# Graphing in Excel

Introduction

This handout provides information about how to use the graphing capabilities in Excel to make nice graphs in a fairly painless fashion. In the first lab, we will go through an example step by step. This handout, however, provides extra details and should be a useful reference throughout the course.

Choosing Graph Type

The two types of graphs we'll be using most often in this course are *scatterplots* and *line graphs*. Scatterplots are very similar to line graphs, except that scatterplots do not have lines connecting the plot points like the line graphs do.

The first step in creating a graph is to select the data that you wish to graph. Do this by dragging the cursor from the upper left cell of the data to the lower right. Next, you need to select Chart from the Insert menu. This will bring up the following dialog box (Chart Wizard).   Select the Chart type from the list on the left then the Chart sub-type from the right. Next click on the Next button in the bottom of the dialog box.
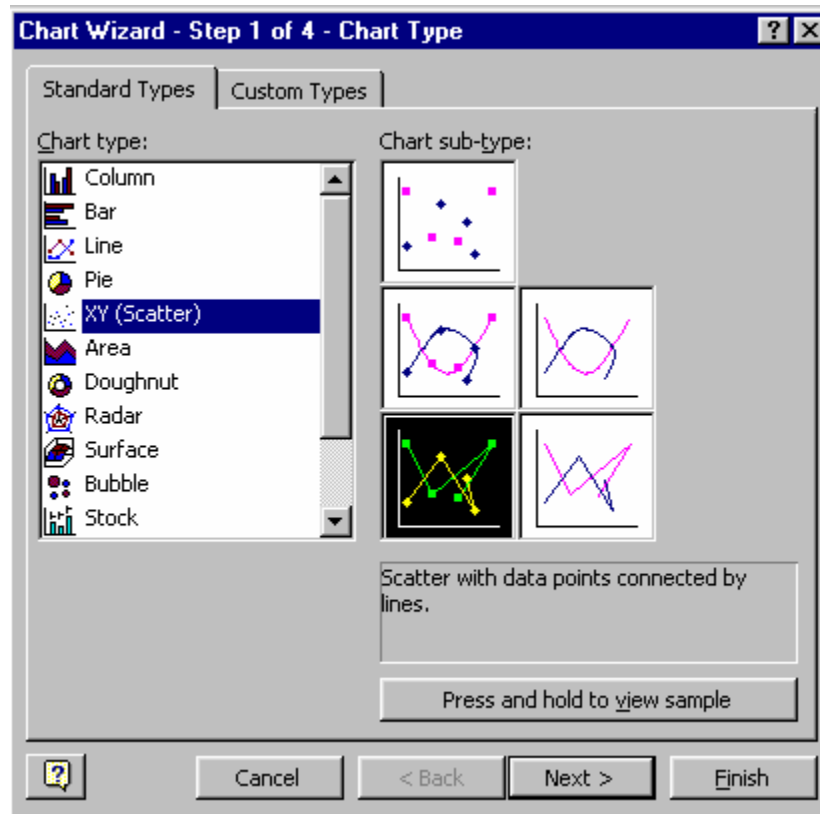
The Next window will show what the graph will look like. Click on the 'Series' tab at the top. Sometimes you may need to specify the category X labels as well as the data labels for the different sets of data. To specify the category X labels, first click on the Series tab at the top of the page. Then click on the button with the red arrow on it to the right of the text box labeled 'Category (X) labels. This will allow you select the cells that you wish to use for the category X labels. In order to specify the data
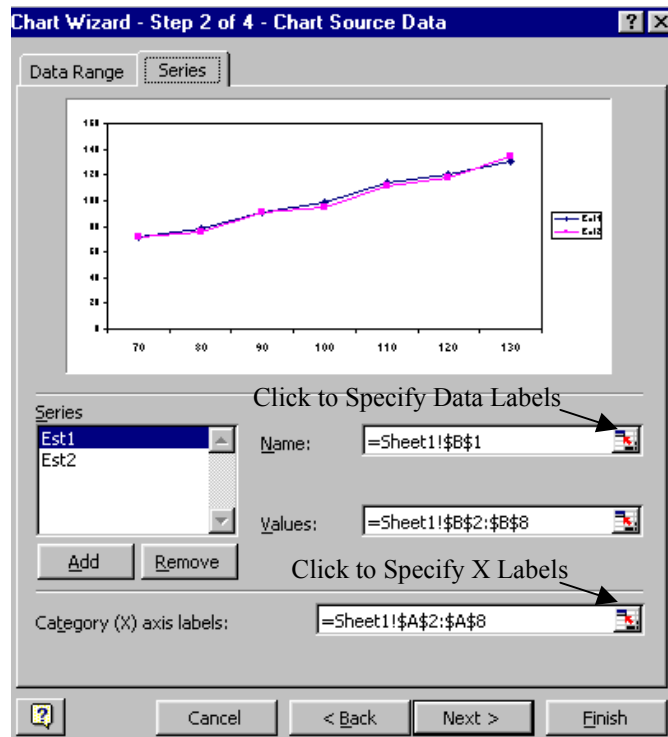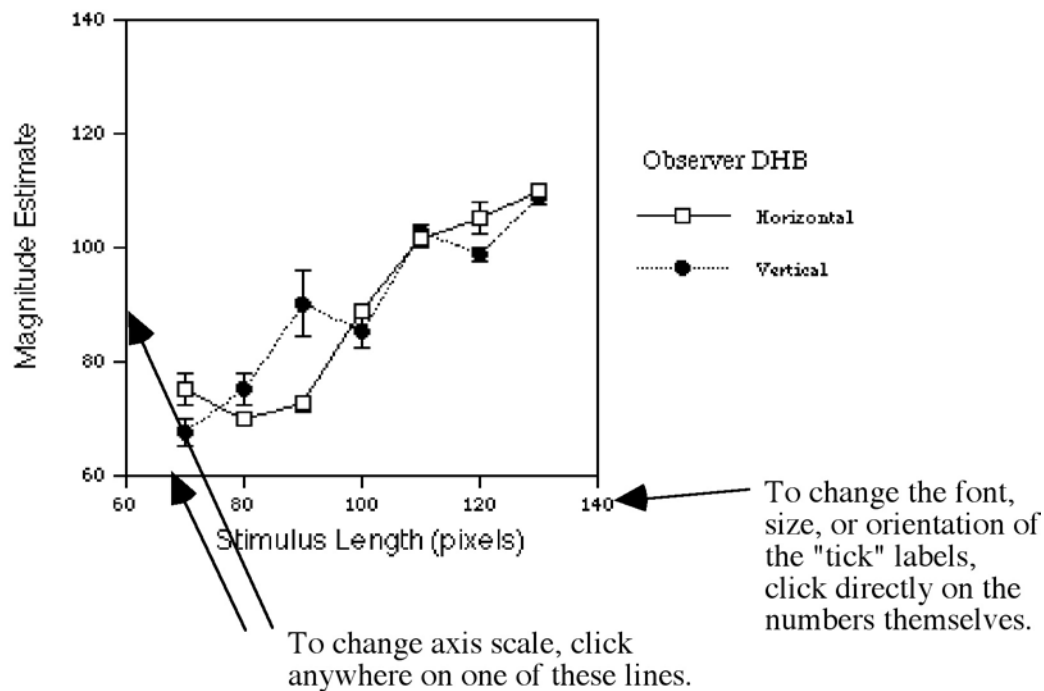

