Plotting data in Microsoft Excel

Eric D. Black
California Institute of Technology
v1.0

1 Introduction

Microsoft Excel is ubiquitous. You have probably used it before. You probably already have it on your computer, or at least something compatible with it like Open Office. And finally, you will probably see it again after you leave this class. For these reasons you may decide to plot your data in Excel, and that is why I am including this guide.

One thing you should be aware of before you start is that unlike Kaleidagraph, Mathematica, Matlab, or many other more science-centric programs, Excel does not allow you to fit your data to nonlinear or arbitrary functions. It does have some simple fitting routines, which we will cover in this lab, but for general use it is much more limited than your other options.

All of the following examples were done in the 2011 version of Excel on a Macintosh.

2 Opening a data file

Choose Open from the File menu. A dialog box will appear where you specify the location of the file. After you specify the file, three more dialog boxes will come up in succession, allowing you to specify how the file is to be interpreted, *i.e.* how the columns are separated (delimited), how the individual entries should be formatted, etc. (Figures 3, 4, and 5).

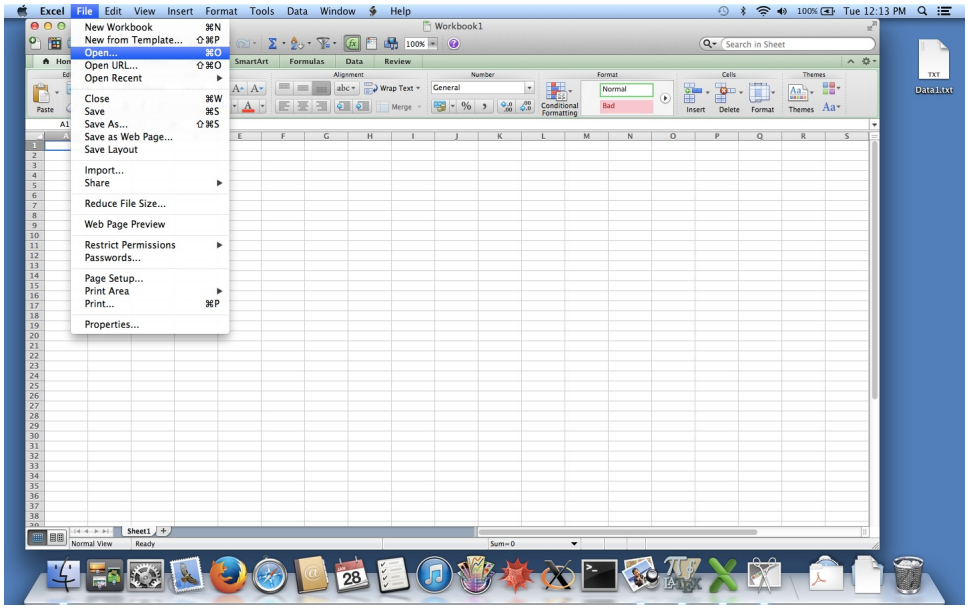

Figure 1: Opening a data file from Microsoft Excel.

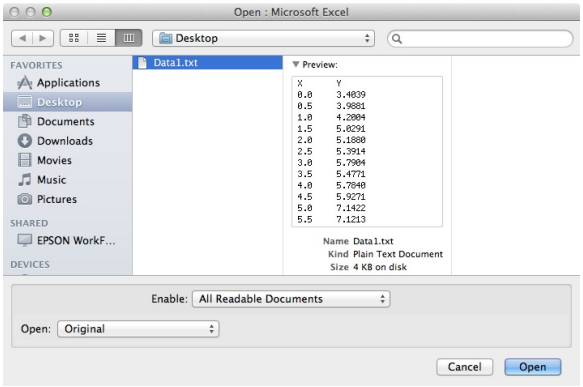

Figure 2: Specifying which file to open. You may have to choose "All Readable Documents" from the "Enable" pulldown menu before you can access a plain-text file.

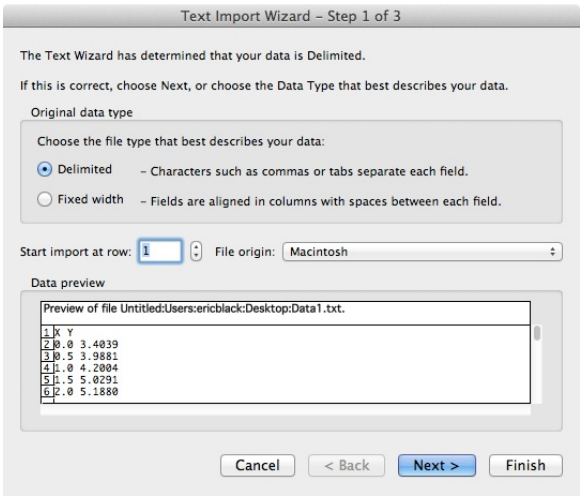

Figure 3: The first of three dialog boxes.

