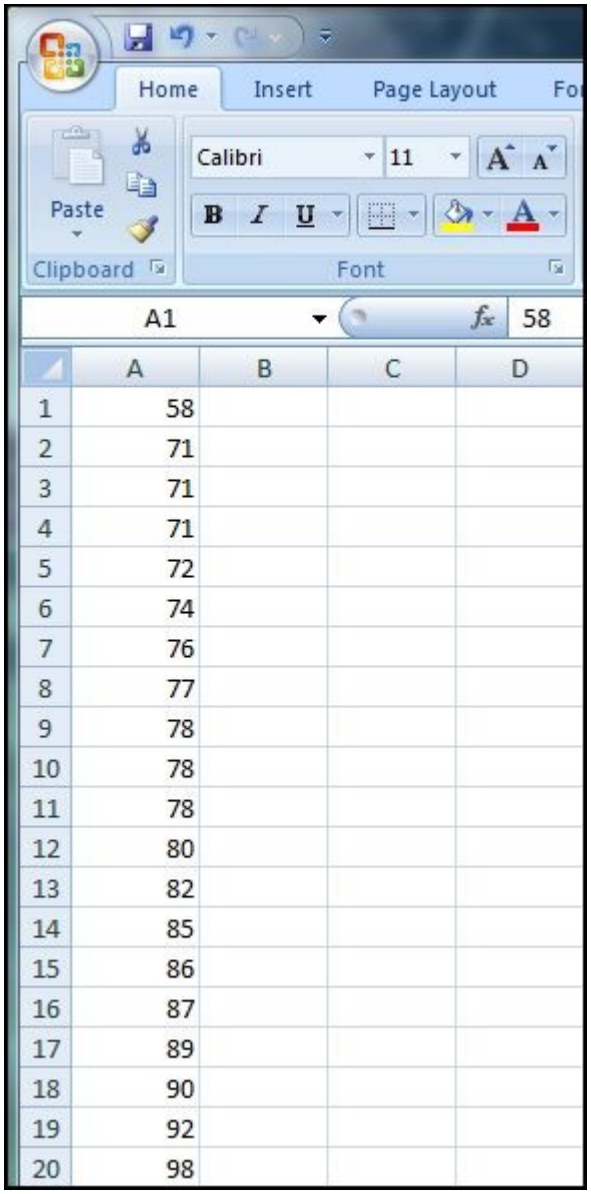
SAMLab Tip Sheet #4
Creating a Histogram

Another great feature of Excel is its ability to visually display data. This Tip Sheet demonstrates how to create a histogram and provides a general overview of how to create graphs, which is covered more specifically in Tip Sheet #5.

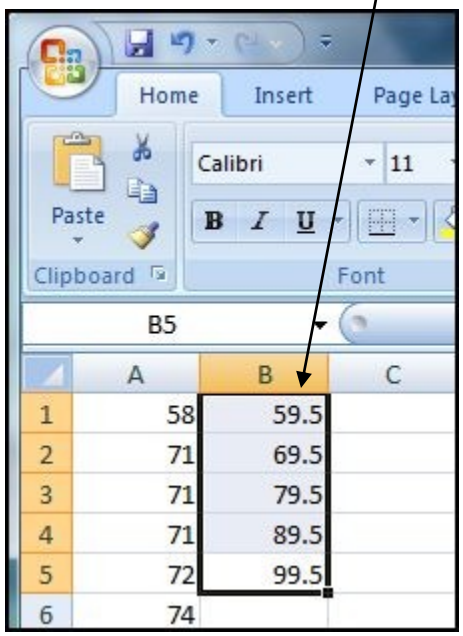
Creating a Histogram

We'll use the data below to learn how to create a histogram. Imagine that these data, which are fictitious, represent students' scores on the final exam of a statistics course.



	A	B	C	D
1	58			
2	71			
3	71			
4	71			
5	72			
6	74			
7	76			
8	77			
9	78			
10	78			
11	78			
12	80			
13	82			
14	85			
15	86			
16	87			
17	89			
18	90			
19	92			
20	98			

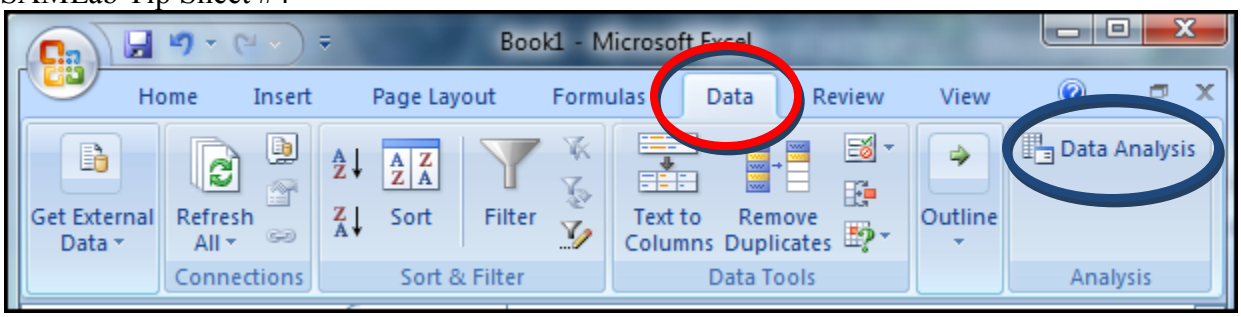
First, we need to decide how many class intervals we want for our histogram. Looking at our data, it looks like 5 intervals of width 10 will work best[1]. Next, we need to find the upper real limits[2] of our intervals and enter them into our spreadsheet (highlighted below)



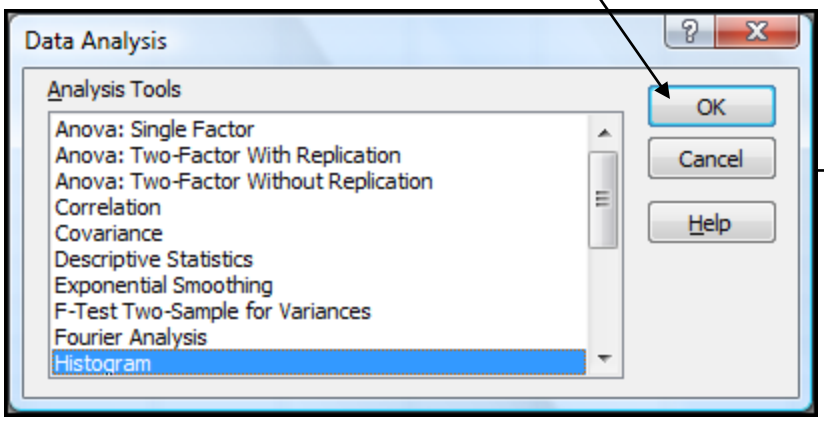
	A	B	C
1	58	59.5	
2	71	69.5	
3	71	79.5	
4	71	89.5	
5	72	99.5	
6	74		

Next, select "Data Analysis," which can be found under the "Data" tab (shown at the top of the next page). If you don't see "Data Analysis…," go to the end of this Tip Sheet for how to make it appear.