series labels, first click on one of the entries in the 'Series' box, then click on the button with the red arrow to the right of the text box labeled 'Name.' This will allow you to select the cell that will be the label for the data series. When the labels are identified, click the 'Next' button. The next dialog box (see below) allows you to specify the title and axis labels.   By clicking on the tabs across the top of the dialog box, you can specify other attributes of the graph such as the axes style and the legend position. When you are done click the 'Next Button.
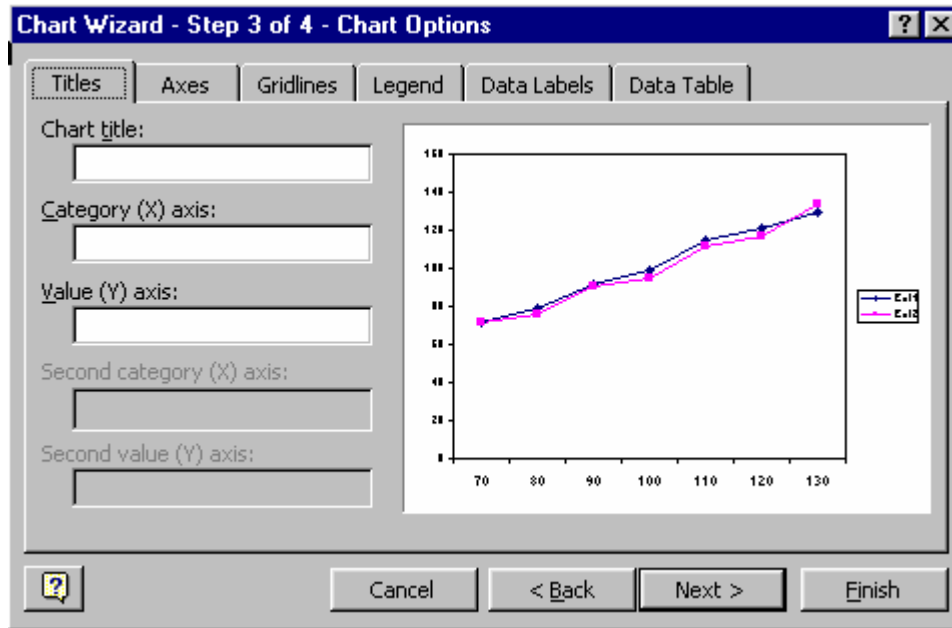
The next window will ask if you want to insert the graph as an object in the data sheet or create new sheet that only contains the graph. Select 'As new sheet.'