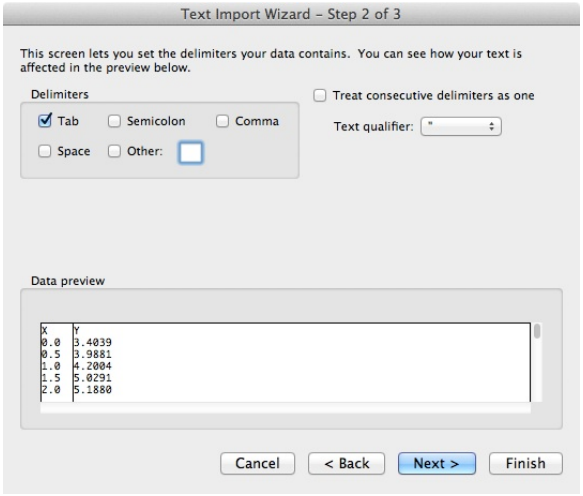

Figure 4: The second of three dialog boxes.

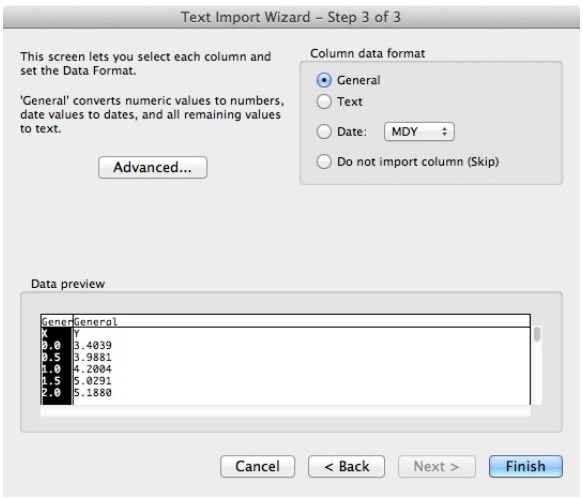

Figure 5: The third of three dialog boxes.

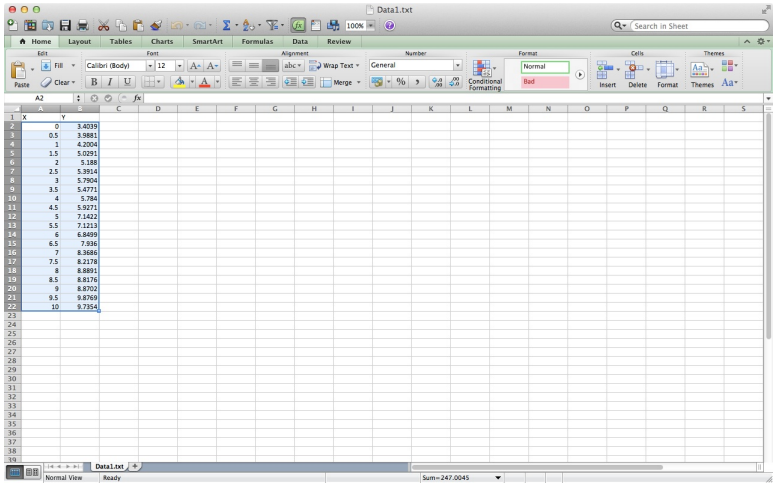

Figure 6: What you get.

3 Making a plot

Plots are called "Charts" in Excel. To make one, first select the data you want to plot, as shown in Figure 6. Then, click on the word Charts at the top of the window. The green bar that contains the words Home, Layout, Tables, Charts, etc. is called the "Ribbon," and each option in the Ribbon contains a menu of its own. In the Charts menu there is a list of icons grouped together under the heading "Insert Chart." These options have pictures associated with them and are largely self-explanatory, and the two that are of interest to us are Line and Scatter. Unfortunately, both have severe limitations.

Line plots, as you might expect, connect each data point with a line. This is sometimes useful if you have a lot of data points and aren't planning on marking each with its own little symbol, but it's less than ideal for the kinds of plots you will be making in this lab. Line plots also don't support horizontal error bars, so if you are planning on including those, Line is right out.

Scatter plots don't "connect the dots" with a line from point to point, and they allow vertical and horizontal error bars. However, you don't get the option of one or the other. You *must* accept both vertical and horizontal error bars.

