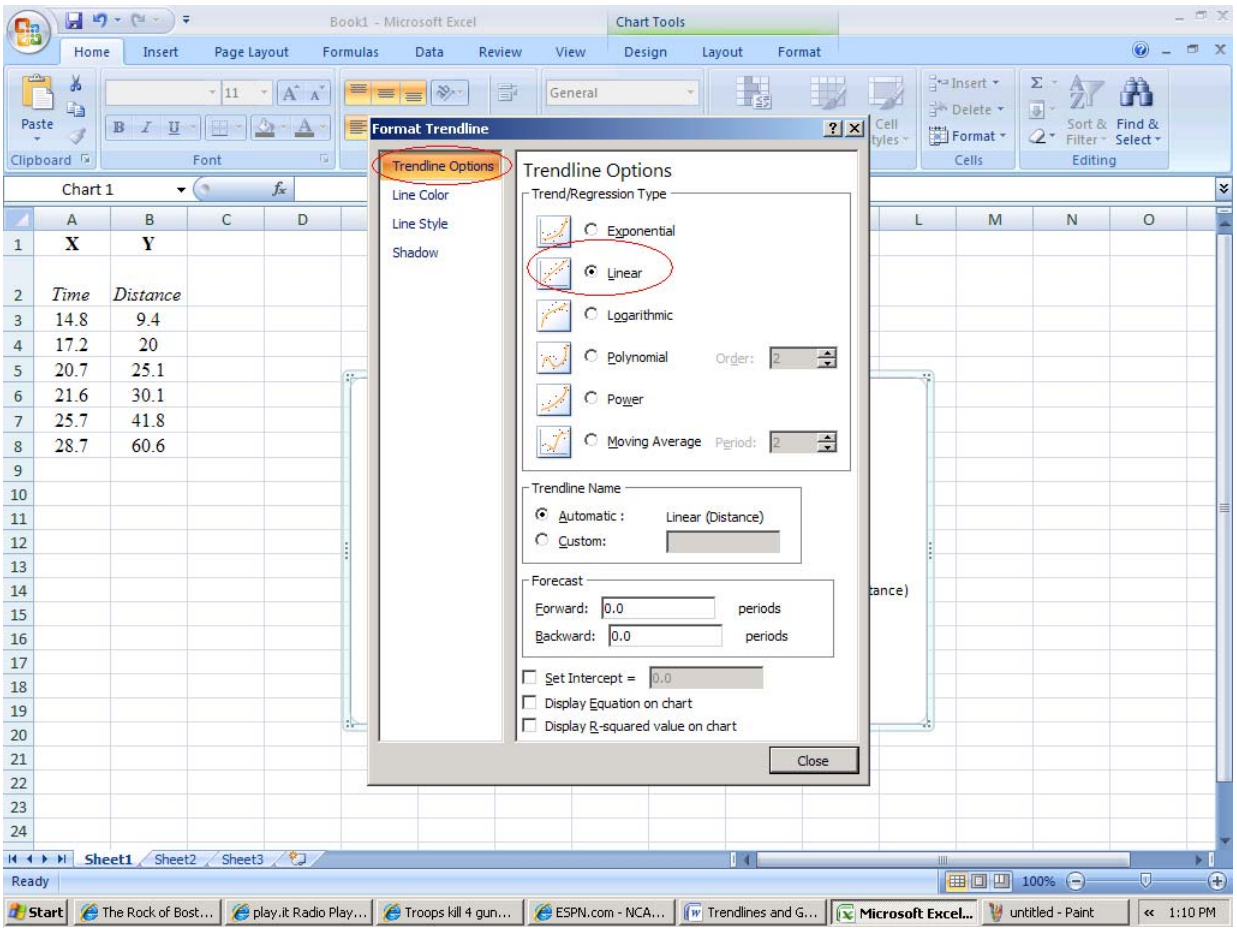
7. Graphing Calculator

The linear regression function is included in just about every calculator made, including the plebe TI-voyager. It is quite convenient to use and some instructors will have you use your TI-voyager in lab. The TI-voyager Guidebook has a description of how to make a plot like we are making. It breaks down the whole process by steps, keystrokes and gives representative displays as you go along.