Changing the Axis Scale

Excel will automatically take a guess at what should be the highest and lowest values to plot on your graph, but often you'll want to change these. For example, you might want to change the axis scales so that several graphs use the same scale, to make it easier to compare data across graphs. To change the axis scale, double-click directly on the line of the axis itself.

To change the font, size, or orientation of the "tick" labels, click directly on the numbers themselves.

To change axis scale, click anywhere on one of these lines.

You should see a window pop up like the one below that says "Format Axis". Click on the 'Scale' tab at the top of the widow to modify the axis scaling.
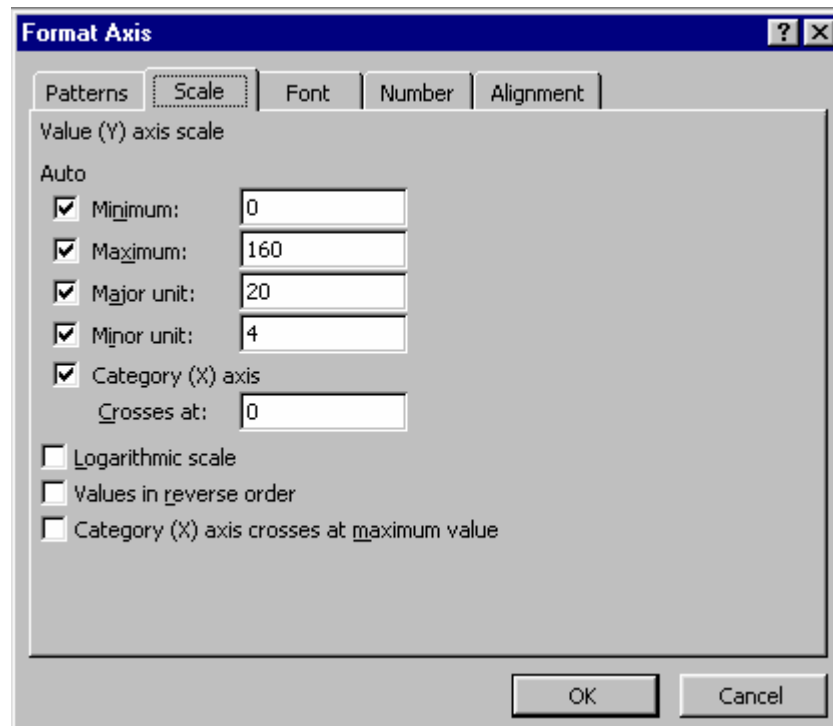
Use the boxes in this window to change the minimum and maximum values that will appear on the graph. Use the box marked "Major unit" to adjust the size of the divisions between the lowest and highest values.