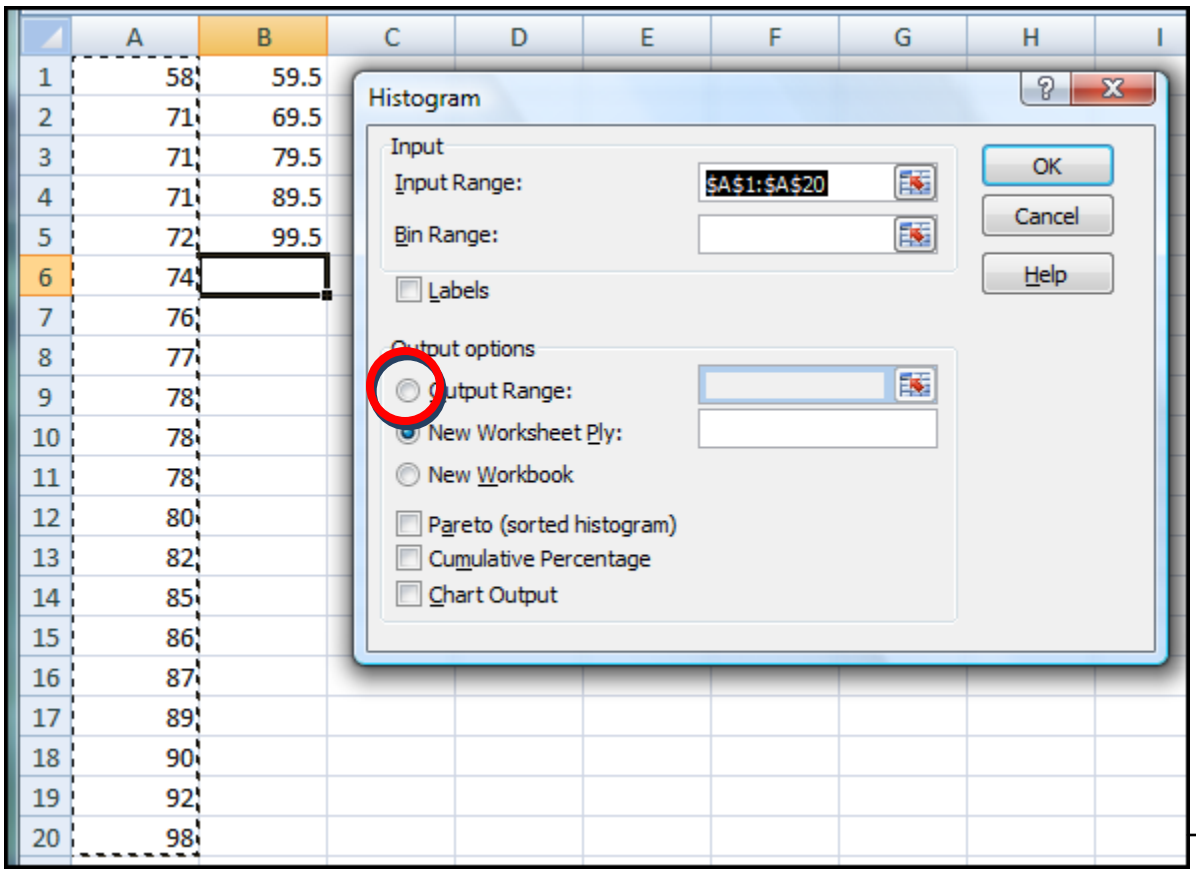
SAMLab Tip Sheet #4

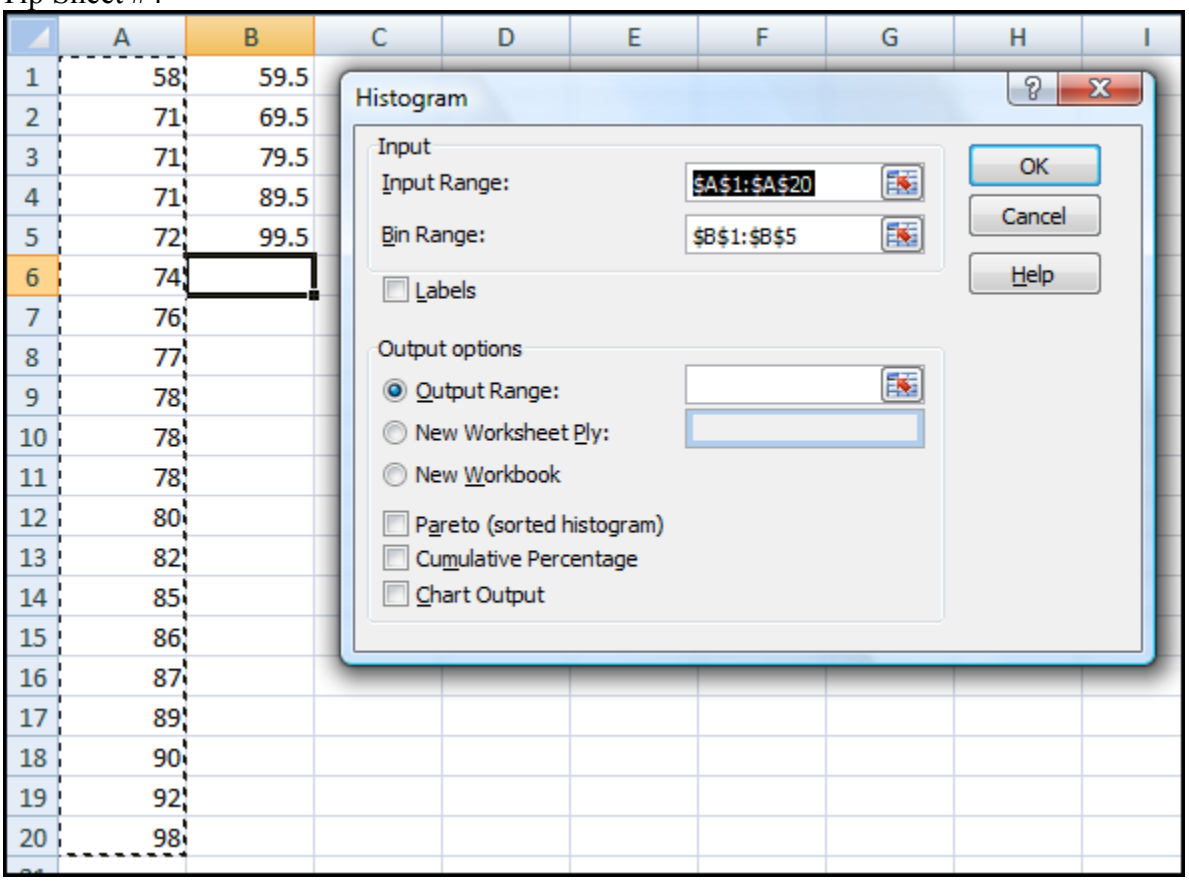


After selecting "Data Analysis," you will see the screen below. Highlight "Histogram" as shown and click OK.