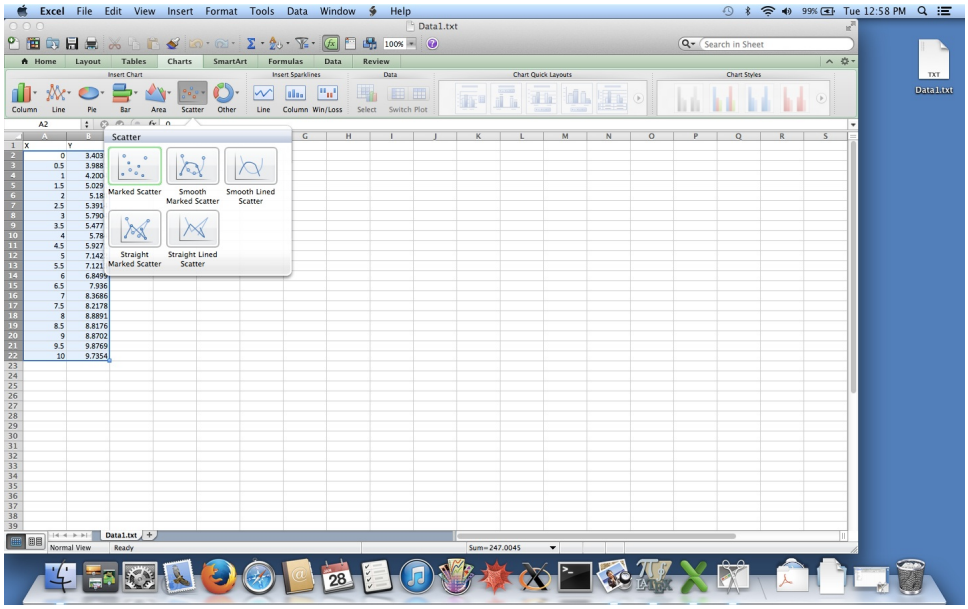

Figure 7: Click on the Charts tab in the Ribbon, then select Scatter as the type of plot.

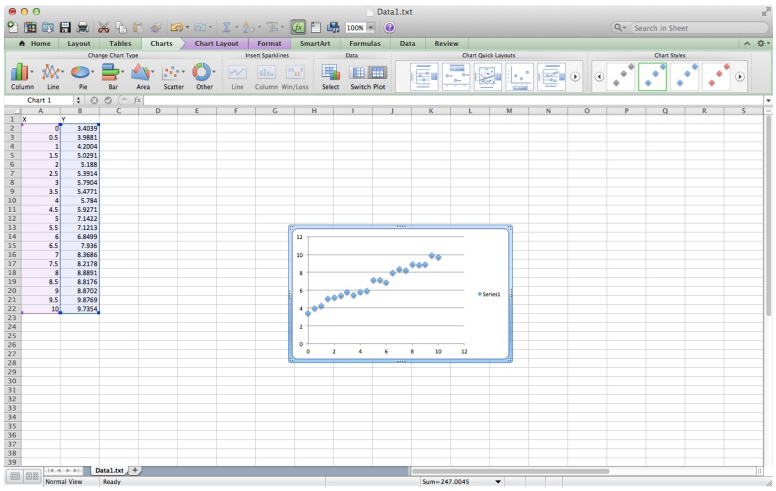

Figure 8: This is what you get.

4 Adding error bars

To add error bars, click on "Chart Layout" in the ribbon, then follow the directions from there.

Excel is a little quirky about this. First of all, it automatically sets your error bars to some value it thinks they should be, so you will have to override this manually. Second, as I mentioned above, your options are somewhat limited. With a scatter plot, you get both vertical and horizontal error bars, whether you want them or not. If you want only vertical error bars, you must choose "Line Plot" as your style and then put up with the lines that connect your data points.

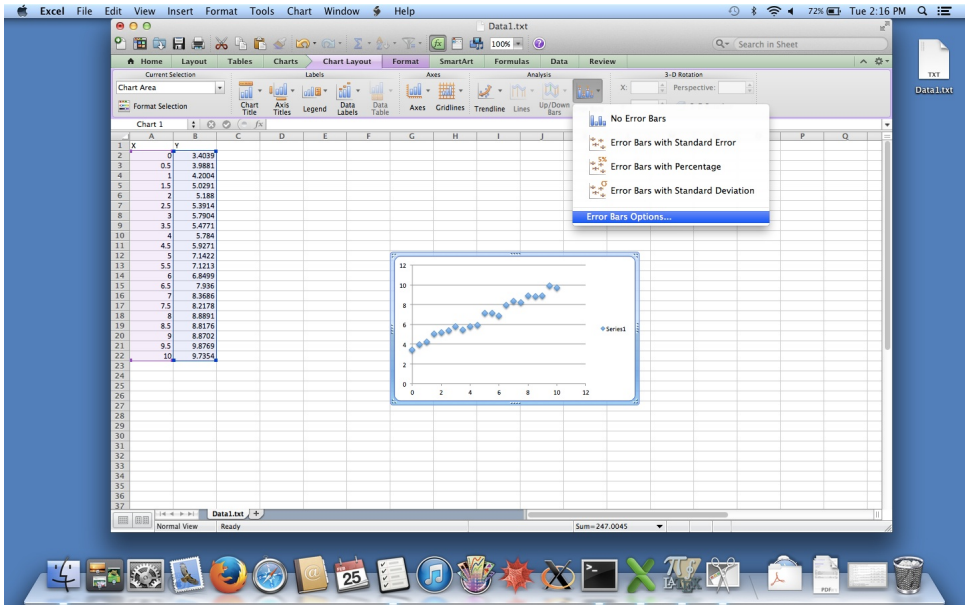

Figure 9: Click on the Chart Layout tab next to the Charts tab in the Ribbon, then select Error Bars. To specify the value of your error bars, as opposed to letting Excel decide what they should be, go straight to Error Bar Options.