Tick Label Attributes

The numbers directly below axes on the graphs are called the "tick labels". You can change the font, size, and orientation of these numbers, as well as several other attributes. As a matter or graphing style, you should use the same tick label settings for both axes. To change the tick label attributes, double-click on the numbers themselves, and adjust the attributes under the 'Font' tab that appears across the top of the 'Format Axis' window.

Axis Labels

If you need to change the title or axis labels after the graph is made, right click on the graph and select 'Chart Options…', then select the 'Titles' tab.

Repositioning Labels, Legends, Titles, etc.

To move something to a different part of the page in Excel, simply click on the object and hold down the mouse button while you drag the cursor to the place you want the object to be.

Changing Plot Symbols

Excel will automatically use different plot symbols when you are graphing several columns at once, but you will probably want to change these first guesses to suit your own tastes. To change what symbol is used, the size of the symbol, whether the symbol is filled or open, etc., double-click on the plot symbol in the graph, and then adjust the look of your symbols using the "Format Legend" window.

Creating dotted or dashed lines to connect your plot symbols

Double click on the plot symbol in the legend that uses the line you wish to change. You should get the "Format Legend Key " box above. Click and hold the mouse button in the "Style" button in the "Line Style" box, then drag the cursor down to select the type of line you want. Click OK to return to the graph.

Adding error bars

To add error bars to your plot points, first be sure that you have already calculated the standard errors. Then, right click on the line in the plot area that corresponds to the data you want to add error bars to and select "Format Data Series…". Next, choose the Y Error Bars tab. A window should appear like the one below. This window gives you several options for changing the appearance of the error bars. The first thing to do is to select the cells in the worksheet that correspond the standard error values. To do this click the button to the right of the 'Custom' text box (see below) and select the appropriate cells. To have both error bars above and below the data point, select the same cells for both the '+' box and the '-' box.
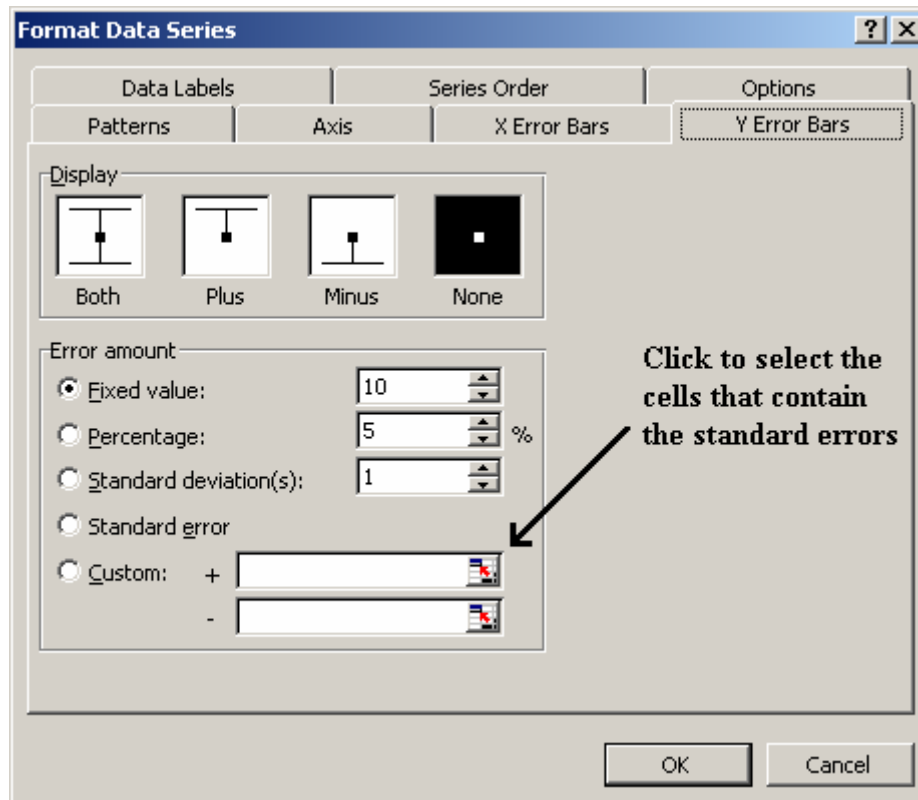
Printing

