# An introduction to plotting data

Eric D. Black
California Institute of Technology

February 25, 2014

## 1 Introduction

Plotting data is one of the essential skills every scientist must have. We use it on a near-daily basis for visualizing our data, drawing conclusions from it, and communicating our results to the rest of the world. In the old days people drew plots by hand, with pen and ink, on graph paper. This was very easy to learn and quick to do, but it was also very limited. If you wanted to make anything more than a minor change to your plot, changing the scale of the axis, for example, or (horrors!) changing it from linear to logarithmic, you basically had to start over and draw a new one. Today we have tools to make all these tasks, as well as the analysis of our data, much easier, and there is no better time than now for you to start learning them.

This week we are going to start learning the basics of plotting and analyzing data, but I'm going to deviate from the usual way this is taught in most freshman-level courses. Many of you have come to this class with some background in research and already know how to generate plots using a particular software package. Others will be entirely new to the subject. Of those who already have some skill, moreover, the particular software you know how to use will vary widely. For this reason, I'm not going to mandate a one-size-fits-all approach. Instead, I'm going to focus on *what* to plot more than on *how* to do the plotting, and I'm going to let you choose what program to do it in.

Along with this handout you should also have a second one describing how to do the exercises in Kaleidagraph, and that is the program I recommend you use if you are new to the subject. If you are already fluent in another program and want to use that, you are free to do so.

## 2 Single-variable plots

Here, we will focus on two-dimensional plots of single-variable data. For the purposes of our discussion, let's assume you have a set of data points of the form,

$$\{x_i, y_i\}$$

where $i$ runs from 1 to $N$, the total number of data points. Moreover, let's assume you suspect there is a connection between the numbers, *i.e.* that $y_i$ depends on $x_i$ somehow.

In this context, $x_i$ is called the *independent variable* and is the one you can control. You have a knob, for example, that allows you to set a particular value for $x_i$ at will, and then measure the outcome of some experiment that that is connected to it. This outcome is $y_i$, and we will call this the *dependent variable* because it is dependent on $x_i$. The obvious thing you will want to do is plot a graph of $y_i$ vs. $x_i$. This graph will be more than a simple illustration. As Edward Tufte points out [1],

> At their best, graphics are instruments for reasoning about quantitative information. Often the most effective way to describe, explore, and summarize a set of numbers - even a very large set - is to look at pictures of those numbers.

**LABORATORY EXERCISE 1:** Import the following data set into your program, and make a plot of $y$ vs. $x$.

| X | Y |
|---|---|
| 0.0 | 3.4039 |
| 0.5 | 3.9881 |
| 1.0 | 4.2004 |
| 1.5 | 5.0291 |
| 2.0 | 5.1880 |
| 2.5 | 5.3914 |
| 3.0 | 5.7904 |
| 3.5 | 5.4771 |
| 4.0 | 5.7840 |
| 4.5 | 5.9271 |
| 5.0 | 7.1422 |
| 5.5 | 7.1213 |
| 6.0 | 6.8499 |
| 6.5 | 7.9360 |
| 7.0 | 8.3686 |
| 7.5 | 8.2178 |
| 8.0 | 8.8891 |
| 8.5 | 8.8176 |
| 9.0 | 8.8702 |
| 9.5 | 9.8769 |
| 10.0 | 9.7354 |

If you don't want to type it in by hand, you can download it at

*http://http://www.pma.caltech.edu/∼phy003/labs/Data1.txt*

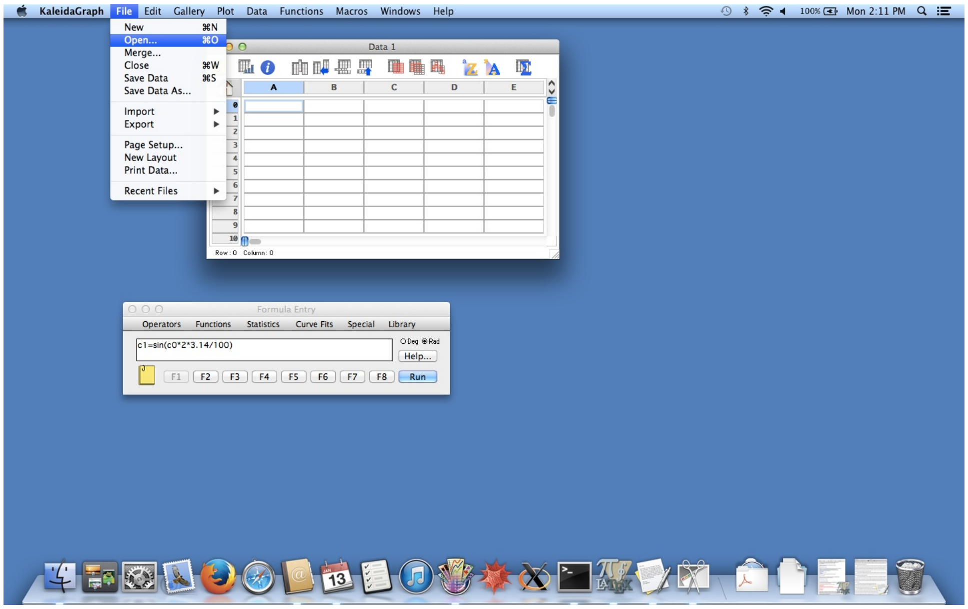
## 2.1 How to do this in kaleidagraph

To reiterate, you don't *have* to use Kaleidagraph to do these exercises. This section is optional. If you would rather use another program you are free to do so, whether on your own computer or one of the lab's. I am including these instructions because I think Kaleidagraph is the easiest to pick up in a short period of time and carries with it the least amount of frustration for a beginner. (There is some level of frustration in all of them.) It is also a full-featured tool that any experimentalist - physicist, biologist, etc. - can use at any stage of his or her career to generate just about any publication-quality plot you need.

KaleidaGraph

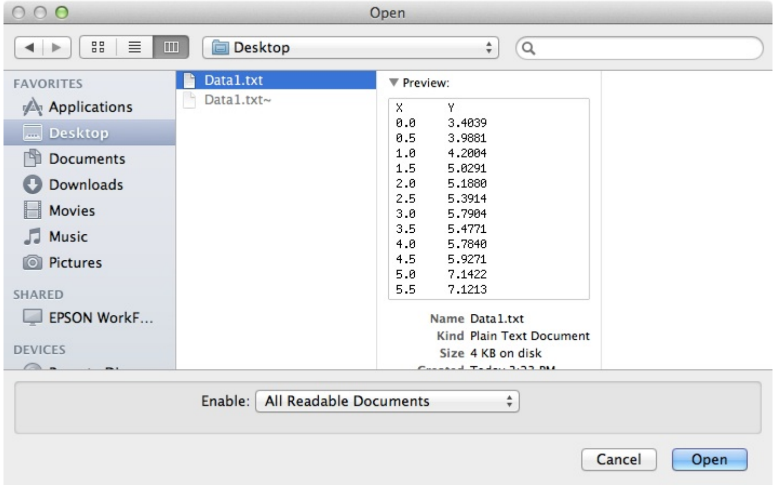
This is what the Kaleidagraph icon looks like. If you are using one of the imacs in the lab, it should already be in the dock. After it starts up, you should see two windows, one for data entry and another for forumla entry. (If one of these does not show up, go to the Windows menu in the menu bar and choose “Show Data” or “Formula Entry” to make it appear.)

At this point you can type your data into the cells of the data window, or you can open an existing data file.



c1=sin(c0*2*3.14/100)

Figure 1: How to open a data file in kaleidagraph. After you choose Open from the File menu, a dialog box will appear that allows you to choose the file you want to open.

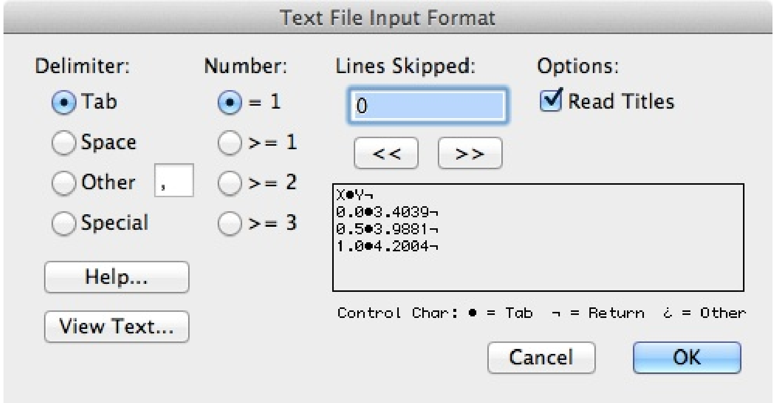


```
X       Y
0.0     3.4039
0.5     3.9881
1.0     4.2004
1.5     5.0291
2.0     5.1880
2.5     5.3914
3.0     5.7904
3.5     5.4771
4.0     5.7840
4.5     5.9271
5.0     7.1422
5.5     7.1213
```

Name Data1.txt
Kind Plain Text Document
Size 4 KB on disk

Enable: All Readable Documents

Figure 2: The dialog box for choosing a file to open.