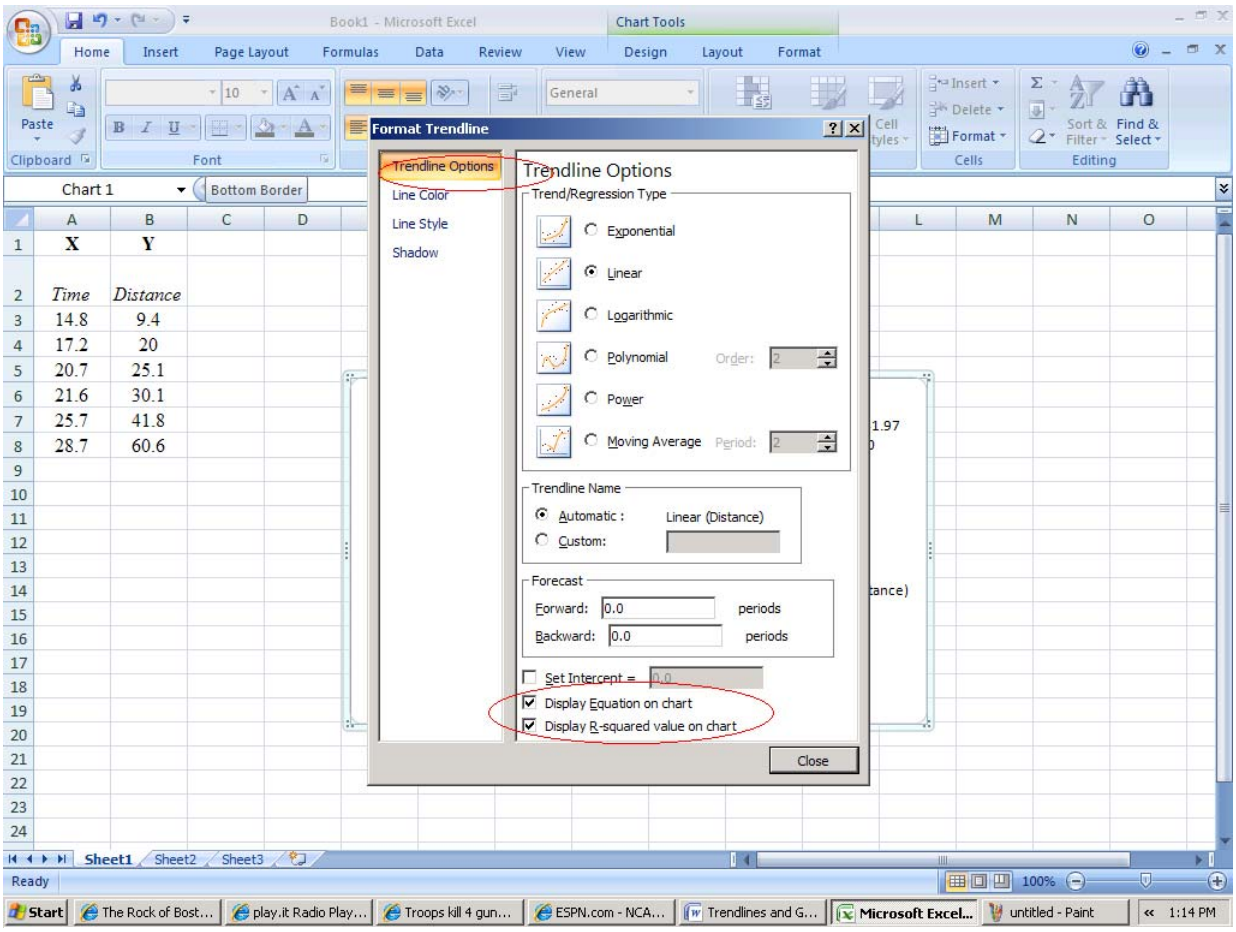
There is always a snake in the Garden of Eden. It is quite cumbersome to get a decent hard copy of your graph from TI-voyager. With GraphLink (an add-on costing another $50, thank you) you can connect your TI-voyager with your PC. Among other things, Graph-Link allows you to capture the screen of your calculator as a .tif file. A .tif file can be loaded into MS Paint and printed from there. As stated in a previous tutorial, the beauty of spreadsheets is their rapid and attractive formatting. The beauty of a calculator is its portability. The best world would have a way to connect these.

