Group members_____

This document assumes you are using the software available in campus computer labs, Microsoft Excel 2010 for PC. If you are using your own computer or the Mac lab, the version of your software may be slightly different. See the instructor if you have questions about your specific version. Write down your answers for all questions in each section before moving on to the next section.

Section 1 – The Parts of Scientific Graphs

Scientific graphs should stand by themselves. Your audience should be able to see all the important information about the data from the graph. Graphs must include:

1. Meaningful axes, with well-chosen tick marks
2. A caption or title
3. A legend only if there is more than one data set

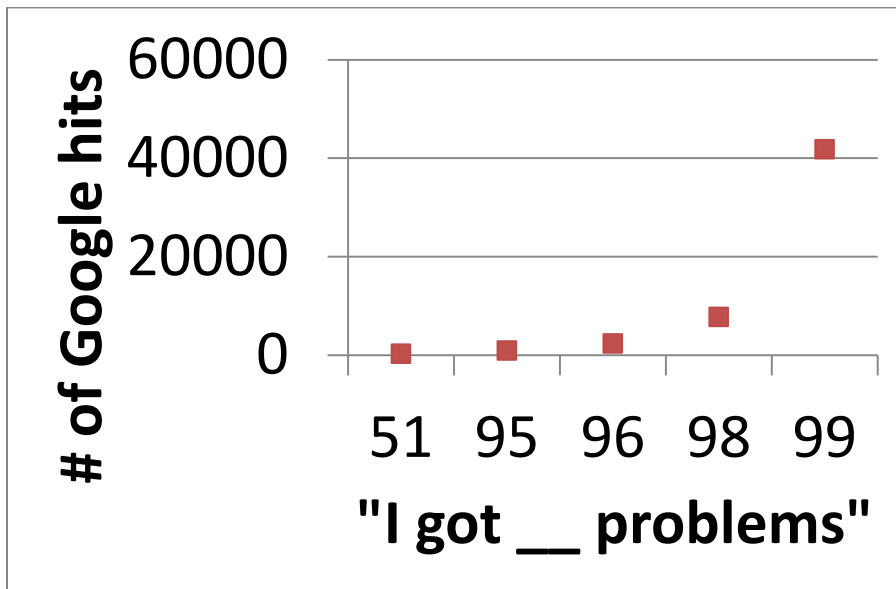We'll go over each of these in detail in later sections. First, inspect this graph:

Figure 1: Based on xkcd 715

What do you think about the presentation? Does it have all the parts it needs? Can you see the data points? Is the font size appropriate? What is the relationship the graph is trying to communicate? Does the number of Google hits increase exponentially for the number of problems?

Section 2: Graph types

The data points in Figure 1 seem to be increasing in an exponential fashion. BUT the axis is squished to give this appearance. The space between 51 and 95 is the same as the space between 95 and 96. This is a "line plot" in Excel, meaning that Excel thinks the x-axis data is names, not numbers, so it does not space the data points correctly. It is confusing because both "Line" charts and "scatter" charts can have data lines drawn on them. Do not ever draw data lines in this course; always select the sub-type with data markers, not lines.
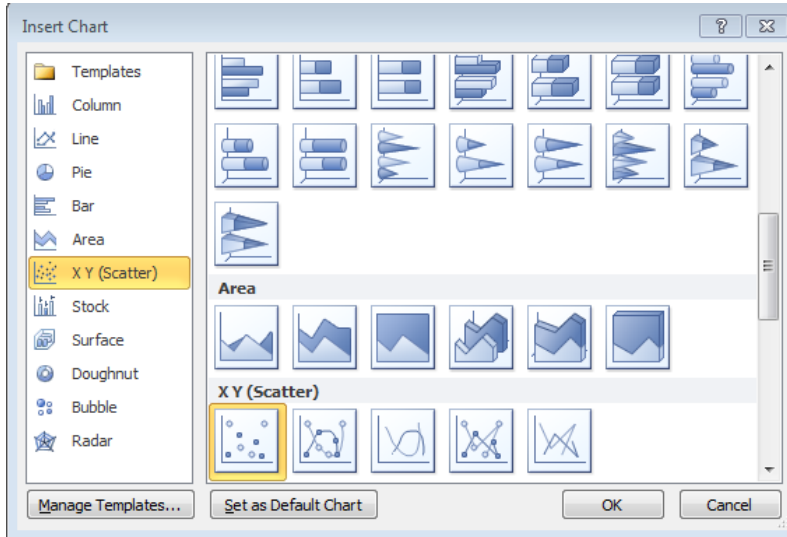
When you insert a chart in Excel, use "Line" if your data is just y numerical values, but use "Scatter" if your data has both x and y values. Column (or bar) graphs and pie (or doughnut) graphs are also frequently used in science, but for this course we'll almost exclusively be drawing Scatter graphs.