Delimiter: ● Tab ○ Space ○ Other , ○ Special

Number: ● = 1 ○ >= 1 ○ >= 2 ○ >= 3

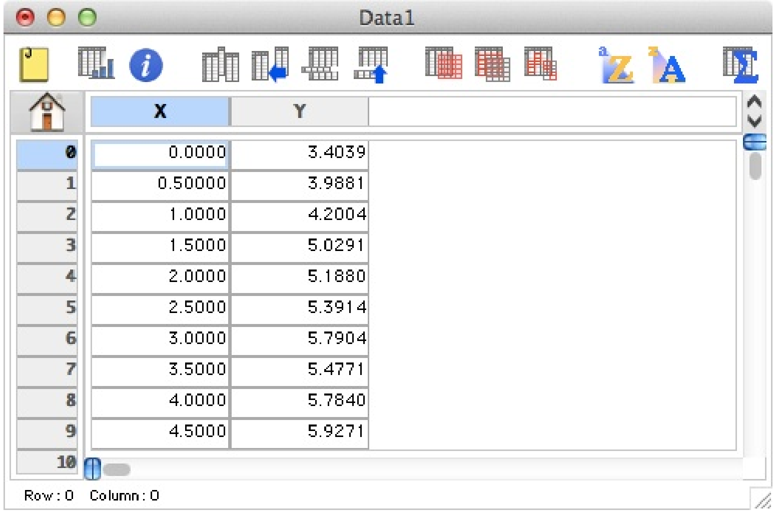
Lines Skipped: 0

Options: ☑ Read Titles

```
X•Y¬
0.0•3.4039¬
0.5•3.9881¬
1.0•4.2004¬
```

Control Char: • = Tab ¬ = Return ¿ = Other

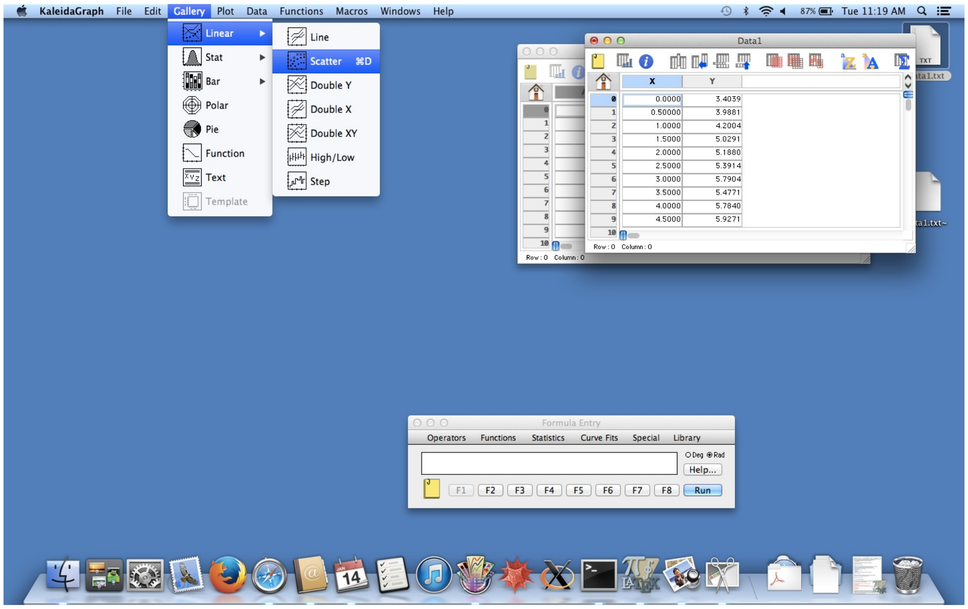
Figure 3: Once you have chosen the file to open, another dialog box appears allowing you to specify how it's read. Most of these are self-explanatory. Note the option to read the first line of data as column titles.



|  | X | Y |
|---|---|---|
| 0 | 0.0000 | 3.4039 |
| 1 | 0.50000 | 3.9881 |
| 2 | 1.0000 | 4.2004 |
| 3 | 1.5000 | 5.0291 |
| 4 | 2.0000 | 5.1880 |
| 5 | 2.5000 | 5.3914 |
| 6 | 3.0000 | 5.7904 |
| 7 | 3.5000 | 5.4771 |
| 8 | 4.0000 | 5.7840 |
| 9 | 4.5000 | 5.9271 |
| 10 |  |  |

Row : 0 Column : 0

Figure 4: The data looks like this once you have opened it. Note how the titles for the columns are preserved, if you choose that option.


KaleidaGraph File Edit Gallery Plot Data Functions Macros Windows Help
Tue 11:19 AM
Linear
Stat
Bar
Polar
Pie
Function
Text
Template
Line
Scatter ⌘D
Double Y
Double X
Double XY
High/Low
Step
Data1
Formula Entry
Operators Functions Statistics Curve Fits Special Library
Deg Rad
Help...
F1 F2 F3 F4 F5 F6 F7 F8 Run

Data1

|  | X | Y |
|---|---|---|
| 0 | 0.0000 | 3.4039 |
| 1 | 0.50000 | 3.9881 |
| 2 | 1.0000 | 4.2004 |
| 3 | 1.5000 | 5.0291 |
| 4 | 2.0000 | 5.1880 |
| 5 | 2.5000 | 5.3914 |
| 6 | 3.0000 | 5.7904 |
| 7 | 3.5000 | 5.4771 |
| 8 | 4.0000 | 5.7840 |
| 9 | 4.5000 | 5.9271 |
| 10 |  |  |

Row : 0 Column : 0

Figure 5: To make a plot, pull down the Gallery menu and choose the type you want. For a set of discrete data points, Scatter (under the Linear submenu) is a good choice.

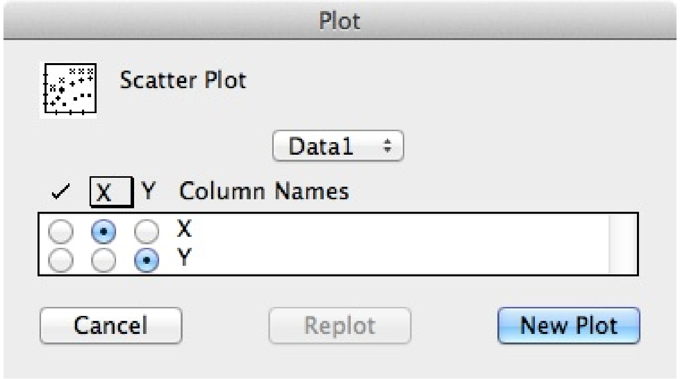


Figure 6: Once you have chosen the type of plot you want, a dialog box will come up to set which column to use as the X and Y axis. Click the appropriate buttons, and then click New Plot.

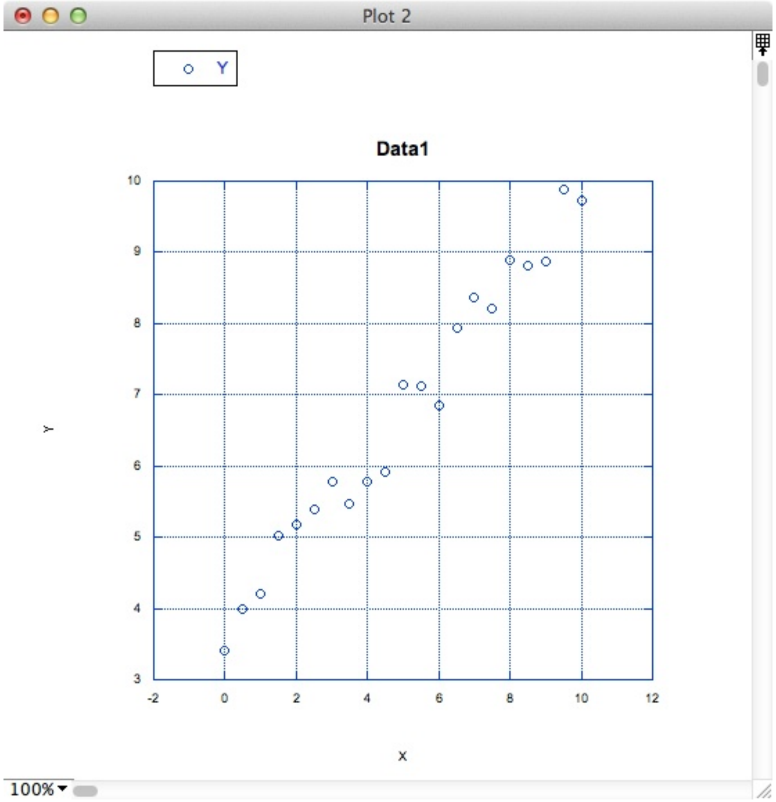


Figure 7: And here is the result.

# 3 Error bars

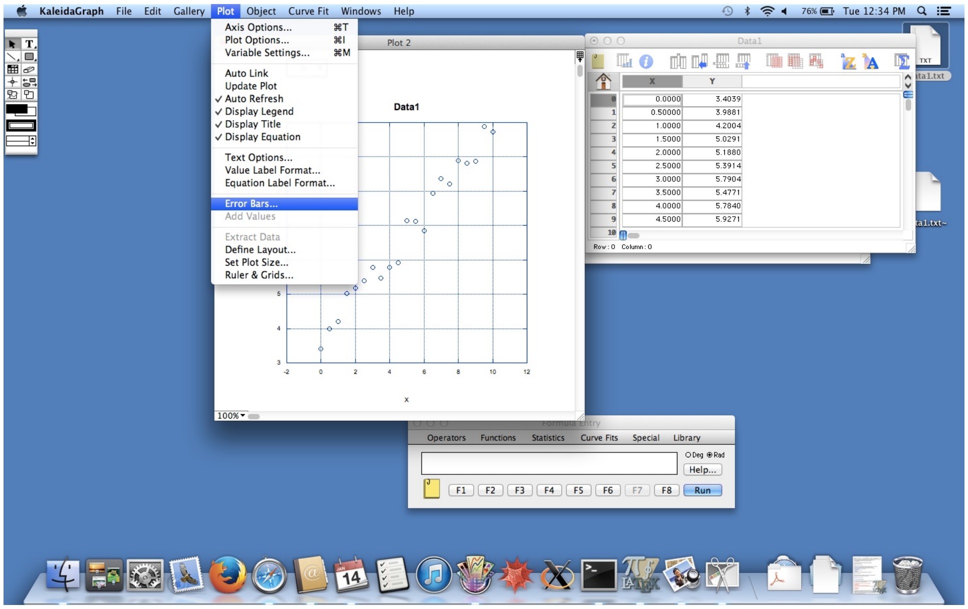
You should have started on your reading from *Taylor* by now, and you have probably even done the first homework set. That is your introduction to uncertainty and what it means, so I am not going to go over that material again here. This section is about how to graphically represent your uncertainty on your plot with *error bars*, little lines above and below your data points whose lengths show to the uncertainty in the data point itself.