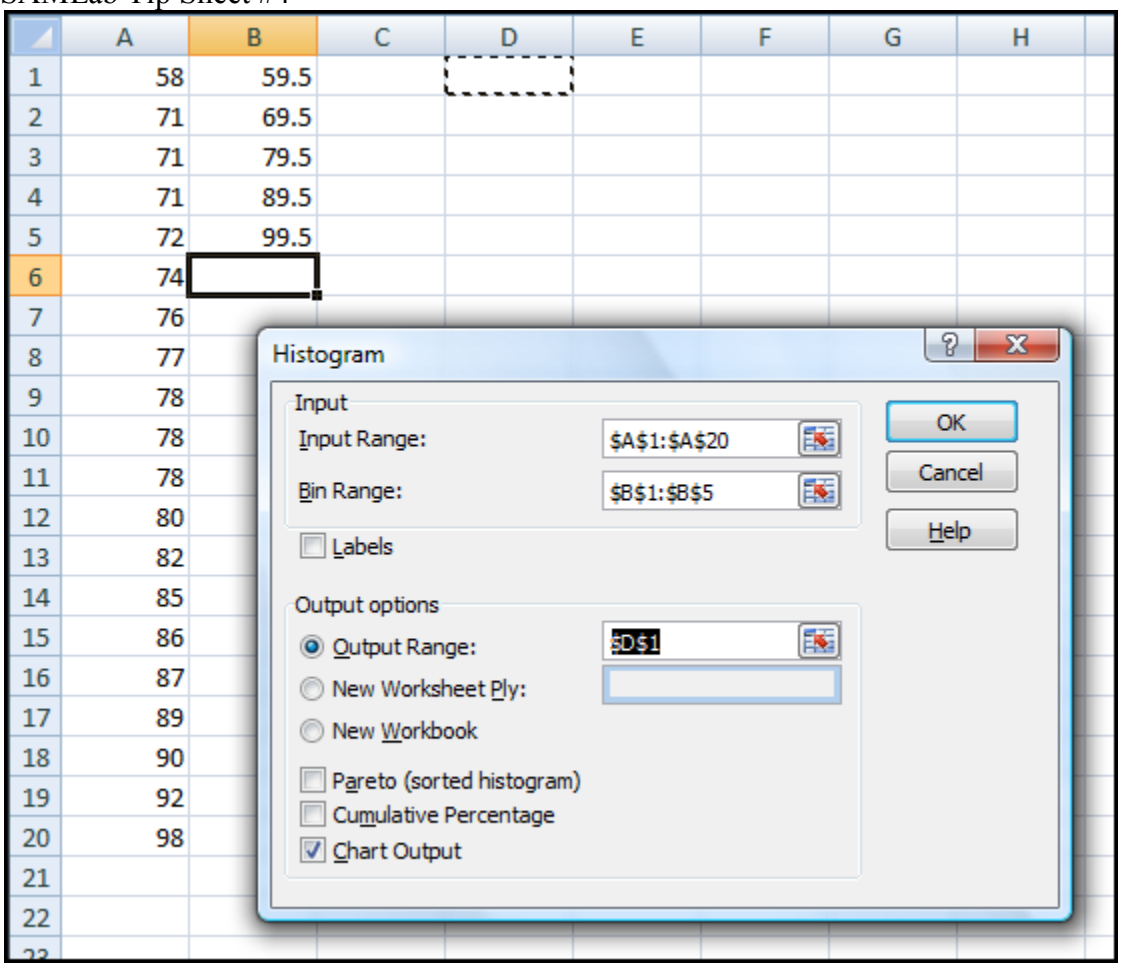


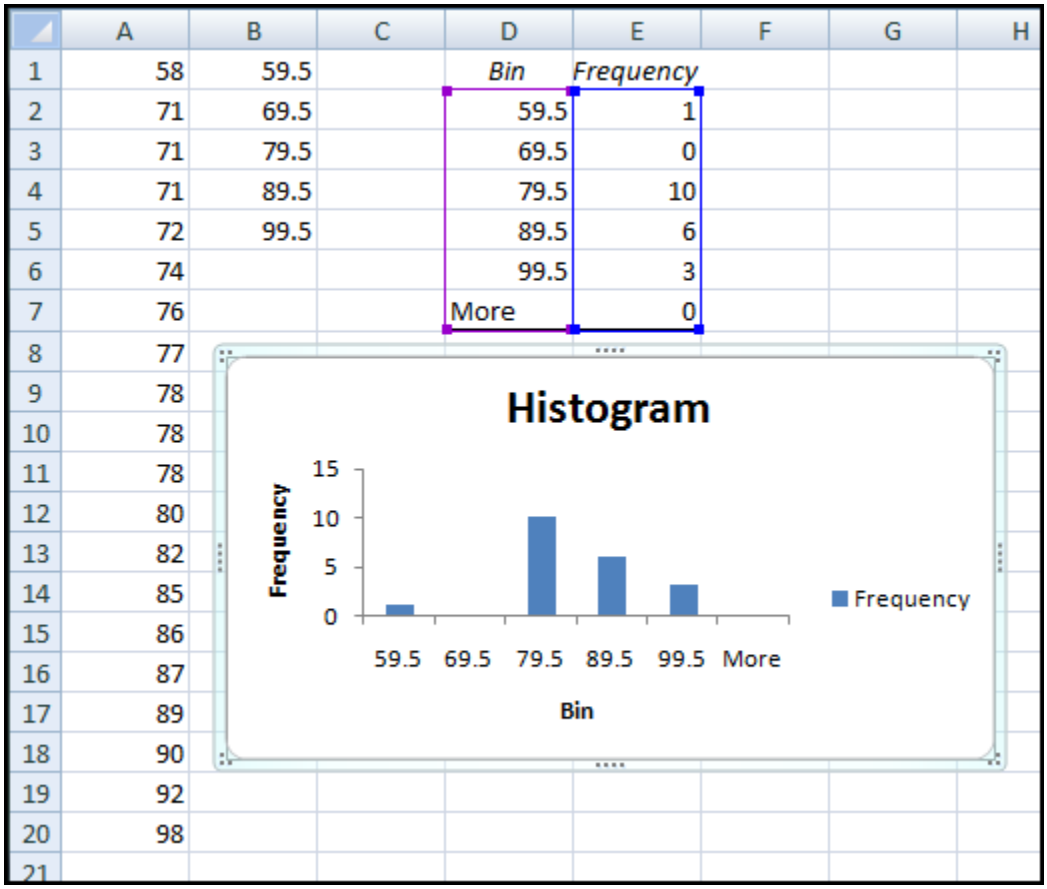
With the cursor flashing in the field labeled "Input Range," select your data set as shown below. For the "Bin Range" field, follow the same procedure but select your upper real limit column instead. Next, click on the radio button beside "Output Range" (circled). Once you've done all of this, your screen should resemble the first image on the next page.