The following graphs are examples of poor graphs. Identify what type the following charts are and what they should be. Then list any problems that make these charts ineffective.
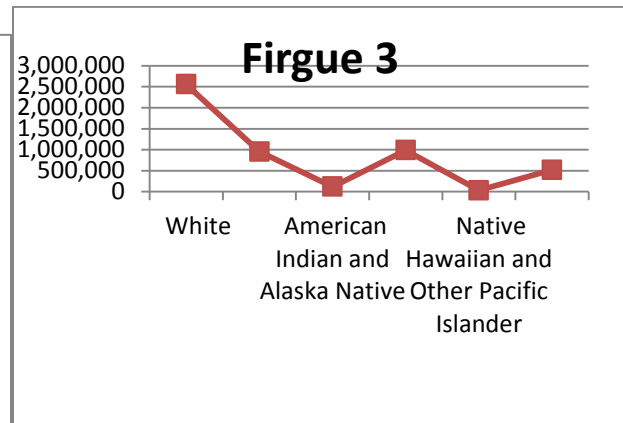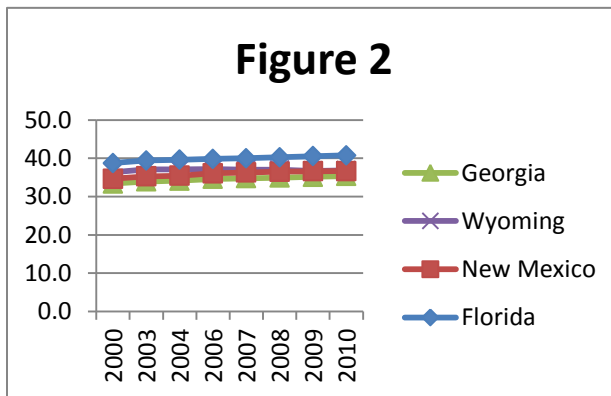
Section 3: Drawing Graphs

Next you are going to redraw the Figure 1 graph correctly, to see whether the data is actually exponential.

Enter the following data into Excel:

| # problems: | 51, | 95, | 96, | 98, | 99 |
| # of Google hits: | 315, | 907, | 2350, | 7780, | 41,800 |

Hopefully you put the data in columns. It's not wrong to put data in rows, but some things are easier in Excel in columns. If you entered the data in rows, please copy and paste as "transpose." Get the instructor to help if needed. **Put your headings at the top of each column.** Now select the data (hold down shift, select cells, and then **insert a chart** that is "scatter type." You should have 0 -120 on the x-axis and 0-45,000 on the y-axis. If not, select a different range or rearrange your columns and try again.

Section 4: Axes

When drawing column (or bar) charts, there is a rule that the axis should start at zero. This rule does NOT apply to scatter or line plots. Your axes should always be selected so that there is not a ton of blank space taking up your graph. Right click on the axis and chose "Format axis" at the bottom of the list. Set the minimum to the fixed value of 40 and the maximum to 100. Always set you axis limits on scatter plots so that your data is visible and there is not a lot of empty space. Also on this screen, you can adjust the spacing between tick marks, control the number format etc. For large numbers, it's easier to view what's going on if you use commas. For very large numbers, you want to use scientific notation. To change the format of the y-axis, select the axis, right click and select format. Under the "Number " menu, Number category, select the desired format. In this case, use 1000 separator and 0 decimal places.
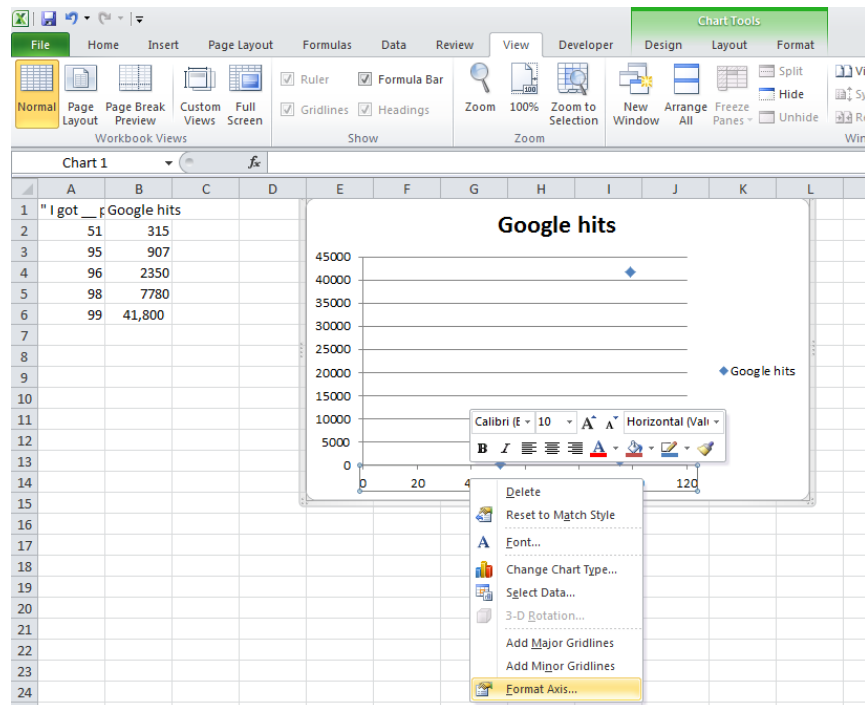
Add axis titles to your graph by going to "Chart Tools" "Layout" and axis titles. I like selecting my heading as my axis title, so that I can adjust the axis titles easily. After selecting the orientation for the title, go to the formula bar (the white line next to the fx symbol) and hit "=", then the cell with your heading in it.