Any decent software will allow you to put error bars on your data points, and it should also allow you to specify them any way you want.

**LABORATORY EXERCISE 2:** Put error bars on your plot. Use $\sigma = 0.33$ for the uncertainty in the Y value of each point.

## 3.1 How to do this in Kaleidagraph

Pull down the Plot menu, and from it choose the Error Bars... option. If this is greyed out, click on your plot window and then try again. Follow the instructions from there.



Data1

|  | X | Y |
|---|---|---|
| 0 | 0.0000 | 3.4039 |
| 1 | 0.50000 | 3.9881 |
| 2 | 1.0000 | 4.2004 |
| 3 | 1.5000 | 5.0291 |
| 4 | 2.0000 | 5.1880 |
| 5 | 2.5000 | 5.3914 |
| 6 | 3.0000 | 5.7904 |
| 7 | 3.5000 | 5.4771 |
| 8 | 4.0000 | 5.7840 |
| 9 | 4.5000 | 5.9271 |
| 10 |  |  |

Row : 0 Column : 0

Figure 8: To put error bars on your data points, pull down the Plot menu, and choose Error Bars. This will pull up a dialog box allowing you to specify whether to put them on the X or Y values (horizontal or vertical), and then a second dialog box allowing you to specify what their values should be.

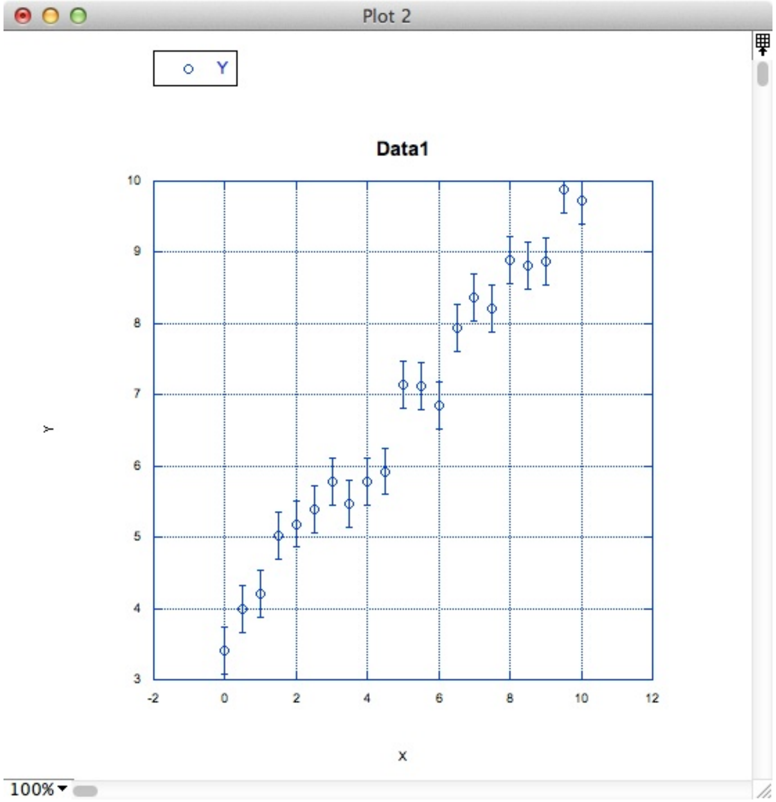


Figure 9: And the result.

# 4 Comparison with a theory curve

Very often you want to plot a theory curve along with your data and see how well the two agree. All modern software will include an automatic routine that will do this for you, but before you haul off and click that button, I want you to understand how it works. I don't want it to be a black box for you.

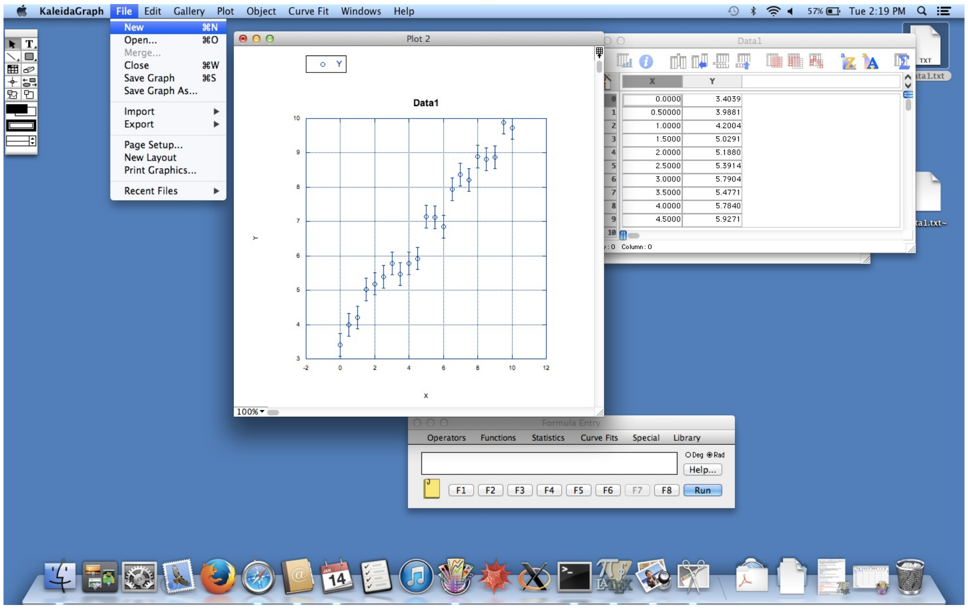
**LABORATORY EXERCISE 3:** Estimate the slope and y-intercept of a line that passes through these data points. Use the form

$$f(x) = Ax + B$$

and add this line to your plot. Adjust $A$ and $B$ as necessary until the agreement looks good.

## 4.1 How to do this in Kaleidagraph

Create a new data window by selecting New from the File menu.



|  | X | Y |
|---|---|---|
| 0 | 0.0000 | 3.4039 |
| 1 | 0.50000 | 3.9881 |
| 2 | 1.0000 | 4.2004 |
| 3 | 1.5000 | 5.0291 |
| 4 | 2.0000 | 5.1880 |
| 5 | 2.5000 | 5.3914 |
| 6 | 3.0000 | 5.7904 |
| 7 | 3.5000 | 5.4771 |
| 8 | 4.0000 | 5.7840 |
| 9 | 4.5000 | 5.9271 |
| 10 |  |  |

:0 Column : 0

Figure 10: Open a new data window. This is where you will generate the data for your theory curve.

Click the top of the first column to highlight it, then pull down the Functions menu, and select Create Series.

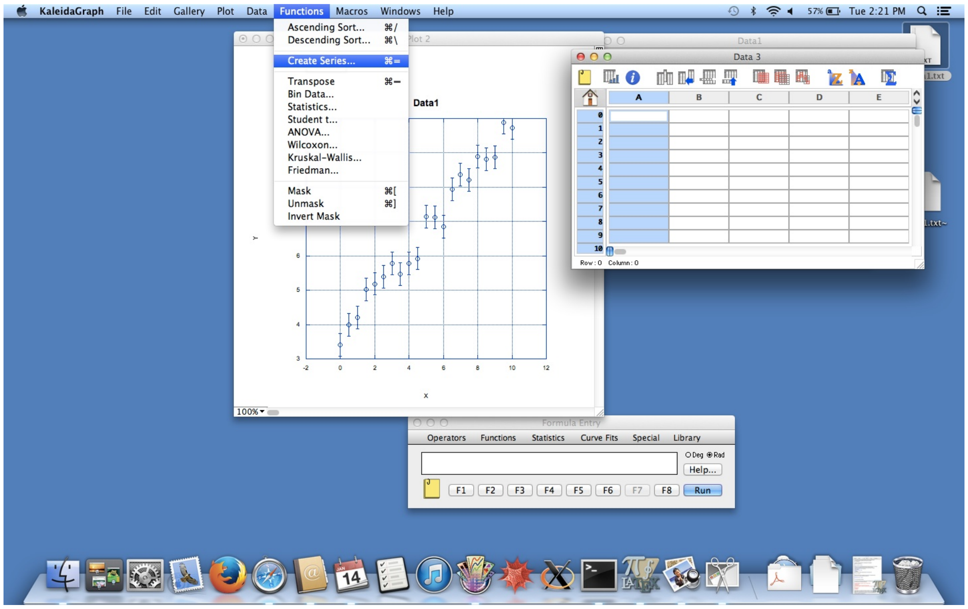
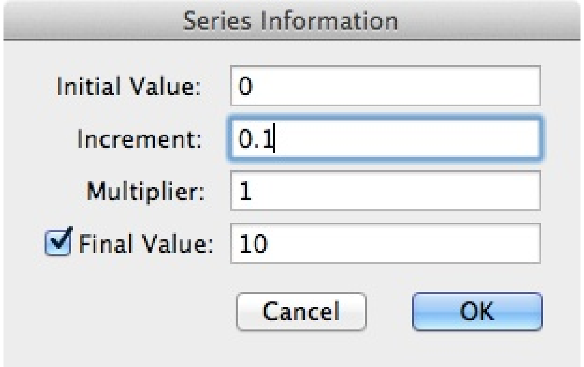


Figure 11: Highlight the first column, then pull down the Functions menu and select Create Series.