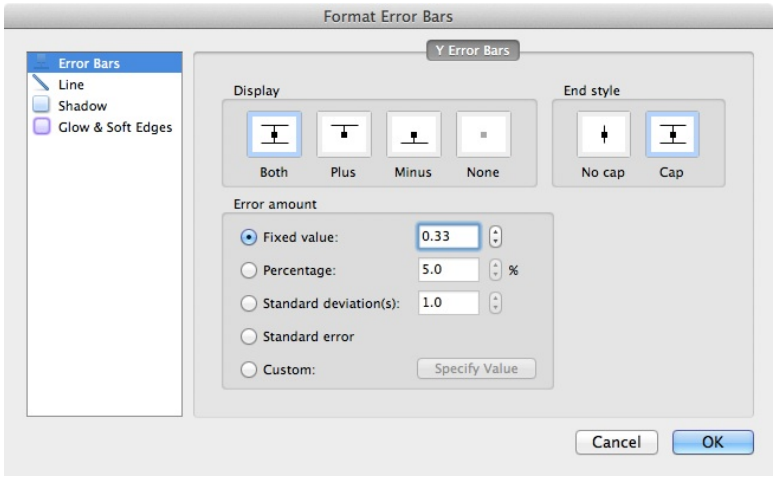

Figure 10: Even though Excel puts error bars on the independent variable (X), you don't get to specify them yet.

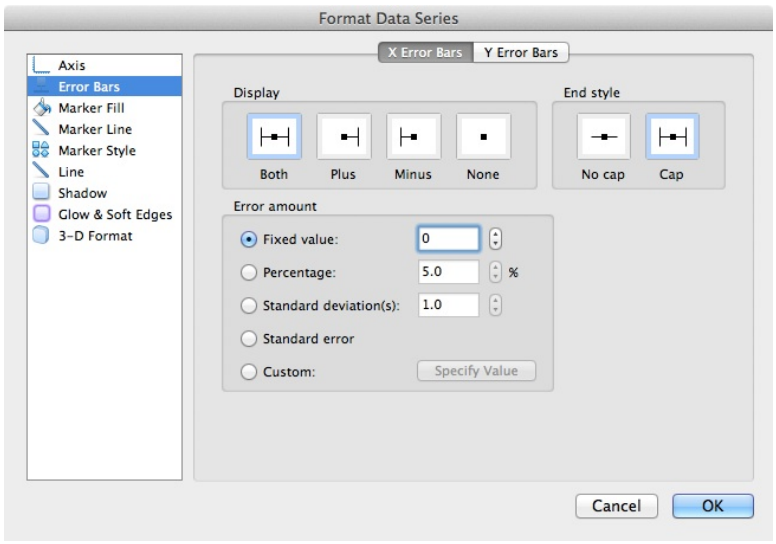

Figure 11: To set the X error bars, click on one of the data points in your plot. That will bring up this dialog box, which now contains a tab at the top that will allow you to adjust either the X or the Y error bars. To suppress the X error bars, choose Fixed Value, and set it to zero.

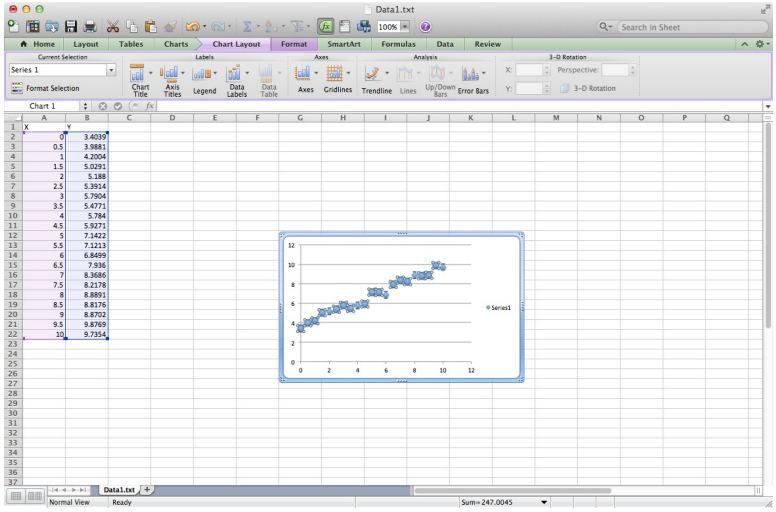

Figure 12: This is what you get.

5 Adding a theory curve

You can add a theory curve by creating a new data set in your existing sheet and adding it to your plot.

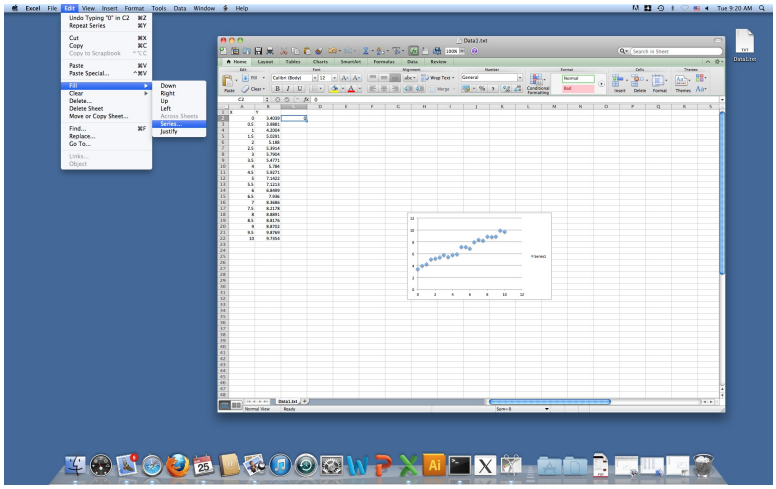

Figure 13: To create a new data set, first enter the first X-value in a particular cell, in this case zero in cell c2. This will "seed" a new series of independent variables. Then choose Fill → Series from the Edit menu. This will bring up a dialog box allowing you to specify the parameters of the series.