8. Practical.

To get to the Linear Regression menu activate the spreadsheet window, then right click on the data points on the chart. SELECT **Add Trendline** and under the tab labeled **Trendline options**. You want a create a linear regression so choose **Linear**. Then SELECT **Close** (see image below).



Now that you have a trendline, you need a display of the data for that trendline. To do this, right click on the trendline and SELECT **Format Trendline**. Once you are in this window SELECT the **Trendline Options** tab. At the bottom, check both **"display equation on chart"** and **"display R-squared value on chart"**. SELECT **Close**. Note: the data will automatically appear over your chart. It is difficult to read the data. So, click on the data, hold down the mouse button and move it to the upper right had corner of the chart. You could have put it anywhere on the chart window, but it is typically placed there. It is a good idea to *save* your file at this point so you won't lose anything if your roommate trips over the cord that connects your computer to the electrical outlet.



9. Caveat Computor

Excel's linear regression function will do exactly what you tell it. But you must tell it the right things. *Independent* means x-values, the ones you plot along the horizontal axis. *Dependent* means y values, the ones you plot on the vertical axis. If you screw up this convention, you will get an extremely wrong result. Except for the formatting of the numbers, your regression output should match that in the next image. If not, check to make sure you are adhering to the conventions. One more item, the y-intercept can be forced to go through zero, or the non-zero value can be computed. In general, you will want to do the latter and compute the value of the y-intercept. This is the default setting in the regression menu, so you usually do not need to do anything here, unless you change the default settings.

10. Analysis?

Taking the generic equation of a straight line to be $y = Ax + B$, the A value (slope) is the 3.4097. The B value (y-intercept) is -41.972. Excel's linear regression function has given us the R-squared value. The closer this number is to 1, the better the fit of the line. Though we won't be using this number. The fact that the r-squared value is close to 1

gives us some comfort. In most cases, the slope will be the quantity sought for, but sometimes we will be interested in the y or x-intercept as well. The x-intercept is obtained by solving the above equation for x and setting y to zero. This gives *-B/A* as the x-intercept.

11. XY Pairs have a Meaningful Relationship.

Now you will generate a second series of y values that correspond to the slope and intercept that the Linear Regression function gave us. In other words, using a few choosen x-values, we use the equation:

y = 3.4097x + (-41.97)

to calculate a second set of y-values. These values are based on an equation of a straight line and will give us the desired result.

Excel 2007 has a linear regression function that is exceptionally easy to use. You need only use the cleverly named functions **=slope()** and **=intercept().** Type in cell E2 "=slope(" and it will prompt you with "=slope(**known y's,** known x's). What it wants you to do is to specify the value ranges for the y-values (in this case "pressure"). You know they want the y-values first because "**y-value"** is **bold-faced** and "x-value" is not. Once you SELECT **B3:B8**, it will **bold-face "x-value".** Now, highlight **A3:A8**, close the parentheses and ENTER. Poof! You have the slope of the linear regression line. Put the label "slope = " in cell D2. Performing the same sequence in G2, but using **=intercept()** will give you the value of the linear regression intercept in the cell. Put the label "slope = " in cell D2.