You can create a series with any resolution you want. A hundred points with a spacing of 0.1 will work fine for our purposes.

Now click on the Formula Entry window, and type in the theoretical formula you want to use to generate your theory curve. This uses a standard spreadsheet-style notation, where the first column is “column zero” or c0, the second is c1, etc. Figure 13 shows how we would write our $f(x) = Ax + B$, with guesses of 0.5 for the slope $A$ and 3.5 for the y-intercept $B$.

Formula entry is really very powerful and is almost a full scripting language. We will just be doing basic things with it here. A full treatment of it is beyond the scope of this lab, but you should be aware that it has a great deal of capabilities.



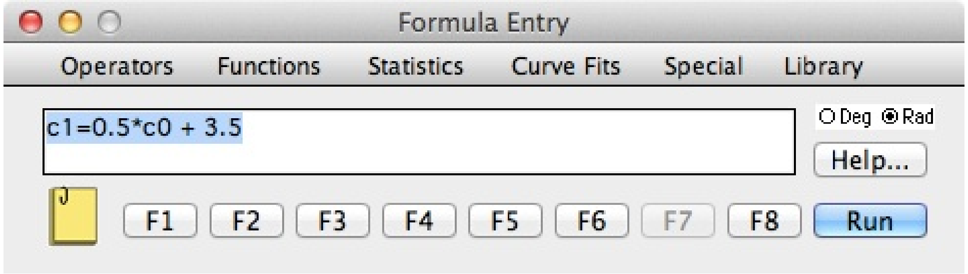
Initial Value: 0

Increment: 0.1

Multiplier: 1

☑ Final Value: 10

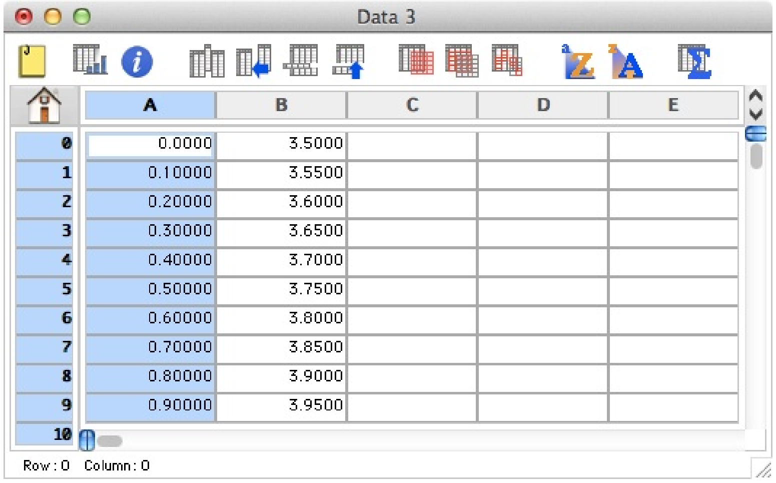
Figure 12: Setting the parameters of the series.



```
c1=0.5*c0 + 3.5
```

○ Deg ● Rad

Figure 13: Defining a formula for your theory curve. In this case it is $f(x) = Ax + B$, where $x$ is the series we defined in the first column, labelled c0, the slope $A$ is set at 0.5, and the y-intercept $B$ is 3.5. (Yes, you can define variables. No, I don't want to get into that here.) The output is saved in the second column, which Kaleidagraph recognizes as c1. Click Run when you are done.

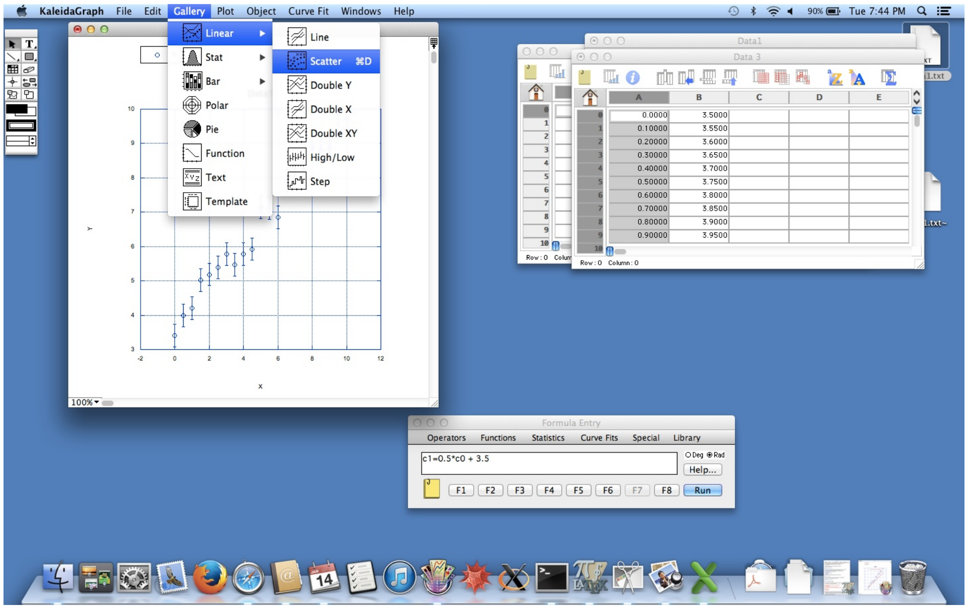


Data 3

|  | A | B | C | D | E |
|---|---|---|---|---|---|
| 0 | 0.0000 | 3.5000 |  |  |  |
| 1 | 0.10000 | 3.5500 |  |  |  |
| 2 | 0.20000 | 3.6000 |  |  |  |
| 3 | 0.30000 | 3.6500 |  |  |  |
| 4 | 0.40000 | 3.7000 |  |  |  |
| 5 | 0.50000 | 3.7500 |  |  |  |
| 6 | 0.60000 | 3.8000 |  |  |  |
| 7 | 0.70000 | 3.8500 |  |  |  |
| 8 | 0.80000 | 3.9000 |  |  |  |
| 9 | 0.90000 | 3.9500 |  |  |  |
| 10 |  |  |  |  |  |

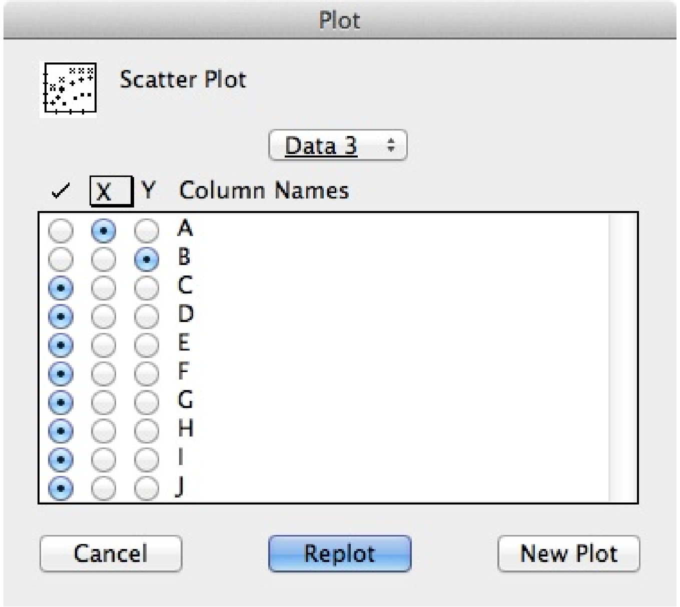
Row : 0 Column : 0

Figure 14: Running your function populates c1 with the result of applying your function to the series in c0.



c1=0.5*c0 + 3.5

Figure 15: To add your newly-generated theory curve to your plot, first click on your plot, then go back to the Gallery menu, and choose Scatter again.



| ✓ | X | Y | Column Names |
|---|---|---|---|
| ○ | ● | ○ | A |
| ○ | ○ | ● | B |
| ● | ○ | ○ | C |
| ● | ○ | ○ | D |
| ● | ○ | ○ | E |
| ● | ○ | ○ | F |
| ● | ○ | ○ | G |
| ● | ○ | ○ | H |
| ● | ○ | ○ | I |
| ● | ○ | ○ | J |

Figure 16: You can use data from many different Data windows for multiple curves on the same plot. Note the pull-down menu that says Data 3. You can select columns from as many data windows as you have to add to your plot, and define whichever you want to be the X and Y variables.