How do we know which way to orient the axes? The x-axis should be your independent variable, the one you control. I typed in "I got __ problems" into Google, filling in the blank with the various values, so I controlled that phrase. The y-axis is your dependent variable, the one you observe. Google spit back the values on the y-axis and I recorded them. Sometimes, you'll need to think about the scientific law or idea that the graph is purporting to show in order to determine what is controlled (x) and what is observed (y). For example, later we will be using a graph to determine values for the variable ΔH° and ΔS° using this equation:

$$\ln(K) = \frac{-\Delta H^\circ}{R}\frac{1}{T} + \frac{\Delta S^\circ}{R}$$
$$y \quad = \quad m \cdot x + \quad b$$

Even without knowing what these symbols mean, you can see that the ln(K) term should be on the y-axis and 1/T data should be on the x-axis to match this linear equation. Decide what should be on the x-axis and which should be on y for the following graphs. Since the slope of a straight line is "rise over run", the units for slope will be your y-axis units divided by your x-axis units. This is another helpful way to check that you are putting data on the correct axes. Remember, ALWAYS include units for variables that have units.

Lab 1, where you dispensed volumes of water and then measured the mass
_____(x)_____(y) slope = density g/mL

Every 5 minutes, you determine the temperature of a beaker of water
_____(x)_____(y)
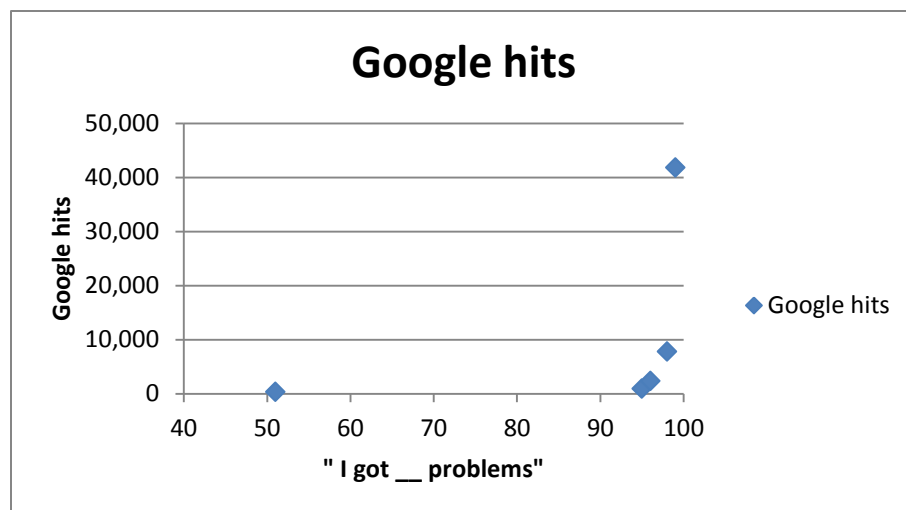
The decline in the number of pirates causes global warming (increase in global temperature).
_____(x) _____ (y)

Section 5: Lines
So if you've followed the steps, your graph should now look like the one below. Try changing it to a Line Graph (like Figure 1) and back to a scatter graph to get familiar with switching graph types.