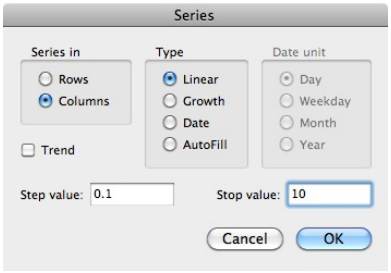

Figure 14: Choose a Step value of 0.1, a Stop value of 10 (these should be self-explanatory), and check the Columns radio button under Series in. This last step ensures that your series fills cells downward instead of accross, which is the default.

If all went well, you just created a series to define your independent variable. To calculate a list of your theory function values, select the first cell to the right of the first number in your new series, and type the following into it

```
=0.5*c2+3.5
```

What's going on here is this, starting a cell entry with an equals sign is code for excel to evaluate whatever mathematical expression you type in next, in this case a guess for a linear function that will fit (or approximate) your data. The slope here is 0.5, the Y-intercept is 3.5, and c2 refers to the value in the second cell in column c, the location of one of your X-values.

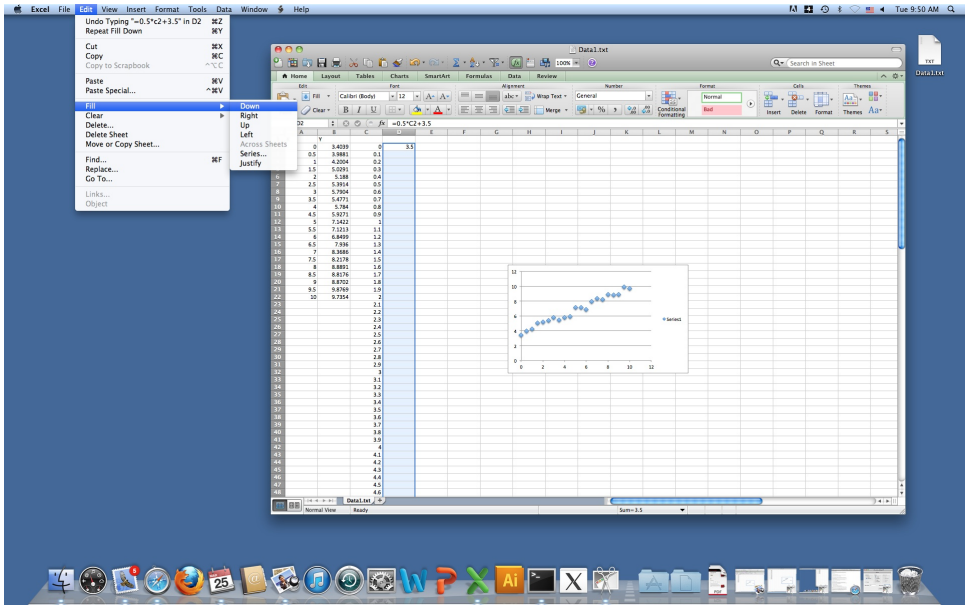

Figure 15: After you have defined the first point in your theory curve by typing your formula (=0.5*c2+3.5) into cell d2, select the whole column you want your theory curve to be in, and choose Fill → Down. This will propagate your formula through all the cells in your selection, updating the dependent-variable reference c2 to c3, c4, etc. as appropriate.

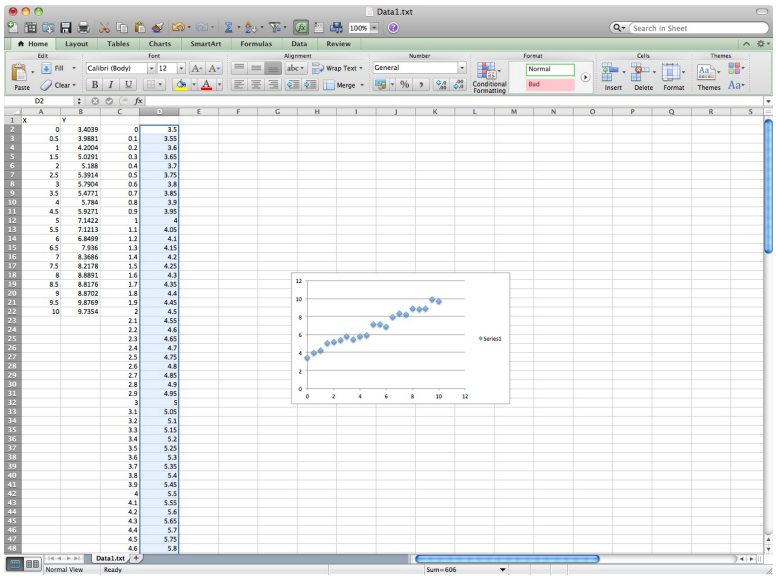

Figure 16: Now you should have your theory function evaluated in column D.