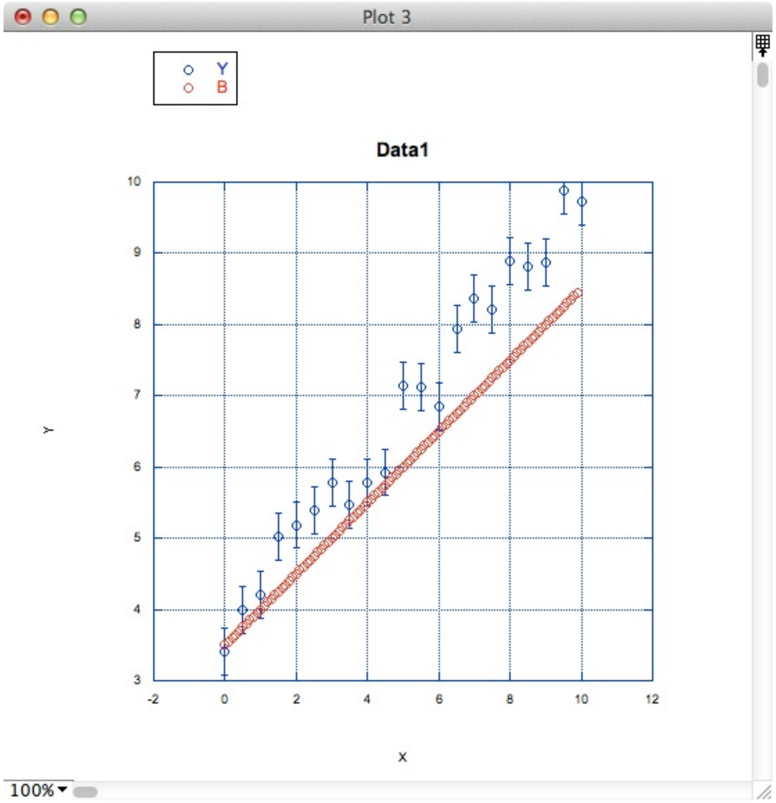


Figure 17: And this is what you should get. (In some cases, adding a theory curve will remove your error bars. If this happens to you, just regenerate them.) At this point the y-intercept looks pretty good, but obviously our initial guess for the slope was too small. Tweak the parameters, and regenerate the plot until the fit looks right.

# 5 Residuals

One way to judge how well a theory curve fits your data is to plot your *residuals*, the difference between the data and the theory for each point, along with the error bars.

$$R_i = y_i - f(x_i)$$

The error bars on $R_i$ are the same as those on $y_i$. If you've gotten that far in *Taylor*, this is because the error on $f(x_i)$ is essentially negligible.

**LABORATORY EXERCISE 4:** Calculate and plot your residuals, with error bars.

## 5.1 Kaleidagraph

You know enough now to be able to do this without me walking you through it. (There are actually several different ways of doing it, all equally valid.) You should get something that looks like this.

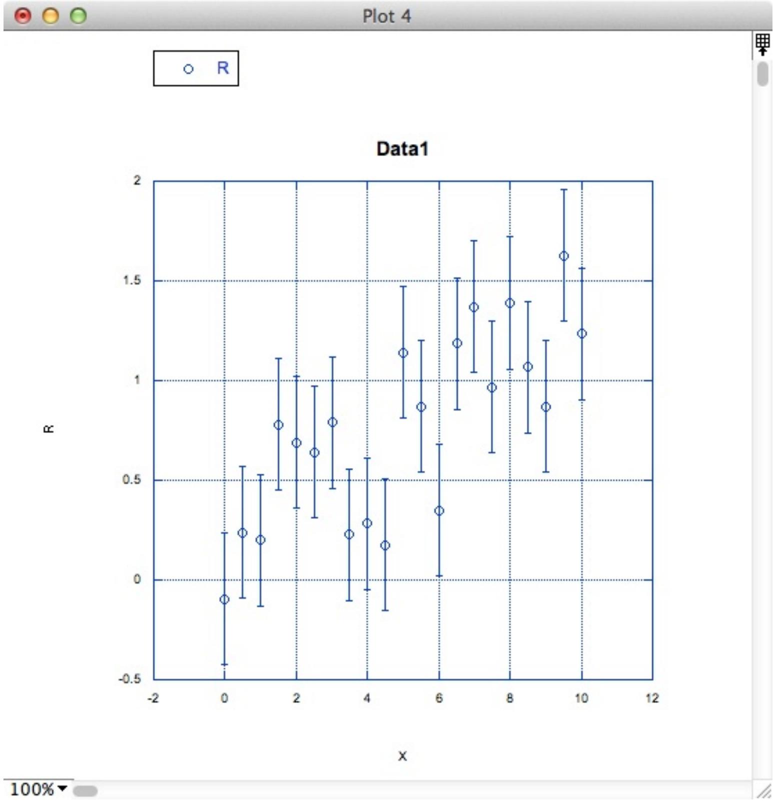


Figure 18: Residuals with error bars.

# 6 $\chi^2$ test

Up to now you have been just looking at your plot and judging by eye whether or not the fit looked right. Now we need a quantitative measure of how good the agreement between your theory curve and data points is. The standard way to do that is with a quantity called the *reduced Chi squared*, or $\chi^2_r$. You will see more of this in your *Taylor* homework, but for now we will simply define it as the mean-squared deviation of the data points from the theory curve, measured in units of the error bars.

$$\chi^2_r \approx \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - f(x_i))^2}{\sigma^2} \tag{1}$$

The distance between a data point and a theory curve that fits its set should be on the order of the uncertainty $\sigma$, so the reduced chi squared should be around one for a curve that fits the data. If its value is much greater than one, that usually means that the theory is way off from the data. I say "usually" because it is also possible to get a large $\chi^2$ by underestimating your error bars, but that seldom happens to careful scientists. More often than not, if you make a mistake in estimating your error bars, you do so by assuming they are bigger than they really are. This makes your $\chi^2$ much less than one.

$\chi^2$ is what computers use to do automated fitting of theory curves to data sets. Every theory curve has a certain number of *parameters* that can be adjusted to change the curve. In our linear case there were two, the slope $A$ and y-intercept $B$. When you ask a computer to do a curve fit for you, it will just adjust the parameters to minimize the $\chi^2$. (It can do this even if you don't supply error bars, but the value of the final, minimized $\chi^2$ won't necessarily be close to one.)

**LABORATORY EXERCISE 5:** Calculate the reduced chi squared for your best-looking theory curve. Does it indicated a good fit? Does it indicate appropriately-chosen error bars?

## 6.1 How to do this in Kaleidagraph

There are, of course, many ways to do this. The most straightforward way is,

1. Insert two new columns into your data using the Insert Column command from the Data menu.

2. Run your theory function on the X values in your data.

3. Define a formula to calculate the individual terms in the sum in Equation 1, e.g. c3=(c1-c2)^2/(0.33^2), where c1 is the column containing the Y values of your data, and c2 is the column containing theory values you just generated.

4. Sum the terms and divide by the total number of terms (20) by defining a function like this and running it,

   cell(23,3)=csum(c3)/20

This will write the result to the cell in the 23rd row and fourth column (well outside of where your data lives), and this will be your chi squared.

# 7 Least-squares fitting

In the case where your theory curve is a straight line, it is simple to calculate the two parameters $A$ and $B$ that minimize $\chi_r^2$ and therefore optimize the fit. Our expression for the reduced chi squared (Equation 1) becomes

$$\chi_r^2 \approx \frac{1}{N}\sum_{i=1}^{N}\frac{(y_i - Ax_i - B)^2}{\sigma^2}$$

To find the parameters $A$ and $B$ that minimize this, we just take the derivatives with respect to $A$ and $B$, and set them to zero.

$$\frac{\partial\chi_r^2}{\partial A} = 0 \text{ and } \frac{\partial\chi_r^2}{\partial B} = 0$$

This gives us

$$\sum_{i=1}^{N}(y_i - Ax_i - B)x_i = 0$$

and

$$\sum_{i=1}^{N}(y_i - Ax_i - B) = 0$$

These two equations are easy to solve for $A$ and $B$, giving us straightforward formulae for the optimal coefficients.

$$A = \frac{N\sum_{i=1}^{N} x_i y_i - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} y_i\right)}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2} \tag{2}$$

and

$$B = \frac{\left(\sum_{i=1}^{N} x_i^2\right)\left(\sum_{i=1}^{N} y_i\right) - \left(\sum_{i=1}^{N} x_i\right)\left(\sum_{i=1}^{N} x_i y_i\right)}{N\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2} \tag{3}$$

This is known as *least-squares fitting* because it finds the parameters that produce the minimum value for the chi squared. (The careful reader may have noted that I have gone back to using an equals sign $=$ where I had been using an approximately-equals sign before $\approx$. That's because the approximation was in the $1/N$ term in front of the sum, which does not matter for zeroing the derivative.)

**LABORATORY EXERCISE:** Using Equations 2 and 3, calculate the parameters $A$ and $B$ for your data. How close are they to the ones you estimated in Section 4?

### 7.1 More general cases

This procedure can be extended to any theoretical function that is a linear combination of individual functions. The usual example is a polynomial of order $n$.

$$f(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n$$

Minimizing the chi squared between this function and a set of data points would yield a set of $n+1$ linear equations, which can be treated by matrix methods to find the parameters $a_0$, $a_1$, etc. Note that the terms do not have to be polynomials, only a linear combination of things, so the procedure would work just as well for a function of the form

$$f(x) = A\sin(2\pi x) + Be^x + C\ln(x)$$

where the parameters to be adjusted for the fit are $A$, $B$, and $C$. The procedure only breaks down if one of your fit parameters falls *inside* the argument of one of your nonlinear functions. For cases like that more sophisticated methods exist, but you already learned the most basic one in Section 4: adjust the parameters by hand until the fit appears to be a good one. That procedure is often referred to as *chi-by-eye*, and good experimentalists use it far more than they will usually admit.