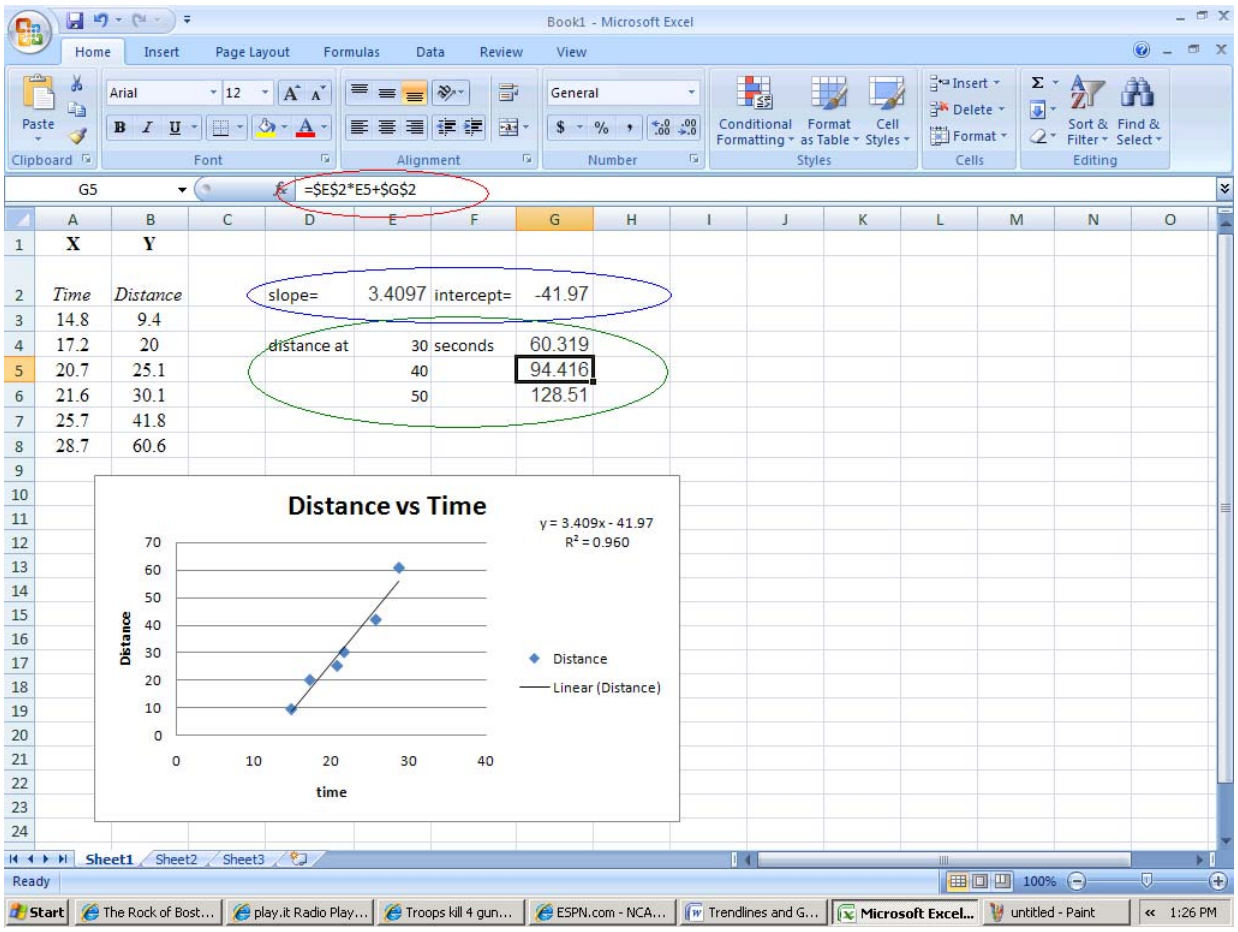
Now, using these values and the equation for a line, you can easily predict the distance during any given time (in seconds).

In cell D4 type **"distance at"**. And in cells E4:E6 enter **30, 40, 50** seconds respectively. In cell F4 type **"seconds"**.

Now you are ready to calculate distances (y-values). You need a formula that keeps the values for slope and y-intercept constant while allowing the x-values to change. The appropriate formula for this type of calculation is:

=$E$2*E4+$G$2

Dollar signs are added to the addresses of the slope and intercept to make them *absolute addresses*. This is done so that the copy and paste buttons will not change them.



12. A Cause for a Celebration?

Your spreadsheet should look something like the one in image above. If it does then Congrats -- You Made it! If it doesn't, and you can't figure out what's wrong, then Seek Help from your instructor or by going to the Chemistry Resource Room, Michelson 100, for help. In any case, save your notebook for a final time.

13. An Additional Exercise or Two or Three

1. Modify your graph to extrapolate the straight line to the x-axis! You will be required to do this in a future chemistry lab experiment report. Right click on the trendline and change "Prediction-Backward" until the line hits the x-axis
2. Starting with the graph in Exercise 1, extrapolate the straight line just a little bit so that it overshoots the highest experimental x-value by a small amount. Adjust the x-axis and y-axis scales so that the upper end of the straight line does not touch the graph boundaries. You may need to use manual rather than automatic scaling to get this accomplished. This enhances the appearance of the graph and is commonly done for publication in technical journals!
3. Suppose you mess up and interchange your x and y-values when doing this exercise.

What is the relationship of the incorrect slope to the correct slope? How are the incorrect x and y-intercepts related to the correct x and y-intercepts?

This page was last revised on 19 Aug 2008.