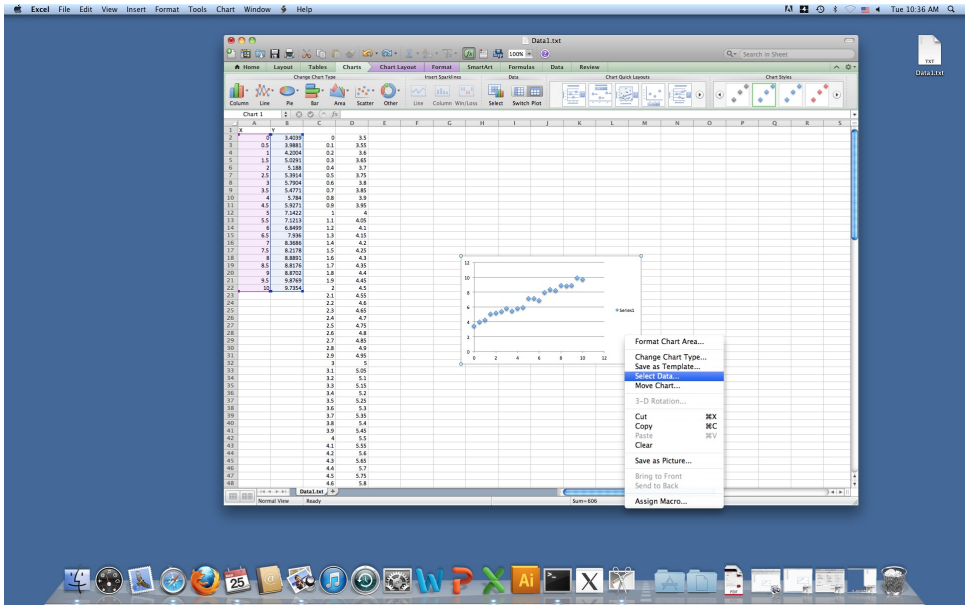

Figure 17: Now that you have defined a table with the values for your theory curve, it is time to add it to your plot using Add Data. Control-click on your plot (right-click for you Windows users), and choose Select Data...

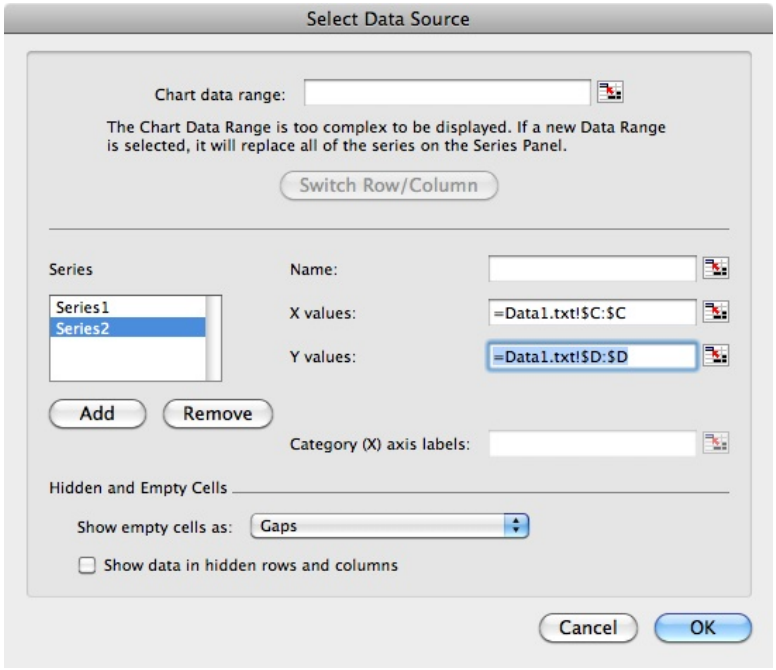

Figure 18: Click Add, then select the appropriate columns for the X and Y values of your theory function.