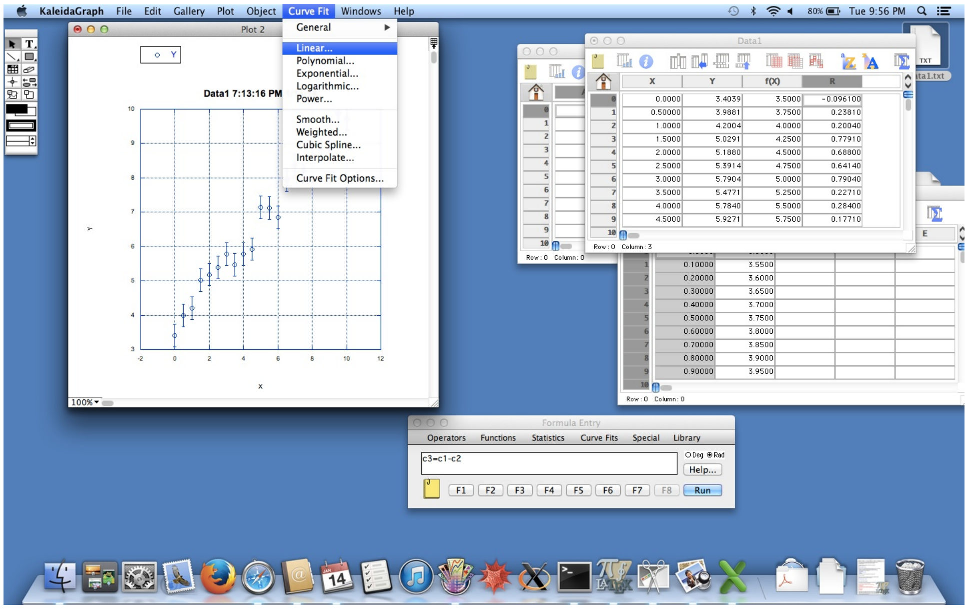
## 8 Automated fitting

Remember how I told you I didn't want you using the curve fitting routines built into the software just yet? Well, now you know enough to understand how they work. Try it out, and see what happens.

**LABORATORY EXERCISE 6:** Using the built-in curve-fitting routine in your software, fit a line to your data, and see how the slope and intercept compare with your own estimates.

## 8.1 How to do this in Kaleidagraph

Curve fitting in Kaleidagraph is very straightforward.



Data1

|  | X | Y | f(X) | R |
|---|---|---|---|---|
| 0 | 0.0000 | 3.4039 | 3.5000 | -0.096100 |
| 1 | 0.50000 | 3.9881 | 3.7500 | 0.23810 |
| 2 | 1.0000 | 4.2004 | 4.0000 | 0.20040 |
| 3 | 1.5000 | 5.0291 | 4.2500 | 0.77910 |
| 4 | 2.0000 | 5.1880 | 4.5000 | 0.68800 |
| 5 | 2.5000 | 5.3914 | 4.7500 | 0.64140 |
| 6 | 3.0000 | 5.7904 | 5.0000 | 0.79040 |
| 7 | 3.5000 | 5.4771 | 5.2500 | 0.22710 |
| 8 | 4.0000 | 5.7840 | 5.5000 | 0.28400 |
| 9 | 4.5000 | 5.9271 | 5.7500 | 0.17710 |
| 10 |  |  |  |  |

Row : 0 Column : 3

|  |  |  |  |  | E |
|---|---|---|---|---|---|
| 1 | 0.10000 | 3.5500 |  |  |  |
| 2 | 0.20000 | 3.6000 |  |  |  |
| 3 | 0.30000 | 3.6500 |  |  |  |
| 4 | 0.40000 | 3.7000 |  |  |  |
| 5 | 0.50000 | 3.7500 |  |  |  |
| 6 | 0.60000 | 3.8000 |  |  |  |
| 7 | 0.70000 | 3.8500 |  |  |  |
| 8 | 0.80000 | 3.9000 |  |  |  |
| 9 | 0.90000 | 3.9500 |  |  |  |
| 10 |  |  |  |  |  |

Row : 0 Column : 0

```
c3=c1-c2
```

Figure 19: To add a curve fit to your plot, click on the plot you want to fit, then pull down the Curve Fit menu, and choose the type of fit you want. In this case, it is Linear.

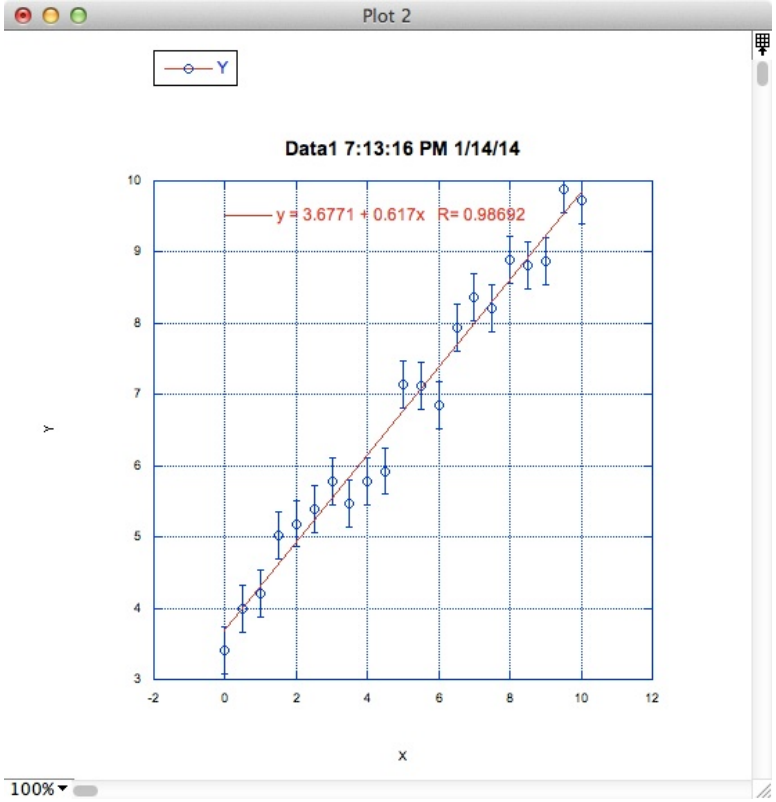


Figure 20: The result is what you expect, and the fitted formula is printed on the plot. The $R$ is a correlation coefficient, which, like $\chi^2$, serves as a measure of the goodness of the fit.

# References

[1] Edward R. Tufte, *The Visual Display of Quantitative Information*, Graphics Pr; 2nd edition (2001).

# A How to plot and fit data in Microsoft Excel

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 appendix.

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

## A.1 How to open a data file in Microsoft Excel

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 23, 24, and 25).

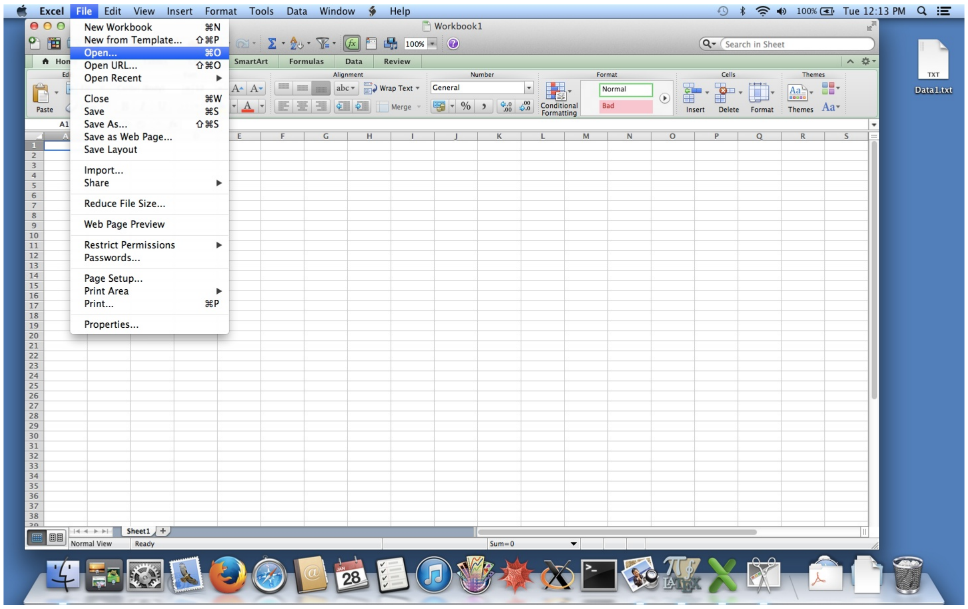
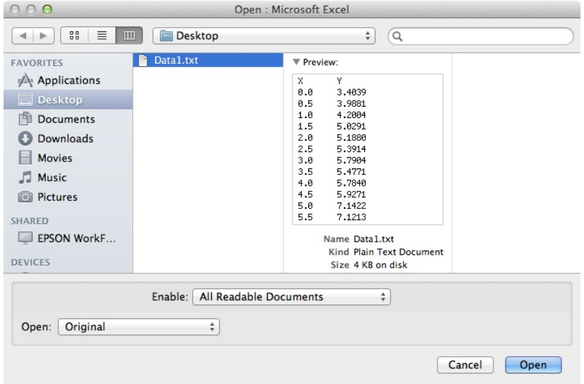


Figure 21: Opening a data file from Microsoft Excel.



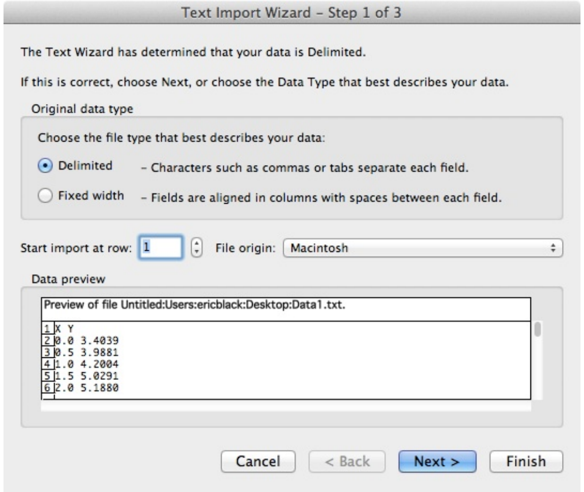
```
X       Y
0.0     3.4039
0.5     3.9881
1.0     4.2004
1.5     5.0291
2.0     5.1880
2.5     5.3914
3.0     5.7904
3.5     5.4771
4.0     5.7840
4.5     5.9271
5.0     7.1422
5.5     7.1213
```

Name Data1.txt
Kind Plain Text Document
Size 4 KB on disk

Enable: All Readable Documents

Open: Original

Figure 22: 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.