To print your graph, choose Print... from the File menu. A window will come up—often you can ignore what's in the window and just choose OK to start printing your graph.

Fitting a Line to Data

In many perception experiments, we are interested in how our perception changes as we manipulate the value of some physical variable. Graphing is one way of summarizing how perception relates to the physical variable. Typically, we plot the values of the physical variable along the horizontal axis and our measured values along the vertical axis.

Fitting a smooth curve through the data is another, more mathematical, way of summarizing measurements. On a graph, this amounts to finding a curve that passes as nearly as possible through the middle of the plotted data points. You might imagine that intricate curves would be better able to fit the data than a straight line. Generally this is true, but perhaps you can see a trade-off here: while highly flexible curves often provide the best fit, the equations for such curves can end up being so complex that their utility as a *summary* of the data suffers. In addition, if there is noise in the data, the curve will fit the noise as well as the underlying "truth." Therefore, it is often of interest to find the straight line that provides the best fit to the data. In this handout, we explain how this may be done using Excel, and how to plot the line using Excel. The graphs below show a hypothetical data set and the corresponding line that provides the best fit.
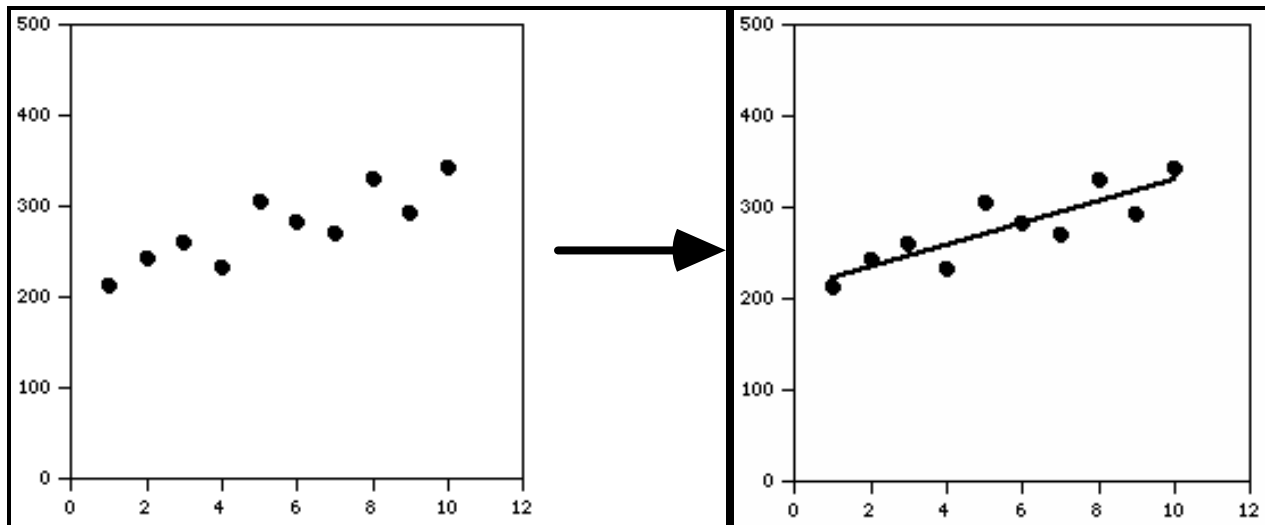
Finding and plotting the best fitting line