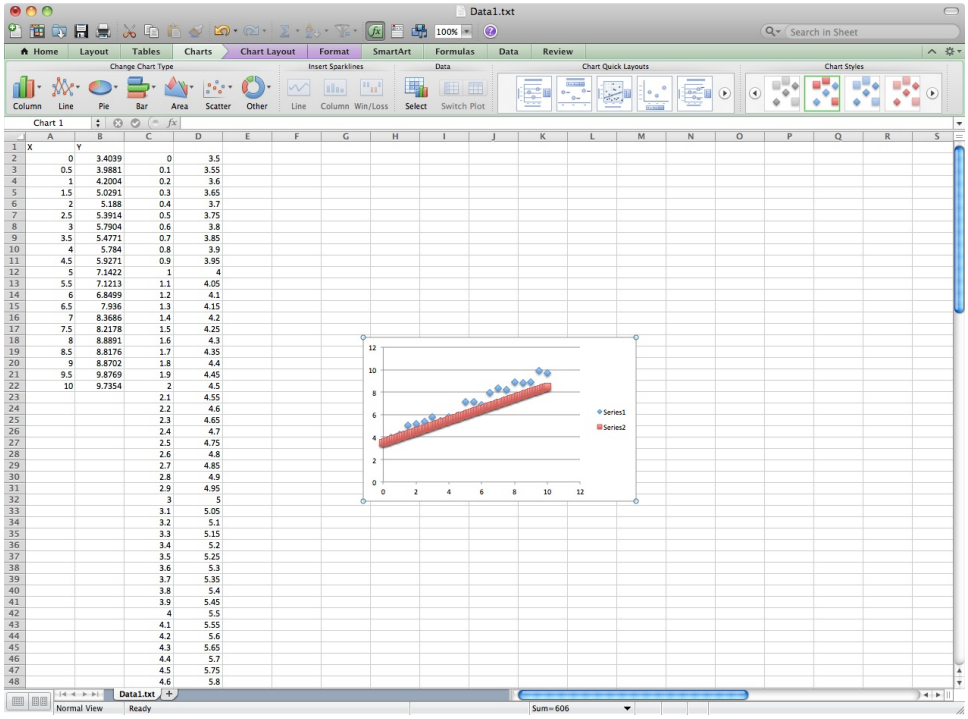

Figure 19: This is the result. You can clean up the formatting as you like.

6 Residuals

You know enough now to be able to do this without me walking you through it. Define a new column; calculate your residuals; and make a new plot. It should look something like this.

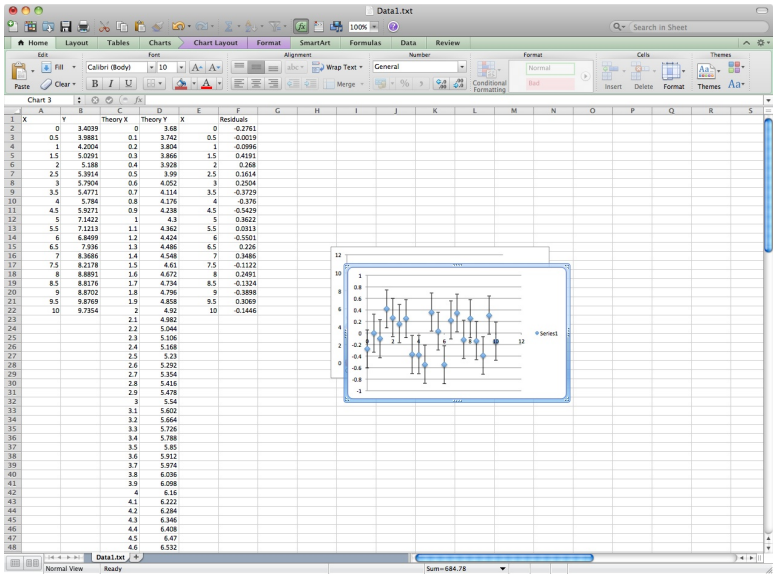

Figure 20: Residuals plotting in Excel.

7 $\widetilde{\chi}^2$ test

Once you know how to perform calculations in Excel, this is pretty easy. The most straightforward way to do this is,

1. Define a new column (in this case Column G) that will contain the squares of the residuals, divided by the square of the uncertainty (σ). The formula will look like this,

   ```
   =(f2/0.3)^2
   ```

   and you will want to fill down to complete your table.

2. In a separate cell (g24 in our example) sum the terms in Column G, and divide by the total number of data points to get the reduced chi squared. The syntax for summing all the values in cells g2 through g22 and then dividing by 21 is

   `=sum(g2:g22)/21`

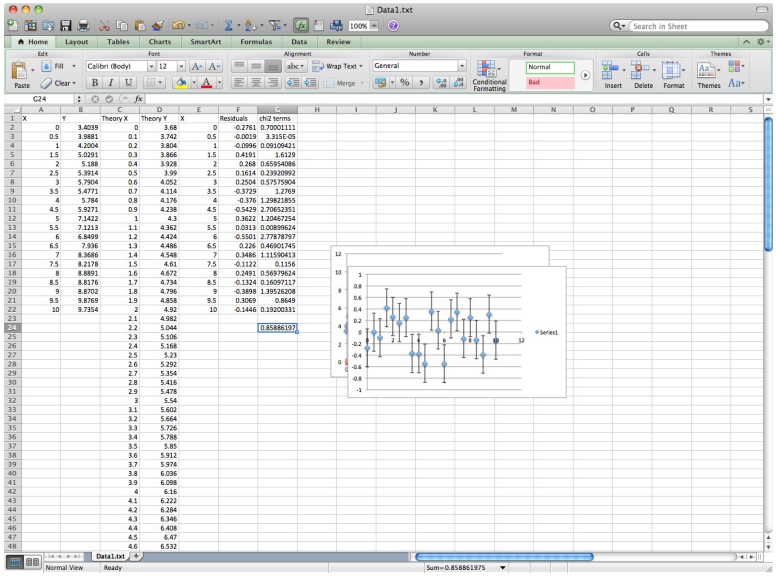

Figure 21: Calculating the reduced chi squared is simple, once you have the residuals.

8 Automated fitting

Excel does have a least-squares fitting feature, but it's not called by that name. It is part of the Trendlines package.