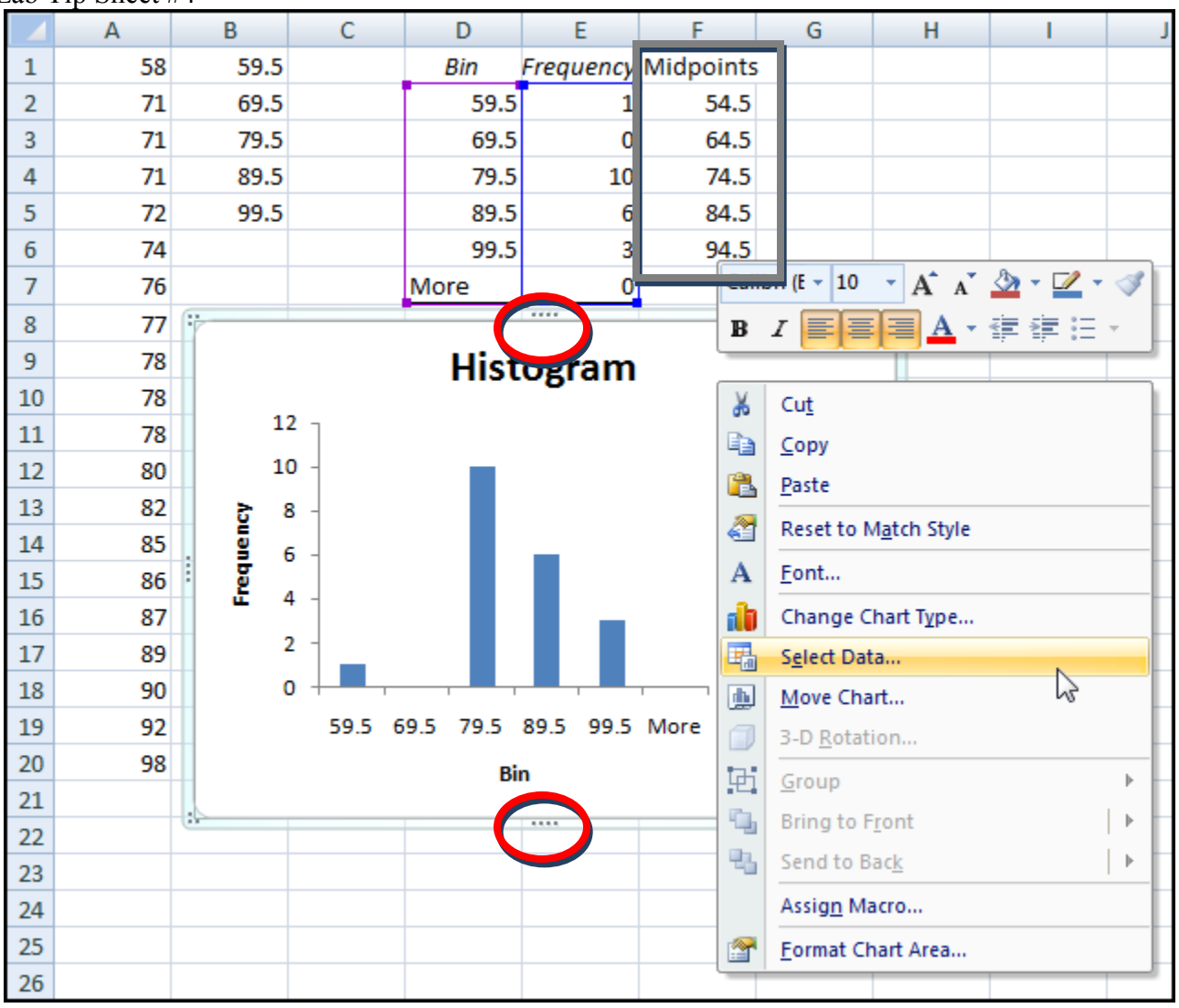
BE CAREFUL HERE because this is a spot where people often make mistakes. We've just clicked on the radio button next to "Output Range," and look which field is highlighted. Excel has highlighted the "Input Range" field. If we were to select a cell for our output at this time, we would actually select a new input range. The next step then is to click in the field next to "Output Range," making the cursor flash in that field and then select a cell on the spreadsheet near your data (shown on the next page). Click in the box next to "Chart Output" so that a check appears. Your screen should now resemble the first image on the next page.





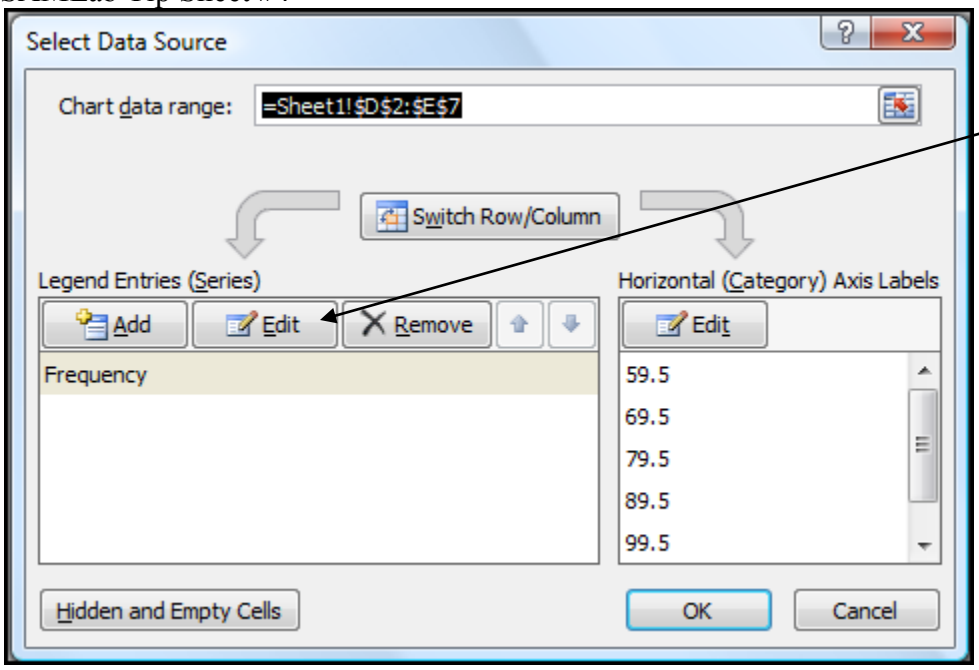
After clicking OK, your spreadsheet should look something like the one on the bottom. You may have to move your chart (the box labeled "Histogram") around so that it is visible. Note that you have also created a **grouped frequency distribution** ("Bin" & "Frequency" columns) along with your histogram. This will be useful in creating grouped frequency distributions (Tip Sheet #5). Next, we have to make some changes to our chart to make it more visually appealing and more technically correct. Continue on to the next page to see how this is done.

SAMLab Tip Sheet #4

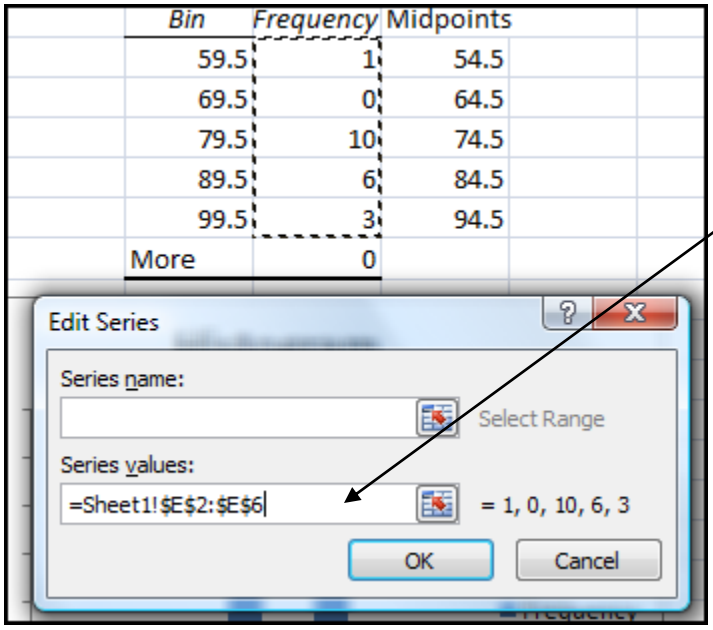


Our first step is to make our chart taller by dragging one of the central "…." (circled in red) away from the center of the chart. Next, we need to make a new column of midpoints (inside grey rectangle) so that we can center the bars of our histogram over the midpoints of the interval later on. Finally, **right click** on the chart and select "Select Data" from the menu.