The Text Wizard has determined that your data is Delimited.

If this is correct, choose Next, or choose the Data Type that best describes your data.

Original data type

Choose the file type that best describes your data:

● Delimited – Characters such as commas or tabs separate each field.

○ Fixed width – Fields are aligned in columns with spaces between each field.

Start import at row: 1 File origin: Macintosh

Data preview

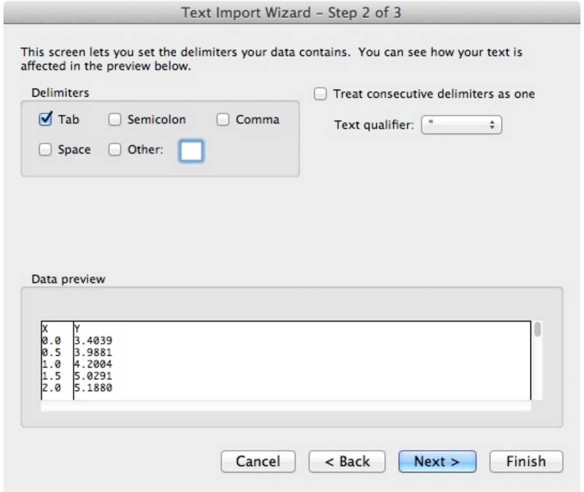
Preview of file Untitled:Users:ericblack:Desktop:Data1.txt.

```
1 X Y
2 0.0 3.4039
3 0.5 3.9881
4 1.0 4.2004
5 1.5 5.0291
6 2.0 5.1880
```

Figure 23: The first of three dialog boxes.

## A.2 How to make a plot

Plots are called "Charts" in Excel. To make one, first select the data you want to plot, as shown in Figure 26. 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



Text Import Wizard – Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

☑ Tab □ Semicolon □ Comma

□ Space □ Other:

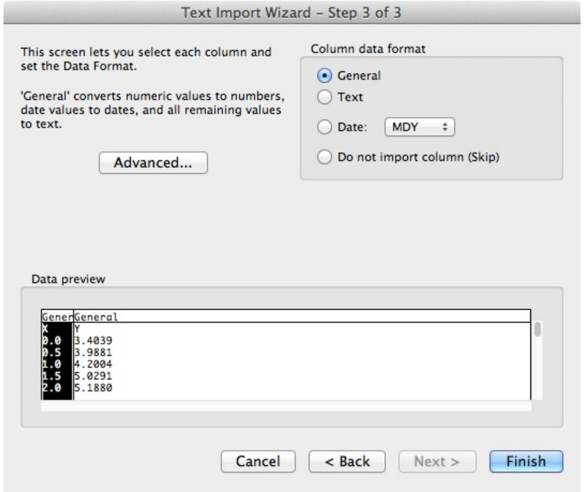
□ Treat consecutive delimiters as one

Text qualifier: "

Data preview

```
X    Y
0.0  3.4039
0.5  3.9881
1.0  4.2004
1.5  5.0291
2.0  5.1880
```

Figure 24: The second of three dialog boxes.



Text Import Wizard – Step 3 of 3

This screen lets you select each column and set the Data Format.

'General' converts numeric values to numbers, date values to dates, and all remaining values to text.

Column data format

● General

○ Text

○ Date: MDY

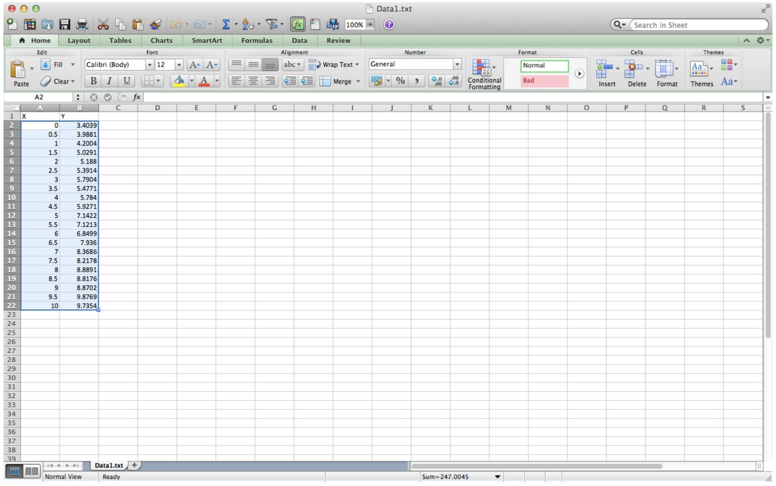
○ Do not import column (Skip)

Data preview

```
GenerGeneral
X    Y
0.0  3.4039
0.5  3.9881
1.0  4.2004
1.5  5.0291
2.0  5.1880
```

Figure 25: The third of three dialog boxes.

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.



| X | Y |
|---|---|
| 0 | 3.4039 |
| 0.5 | 3.9881 |
| 1 | 4.2004 |
| 1.5 | 5.0291 |
| 2 | 5.188 |
| 2.5 | 5.3914 |
| 3 | 5.7904 |
| 3.5 | 5.4771 |
| 4 | 5.784 |
| 4.5 | 5.9271 |
| 5 | 7.1422 |
| 5.5 | 7.1213 |
| 6 | 6.8499 |
| 6.5 | 7.936 |
| 7 | 8.3686 |
| 7.5 | 8.2178 |
| 8 | 8.8891 |
| 8.5 | 8.8176 |
| 9 | 8.8702 |
| 9.5 | 9.8769 |
| 10 | 9.7354 |

Figure 26: What you get.

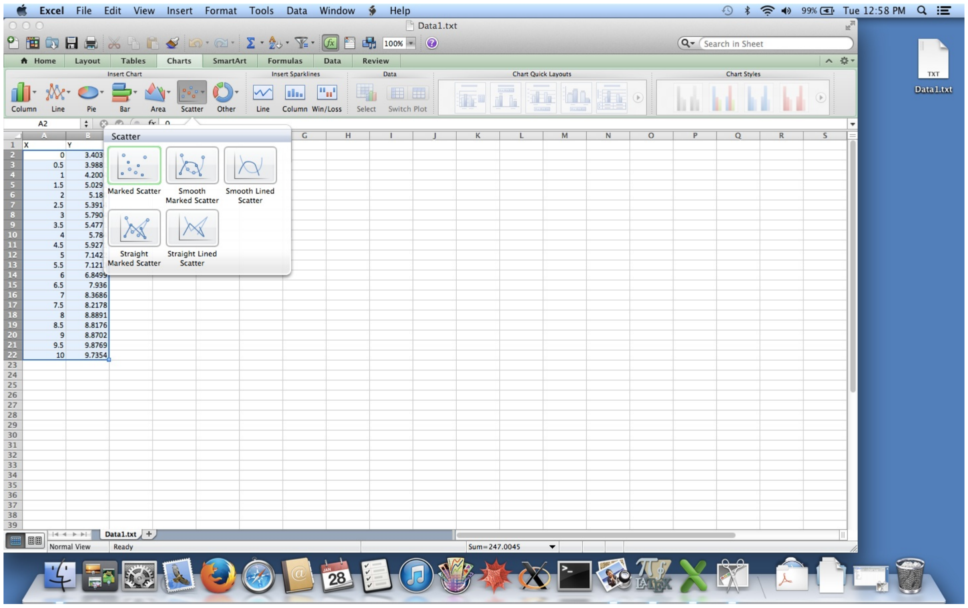
Line plots, as you might expect, connect each data point with a line. This is sometimes useful if you have a lot of data points and aren't planning on marking each with its own little symbol, but it's less than ideal for the kinds of plots you will be making in this lab. Line plots also don't support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

## A.3 Adding error bars

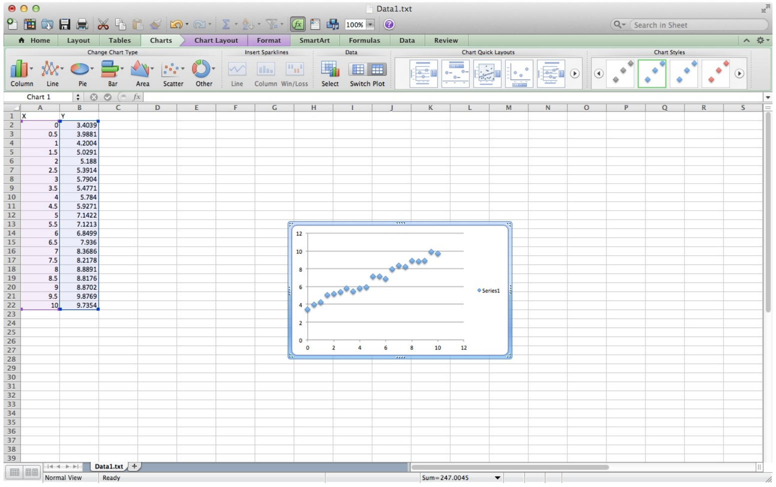
To add error bars, click on "Chart Layout" in the ribbon, then follow the directions from there.