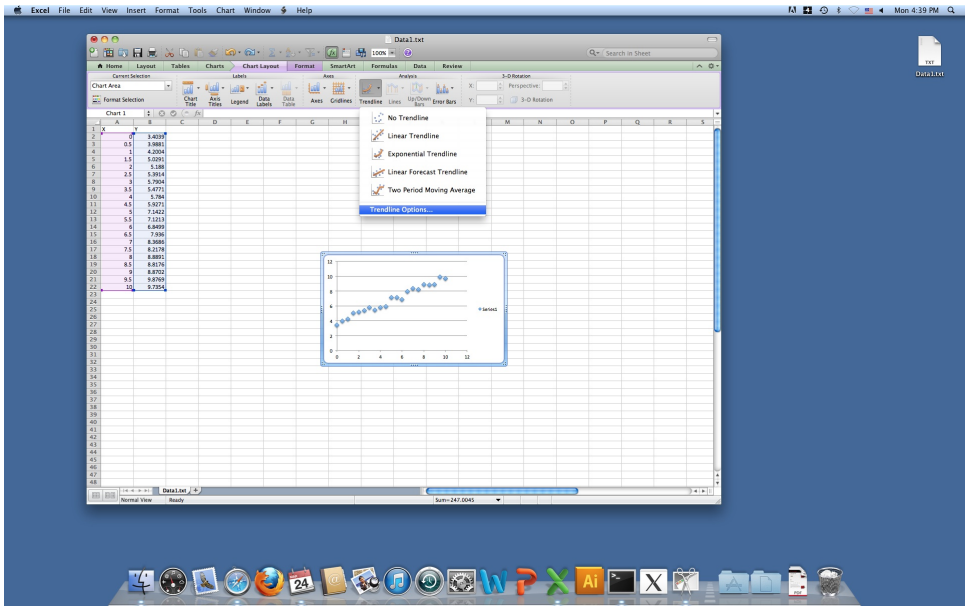

Figure 22: To perform a linear least-squares fit, click on the "Trendline" button in the "Chart Layout" ribbon. You will want to choose "Trendline Options..." from the pulldown menu.

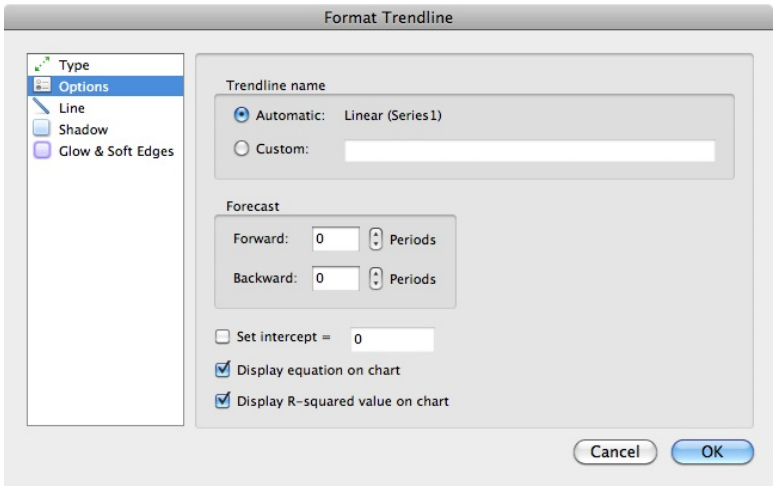

Figure 23: Under Options, select "Display equation on chart." You can also have it display the R-squared value of the fit, which is not the same thing as the reduced χ^2 but serves a similar purpose, *i.e.* it provides a measure of the goodness of the fit.

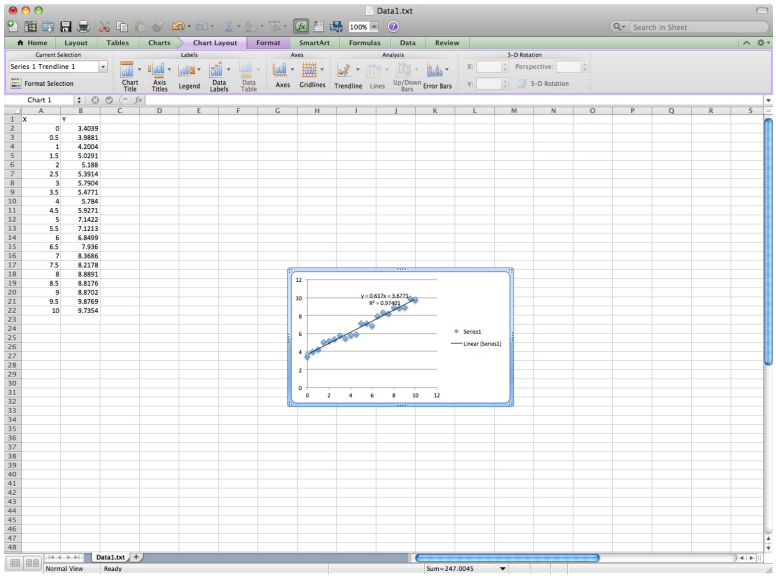

Figure 24: This is what you get.