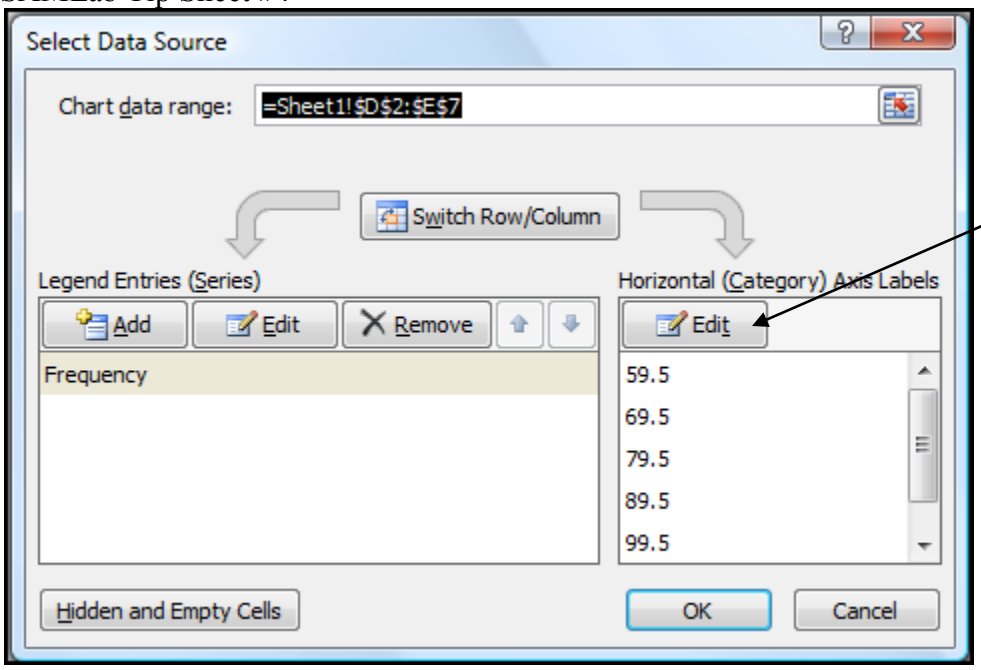
SAMLab Tip Sheet #4



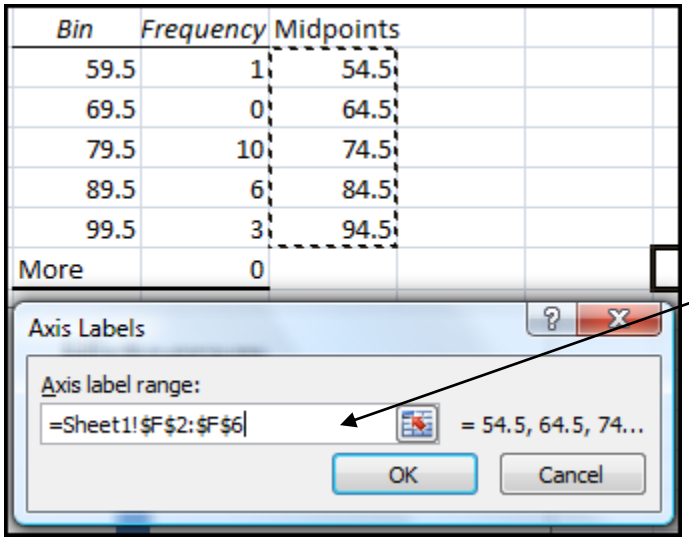
First, click on the "Edit" button under the "Legend Entries (Series)" section.



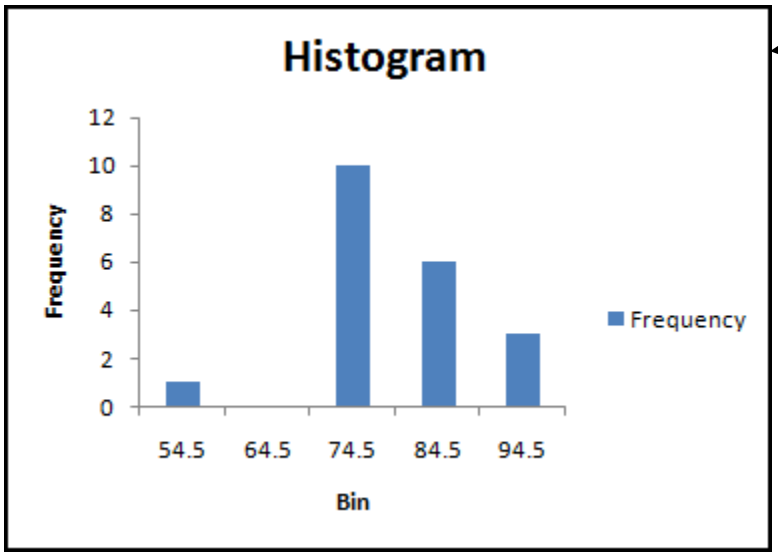
A box titled "Edit Series" will open. Under "Series values," delete the current contents in the field. Now, with your cursor flashing in the "Series values" field, select the cells under the "Frequency" heading, but DO NOT select the cell to the right of "More." Click OK.



Now, select "Edit" under the "Horizontal (Category) Axis Labels" section.

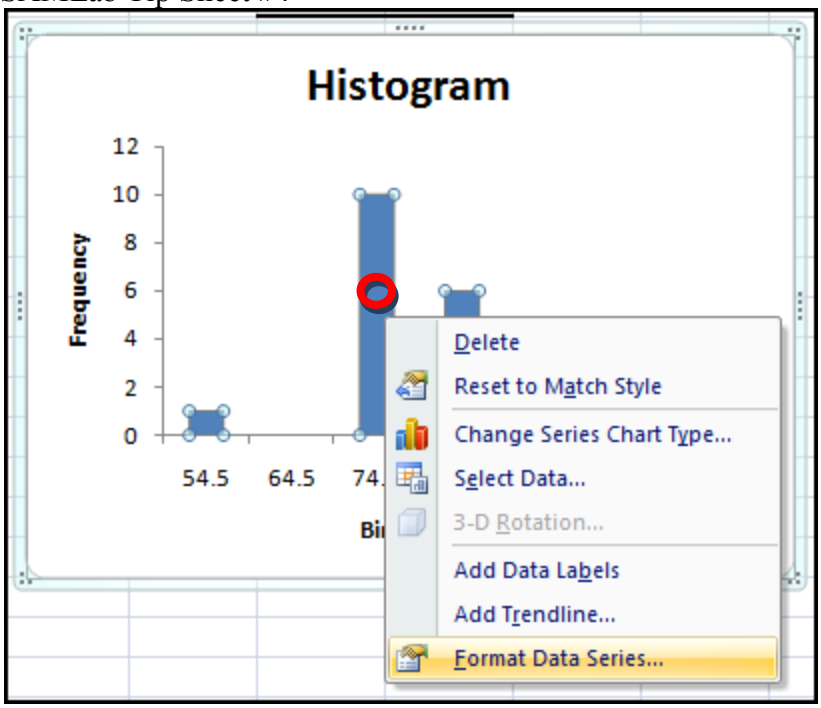


A box titled "Axis Labels" will open. Delete the contents of the field under the label "Axis label range." Now, with your cursor flashing in the "Axis label range" field, select the cells under the "Midpoints" heading. Click OK. You will come back to the "Select Data Source" window that you started in. Click OK to finish.