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



| | A | B |
|---|---|---|
| 1 | X | Y |
| 2 | 0 | 3.403 |
| 3 | 0.5 | 3.988 |
| 4 | 1 | 4.200 |
| 5 | 1.5 | 5.029 |
| 6 | 2 | 5.18 |
| 7 | 2.5 | 5.391 |
| 8 | 3 | 5.790 |
| 9 | 3.5 | 5.477 |
| 10 | 4 | 5.78 |
| 11 | 4.5 | 5.927 |
| 12 | 5 | 7.142 |
| 13 | 5.5 | 7.121 |
| 14 | 6 | 6.8499 |
| 15 | 6.5 | 7.936 |
| 16 | 7 | 8.3686 |
| 17 | 7.5 | 8.2178 |
| 18 | 8 | 8.8891 |
| 19 | 8.5 | 8.8176 |
| 20 | 9 | 8.8702 |
| 21 | 9.5 | 9.8769 |
| 22 | 10 | 9.7354 |

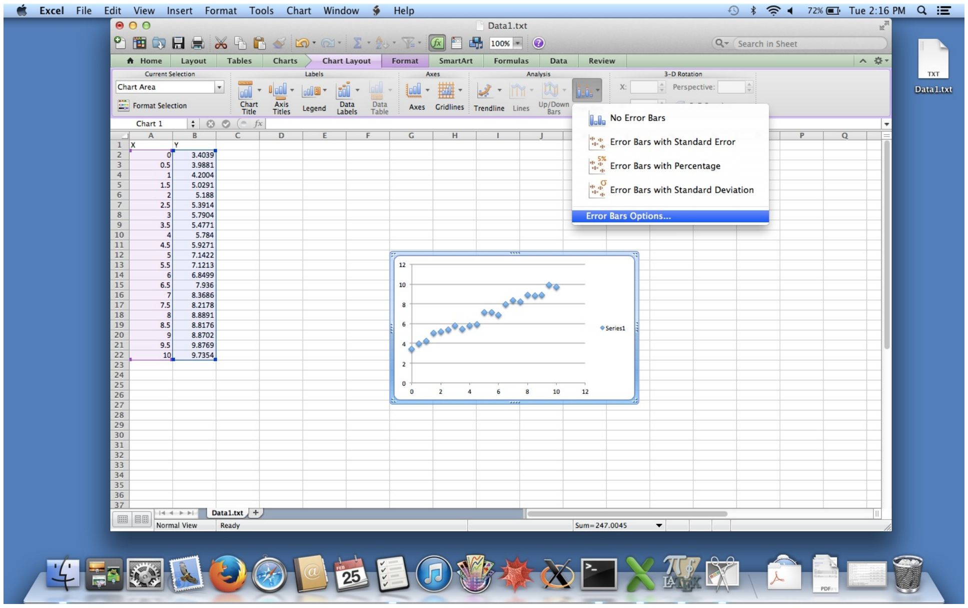
Figure 27: Click on the Charts tab in the Ribbon, then select Scatter as the type of plot.



| | A | B |
|---|---|---|
| 1 | X | Y |
| 2 | 0 | 3.4039 |
| 3 | 0.5 | 3.9881 |
| 4 | 1 | 4.2004 |
| 5 | 1.5 | 5.0291 |
| 6 | 2 | 5.188 |
| 7 | 2.5 | 5.3914 |
| 8 | 3 | 5.7904 |
| 9 | 3.5 | 5.4771 |
| 10 | 4 | 5.784 |
| 11 | 4.5 | 5.9271 |
| 12 | 5 | 7.1422 |
| 13 | 5.5 | 7.1213 |
| 14 | 6 | 6.8499 |
| 15 | 6.5 | 7.936 |
| 16 | 7 | 8.3686 |
| 17 | 7.5 | 8.2178 |
| 18 | 8 | 8.8891 |
| 19 | 8.5 | 8.8176 |
| 20 | 9 | 8.8702 |
| 21 | 9.5 | 9.8769 |
| 22 | 10 | 9.7354 |

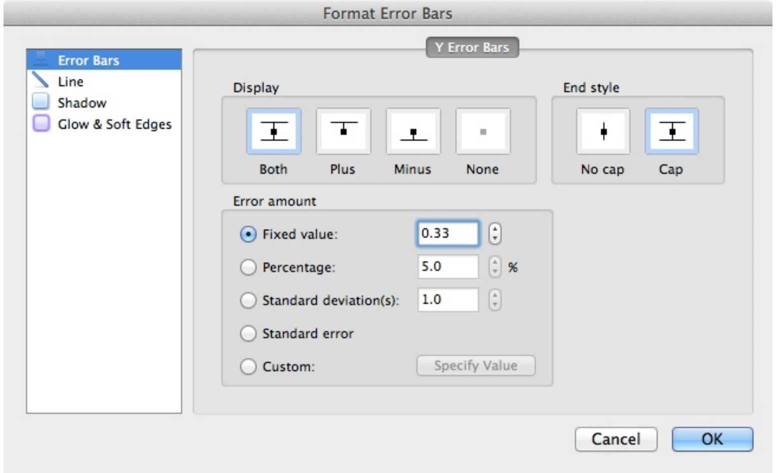
Figure 28: This is what you get.

connect your data points.



| X | Y |
|---|---|
| 0 | 3.4039 |
| 0.5 | 3.9881 |
| 1 | 4.2004 |
| 1.5 | 5.0291 |
| 2 | 5.188 |
| 2.5 | 5.3914 |
| 3 | 5.7904 |
| 3.5 | 5.4771 |
| 4 | 5.784 |
| 4.5 | 5.9271 |
| 5 | 7.1422 |
| 5.5 | 7.1213 |
| 6 | 6.8499 |
| 6.5 | 7.936 |
| 7 | 8.3686 |
| 7.5 | 8.2178 |
| 8 | 8.8891 |
| 8.5 | 8.8176 |
| 9 | 8.8702 |
| 9.5 | 9.8769 |
| 10 | 9.7354 |

Figure 29: 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.



● Fixed value: 0.33

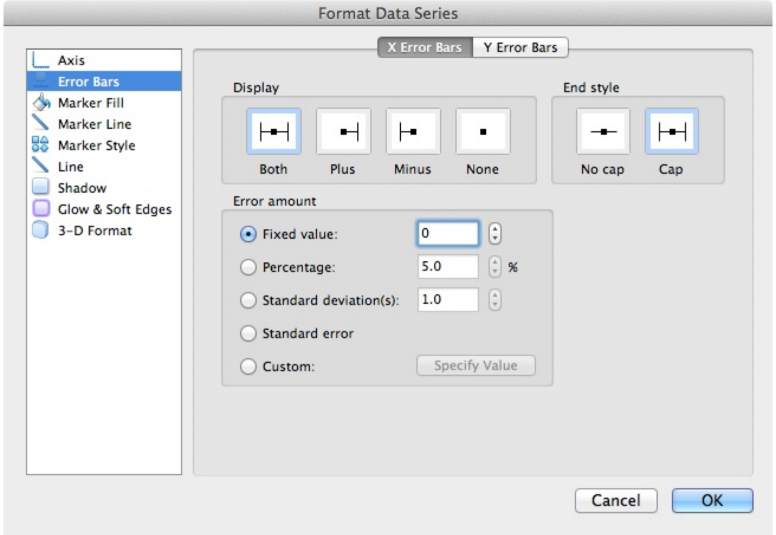
○ Percentage: 5.0 %

○ Standard deviation(s): 1.0

○ Standard error

○ Custom:

Figure 30: Even though Excel puts error bars on the independent variable (X), you don't get to specify them yet.



● Fixed value: 0

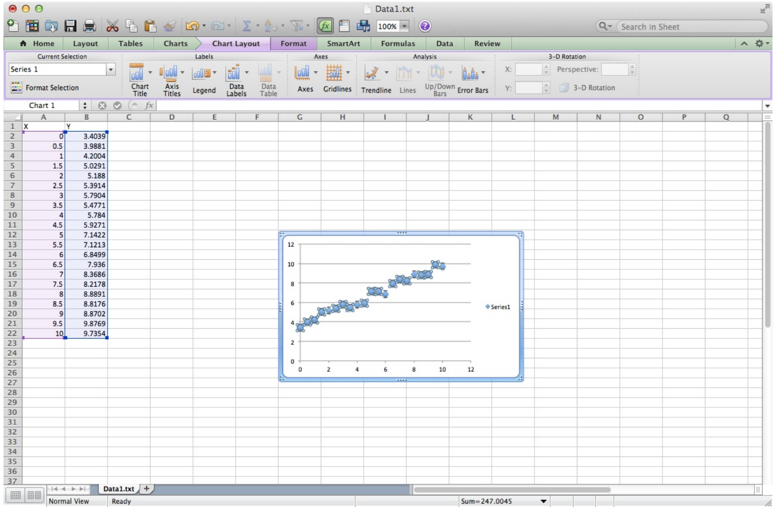
○ Percentage: 5.0 %

○ Standard deviation(s): 1.0

○ Standard error

○ Custom:

Figure 31: 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.



| X | Y |
|---|---|
| 0 | 3.4039 |
| 0.5 | 3.9881 |
| 1 | 4.2004 |
| 1.5 | 5.0291 |
| 2 | 5.188 |
| 2.5 | 5.3914 |
| 3 | 5.7904 |
| 3.5 | 5.4771 |
| 4 | 5.784 |
| 4.5 | 5.9271 |
| 5 | 7.1422 |
| 5.5 | 7.1213 |
| 6 | 6.8499 |
| 6.5 | 7.936 |
| 7 | 8.3686 |
| 7.5 | 8.2178 |
| 8 | 8.8891 |
| 8.5 | 8.8176 |
| 9 | 8.8702 |
| 9.5 | 9.8769 |
| 10 | 9.7354 |

Figure 32: This is what you get.