To make graphs of data like the one on the left above, we typically need two columns of data: one for the values on the horizontal ("x") axis and the other for the values on the vertical ("y") axis. Because we want a line that describes how the y-values change when we vary the x-values, we will need both those columns in estimating the best-fitting line through the data. Accordingly, the equation for a straight line contains both a "y" variable (which in our case represents the column of y-values) and an "x" variable (for the column of x-values):

$$y = m*x + b.$$

In our terminology, **y** is the measured value for some dependent variable (e.g. perceived length, threshold intensity, reaction time, etc.) for a particular independent variable **x** (e.g. line length, masker intensity, number of items, etc.). The constant **m** is a number that determines the slope of the line that relates the *dependent variables* to the *independent variables.* The constant **b** is another number and determines where the line crosses the y-axis. What we want to do is use Excel to find the values for **m** and **b** that make the best fitting line. We often call **m** and **b** the parameters of the straight line that best fits the data in the "y" column.

Finding and plotting the best fitting parameters for your data is really simple. First, right click the mouse on one of the data points in your graph and choose 'Add Trend Line …'. This will bring up a window (shown below).
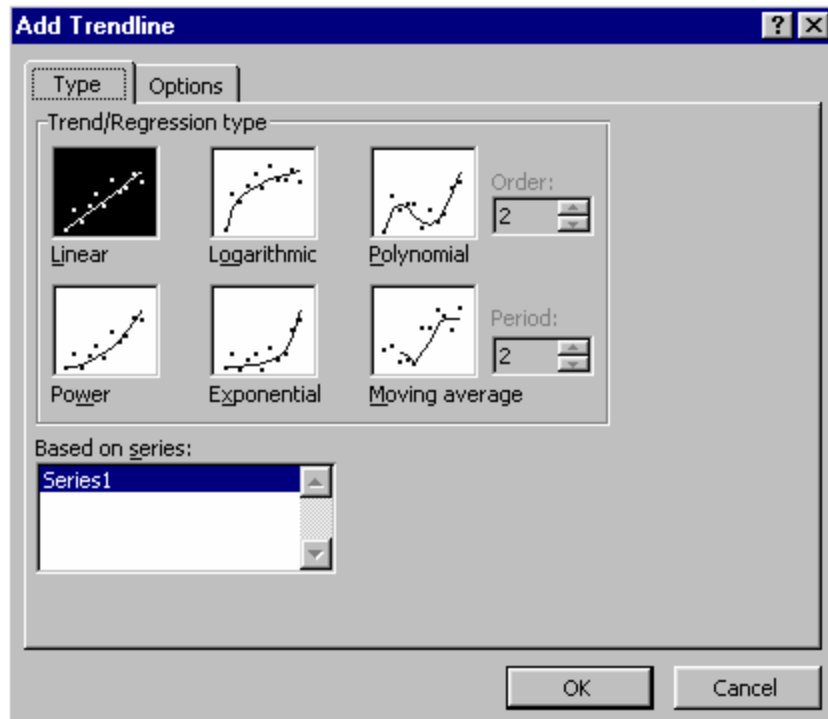
First, you need to select the type of line to fit the data. So far, we have talked about using a linear line to fit the data, so make sure that the 'Linear' box is darkened. Next click the 'Options' tab at the top of the window (see below).

In order to find the parameter values of the best fitting line, make sure that the 'Display equation on chart' button is checked. If you want to know how well the line describes the data, check the 'Display R-squared value on chart' button. The R-squared value is a number from 0 to 1 that indicates how well you can predict the y values from the line. An R-squared of '1' indicates perfect prediction, whereas an R-squared of '0', indicates the line has no predictive value. Finally, click the 'OK' button to plot the line. If you want to change the way the line looks, right click the mouse on the line and choose, 'Format Trend Line…'; then, in the window that appears, choose the 'Patterns' tab.