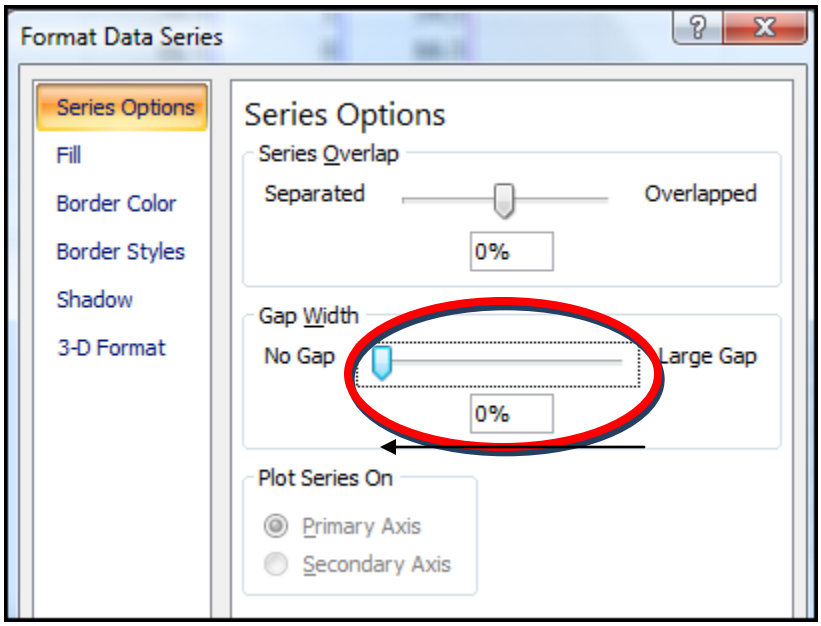


Your histogram should now look like this. It doesn't look quite right yet, however. We still have a few more changes to make.

SAMLab Tip Sheet #4

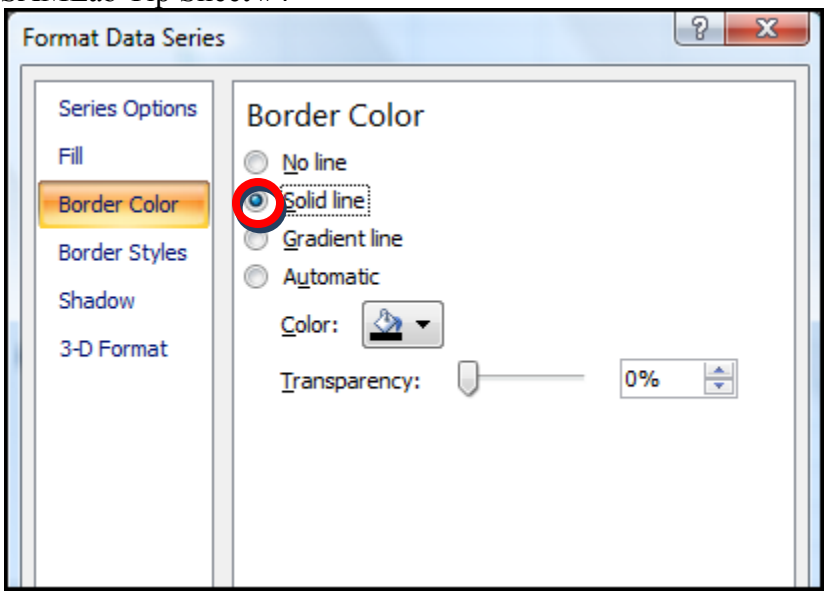


Right click on one of the bars in the chart (for example, where the red circle is located to the left) to bring up a menu. Select "Format Data Series…" from the menu.



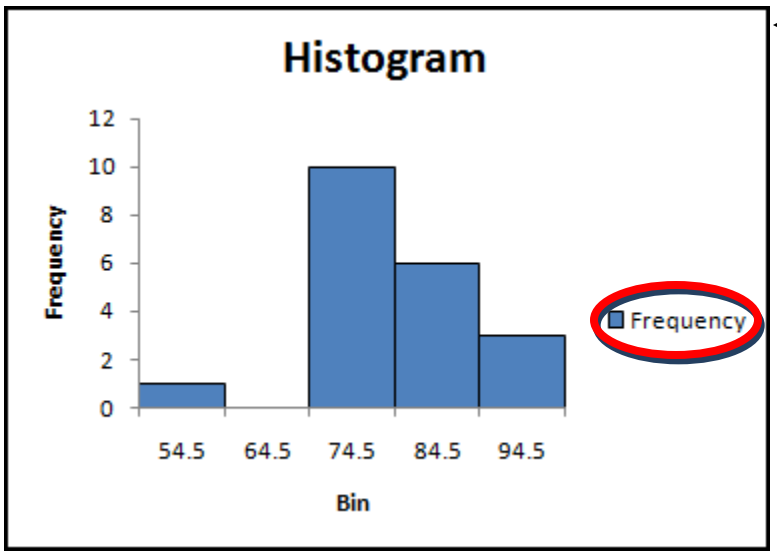
The "Format Data Series" window will open. Click and hold on the slider under "Gap Width" and drag the slider all the way to the left so that the value in the box underneath is at 0% (circled to the left).

SAMLab Tip Sheet #4

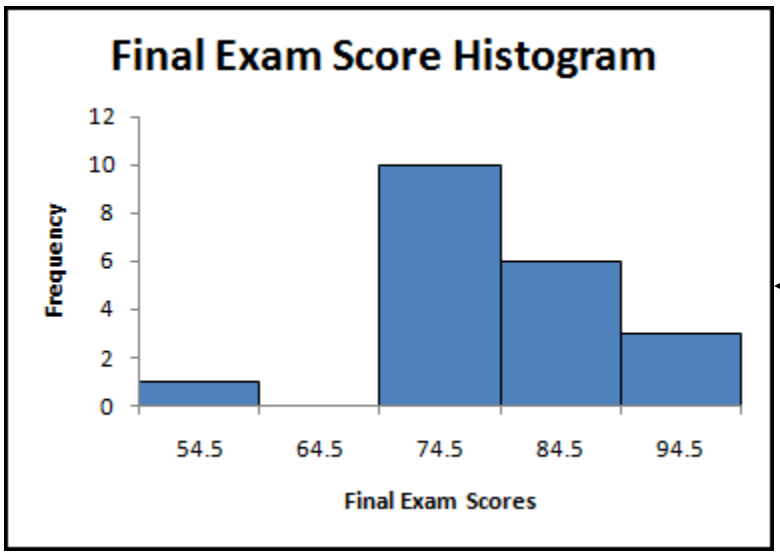


Now, in the same window, click on the menu "Border Color" located to the left of the window (highlighted to the left). Select the radio button "Solid line." Now click on the "Close" button to return to the spreadsheet.