The next step is to see the relationship between the data points. There are two different types of lines that can be drawn on graphs, data lines and trendlines. Data lines connect the dots between your data points. They can be useful if you have huge data sets, but you should *never draw data lines for this course*. We will always be drawing trendlines. A trendline is a mathematical fit to our data. In lab, you will usually be drawing graphs based on a scientific law. Linear trendlines are the most common. **Insert a linear trendline** on your graph by selecting your data points and then right clicking and insert trendline. Display the

Google hits

equation of your trendline and your R^2 value by selecting the appropriate checkboxes in the Trendline window. Now we'll change the units of the trendline equation to be scientific notation, just so you know how to do that. Select your equation, right click and format trendline label, then set your number type to scientific. You can also use forecast forward and backward if you need the line to go beyond the data you are fitting. Practice doing that now. R^2 is a correlation constant. It gives you a sense of whether your line is a good match to the data points. What makes an R^2 value good depends on what type of data you are dealing with. A value of 1 is a perfect match, and for analytical chemistry 0.99999 is a good R^2 but for stock market analysis $R^2 = 0.8$ is considered a good fit. What do you think about the linear fit for your data?

Google hits

$y = 3.30E+02x - 1.83E+04$
$R^2 = 1.48E-01$

" I got __ problems"
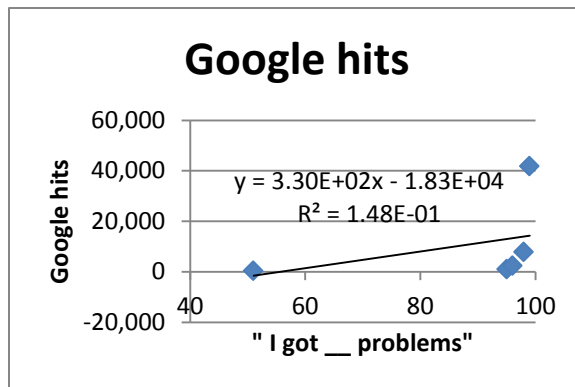
Hopefully you recognized that that line is a terrible fit to the data points. Format the trendline (right click on it) and change to exponential and you should see that the R^2 value increases a tiny bit, but it is still a completely meaningless fit.

Google hits

$y = 9.33E+00e^{6.55E-02x}$
$R^2 = 5.07E-01$

" I got __ problems"
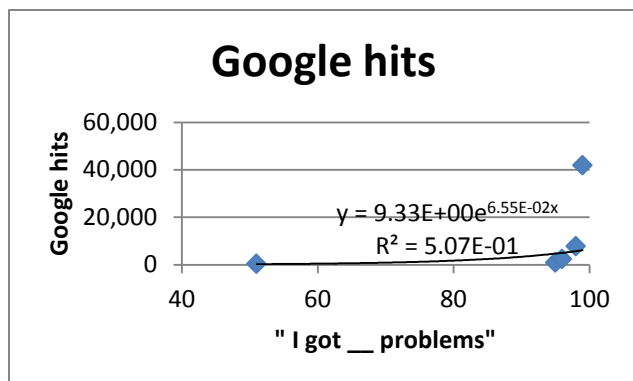
Someone makes the statement "There are exponentially more mentions on the internet when discussing larger numbers of problems." Is this: a) a scientific law b) a scientific theory or c) outside the definition of science? Remember the definitions of laws, theories and science we discussed in class. Does your graph support the statement? Is this a measure of precision or accuracy?
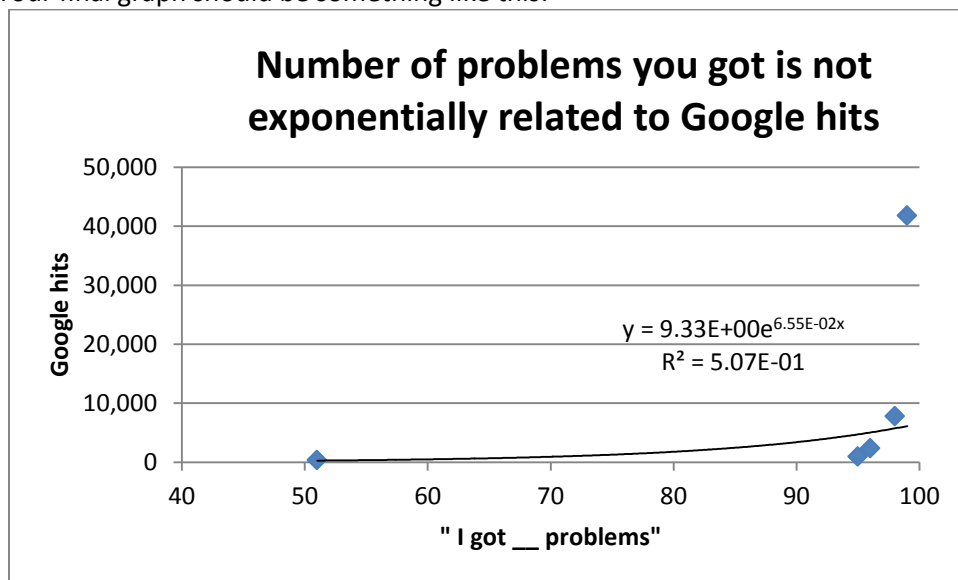
Section 6: Legends, graph titles and captions
Excel has automatically inserted a chart title based on our y-axis label, but that is not a good title. All graphs need descriptions that explain the graph. Scientific writing usually uses captions, complete sentences that explain the relevant experimental details needed to understand the graph. For visual presentations however, scientists often use titles. Titles should NOT be "Figure #" and they should NOT just repeat the axis labels. "Graph of X versus Y" is ok for a middle school science project title, but not for college level work. The reader can see your axes, so there's no point repeating them. Think about why you are drawing the graph and then describe it to the reader in the title. Go to Chart Tools, Layout, Chart title, and change the title to a meaningful summary.