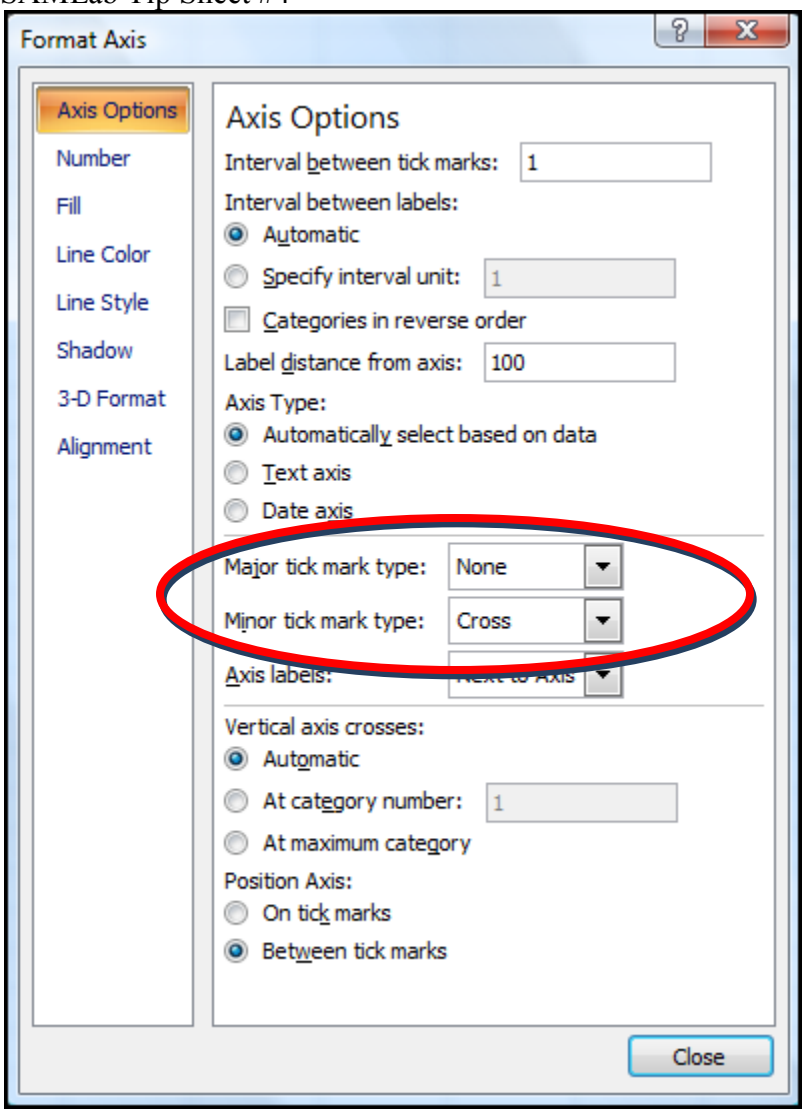
Your histogram should now look like the one to the left. Now, we want to change some titles. To change the title of the x-axis, click on the current title (named "Bin"). A textbox will appear and you can enter your new title. For our example, we will call the x-axis "Final Exam Scores." To change the title of the histogram, do the same thing: click on the current title (named "Histogram") and a textbox will appear. Type in a new title (we will use "Final Exam Score Histogram"). Finally, we don't need that legend, so **right click** on the legend box on the graph (circled in red to the right) and select "Delete."



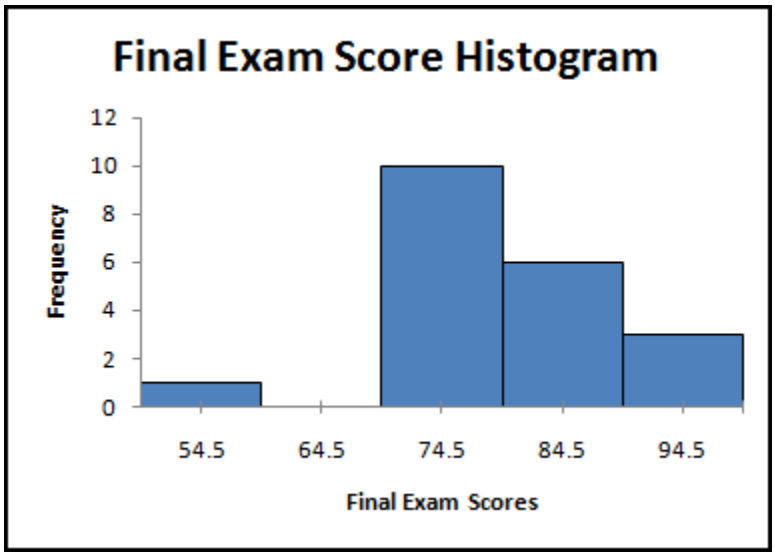
Your histogram should now resemble the one to the left. Now there is just one more thing to do.



SAMLab Tip Sheet #4



Right click on the x-axis of your chart (the easiest way to do this is to right click on one of the numbers in your x-axis). Select "Format Axis…: from the menu that comes up. You should see the window shown on the left. In the "Major tick mark type" menu, select "None." In the "Minor tick mark type" menu, select "Cross" (these areas are circled in red to the left). Click "Close" at the bottom of the window. This will return you to the main spreadsheet window.



Your histogram should look very similar to if not exactly like the one shown on the left. The histogram is now technically correct (the bars are centered on the midpoints of the intervals and each bar covers an entire interval). You certainly can still make changes by resizing the chart of fiddling with the y-axis but you can stop here if you like.

Notes

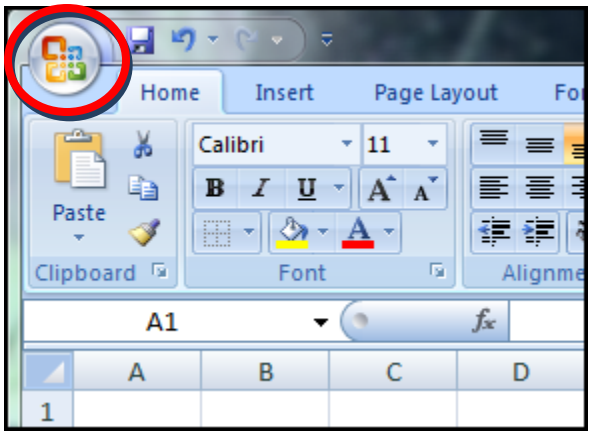
[1]To find the interval width (i), we need to divide the range of our data by the number of intervals (equation below). Dividing 40 by 5 gives 8 as our interval size. Generally, however, widths of 2, 5, 10, & 20 are preferred because they are easier to work with so we round our interval size up to 10.

$$i = \frac{Range}{\#\ of\ intervals}$$

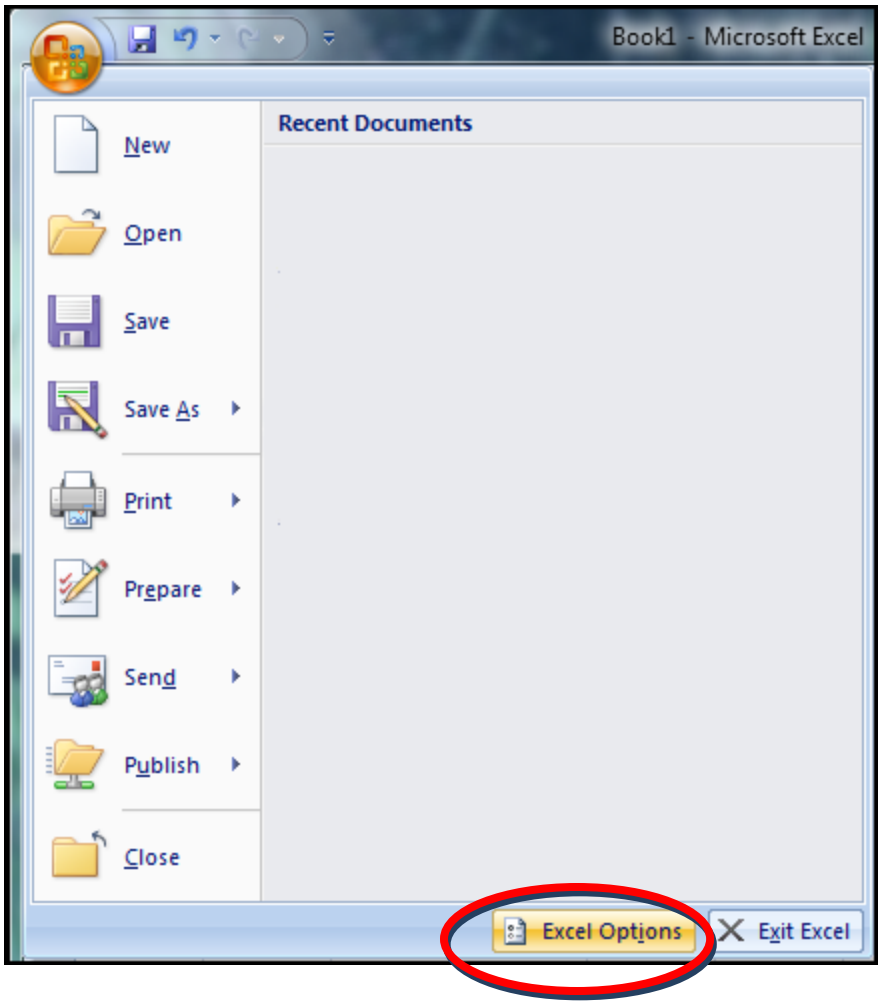
[2]Our intervals are 50-59, 60-69, 70-79, 80-89, 90-99. These numbers represent the upper apparent limits of our class intervals. The upper real limits are just ½ unit above our upper apparent limits. Thus, they are 59.5, 69.5, 79.5, 89.5, & 99.5. We have to use the upper real limits in order for Excel's Histogram function to work the way we want it to.

SAMLab Tip Sheet #4

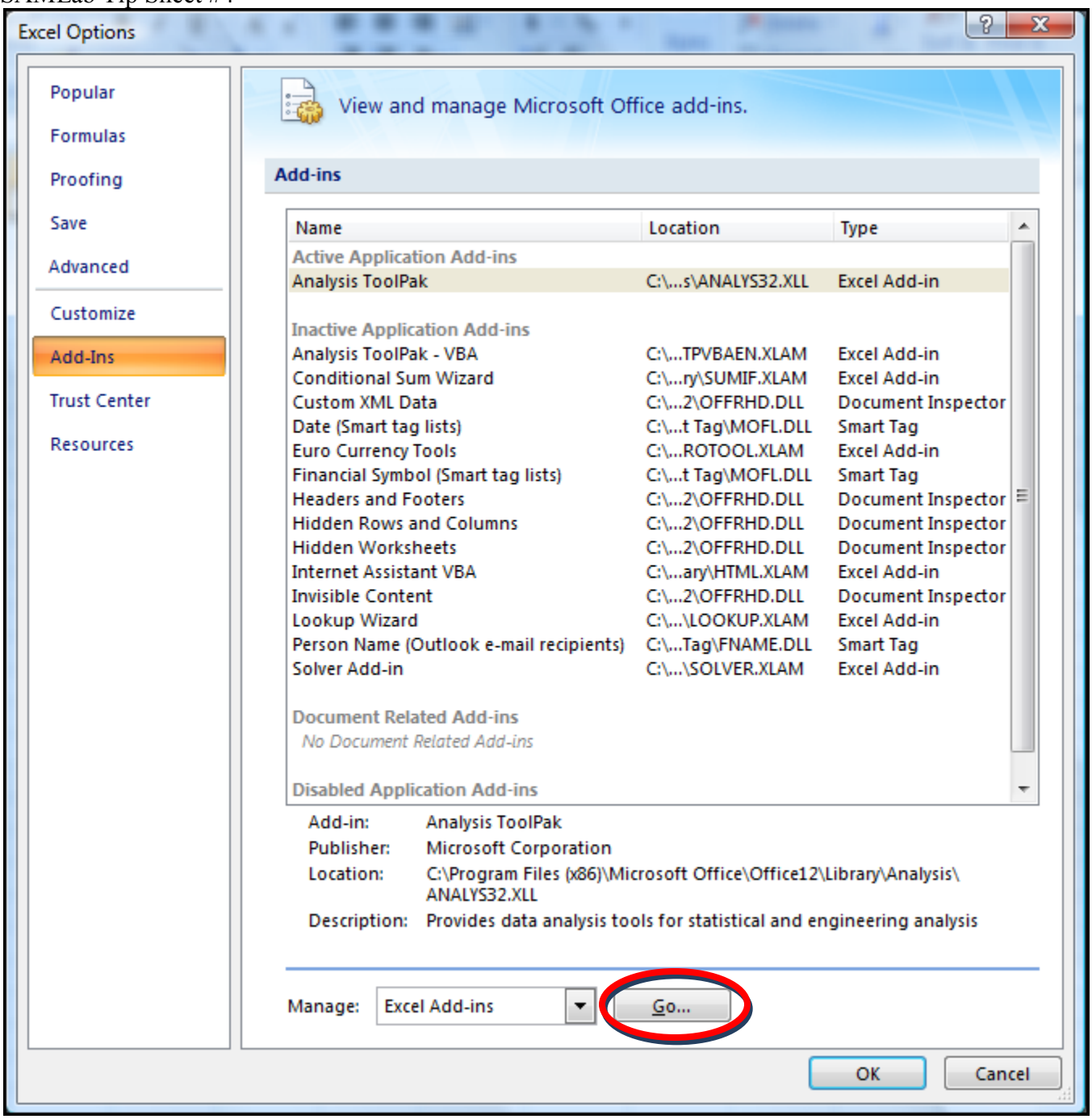
Making "Data Analysis" Appear in the Data Tab (Note: You *may* need your Excel software CD)



First, click on the Office Button (circled to the left). The following menu will come up. Select "Excel Options" at the bottom of the menu (circled below).

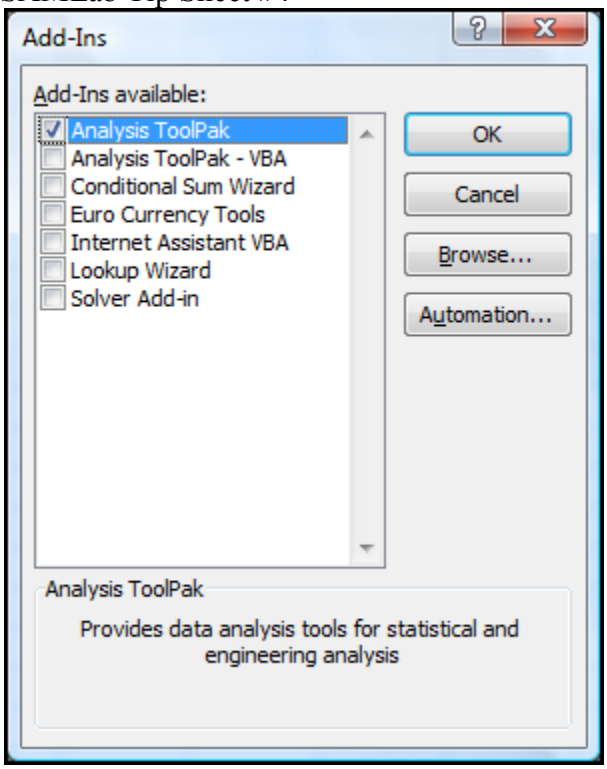


SAMLab Tip Sheet #4



In the "Excel Options" menu, select "Add-Ins" (currently highlighted above) from the options on the left. Make sure that the menu to the right of "Manage" is set to "Excel Add-ins" and click "Go…" (circled above). The window on the next page will come up.

SAMLab Tip Sheet #4



In the "Add-Ins" window, select "Analysis ToolPak" (like in the picture to the left). Click OK. Excel will now install the Data Analysis option which you will find under the "Data" tab (explained earlier in this tip sheet).