Your graph should have a legend if and only if there are different data sets being graphed. It should not have a legend for one data set and a trend line, so delete the legend on the current graph. When you do have multiple data sets, make sure the legend can distinguish them in the form you are turning in. Most
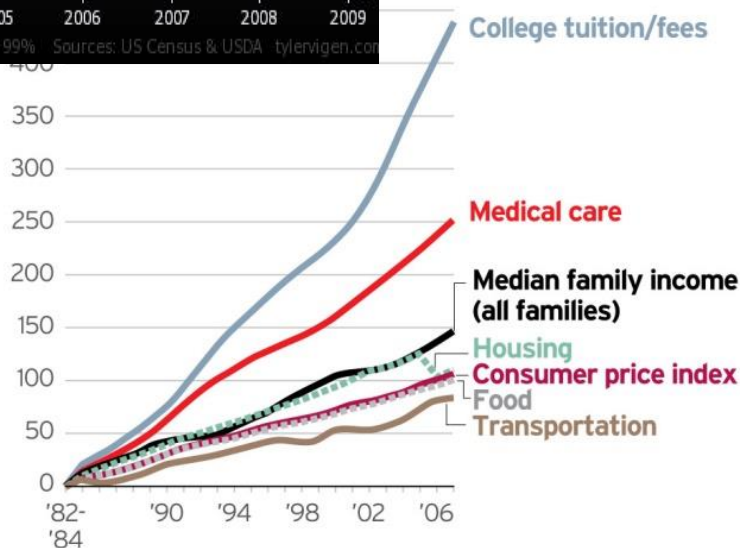
campus computer labs only have black and white printers, so if you print a blue squares and black squares, they'll look the same. Always make sure your data markers are distinct on the format you are turning in. Your final graph should be something like this:

Number of problems you got is not exponentially related to Google hits

$y = 9.33E+00e^{6.55E-02x}$
$R^2 = 5.07E-01$

Come up with title for the following graphs, and comment on whether the legends are useful or not.
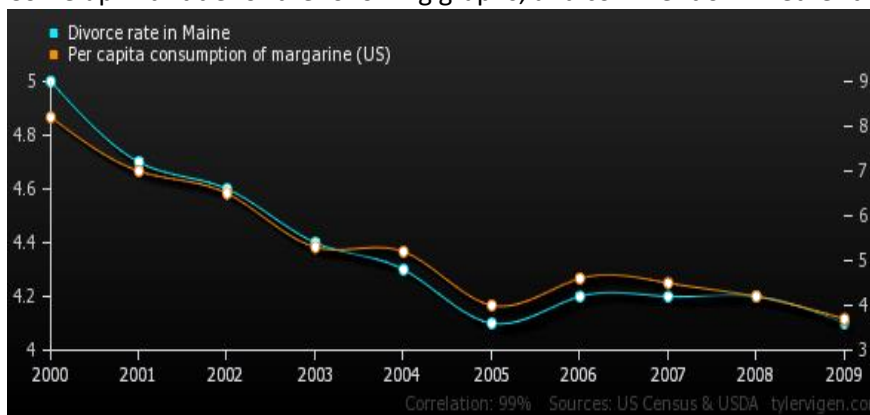
Notice that the top graph has 2 different y axes. The left axis corresponds to divorce and the right axis corresponds to margarine purchases. I am happy to teach fancy graphing techniques like this to anyone who wants to learn, but you do not need to know this for General Chemistry. Focus on making sure you understand the basics: graph types, axes, trendlines and labels.