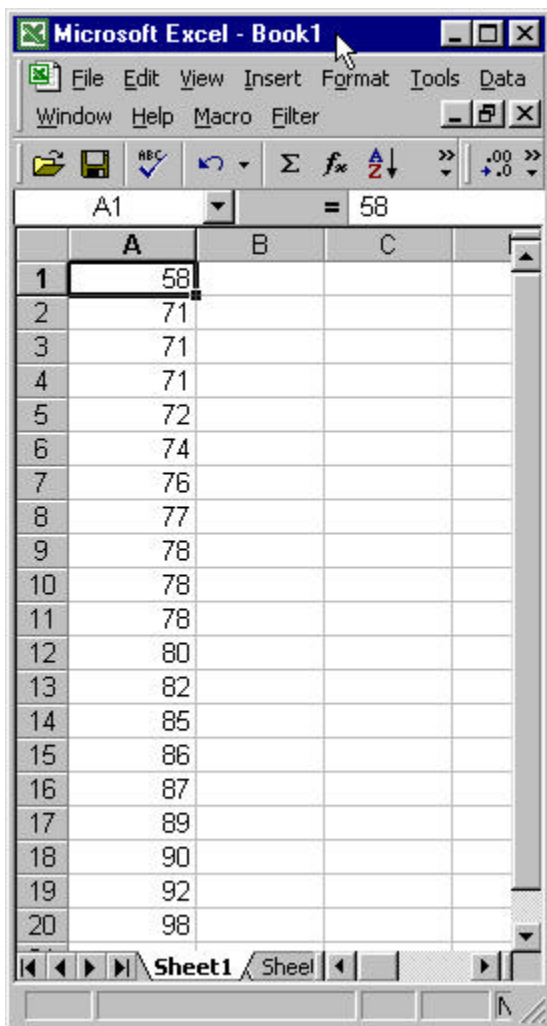**Creating a Histogram**

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.
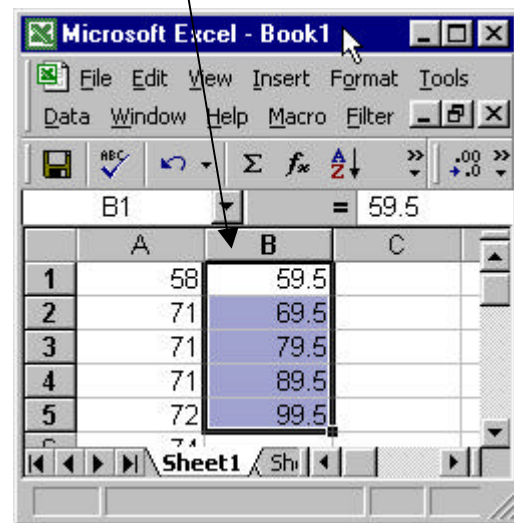
Creating a Histogram

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.

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).

Next, select "Data Analysis" from the Tools Menu (Shown Below). If you don't see "Data Analysis…," go to the end of this Tip Sheet for how to make it appear.

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

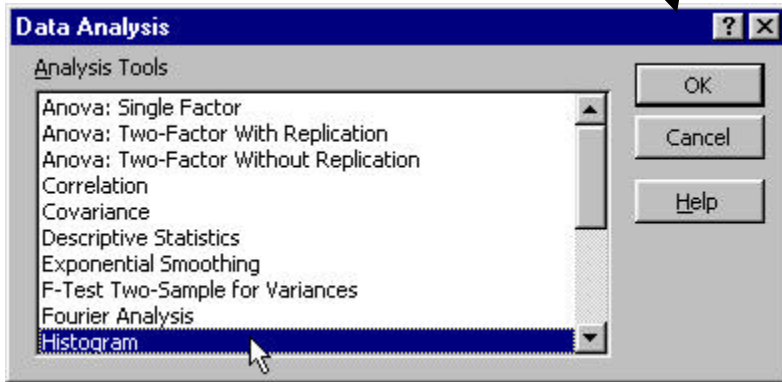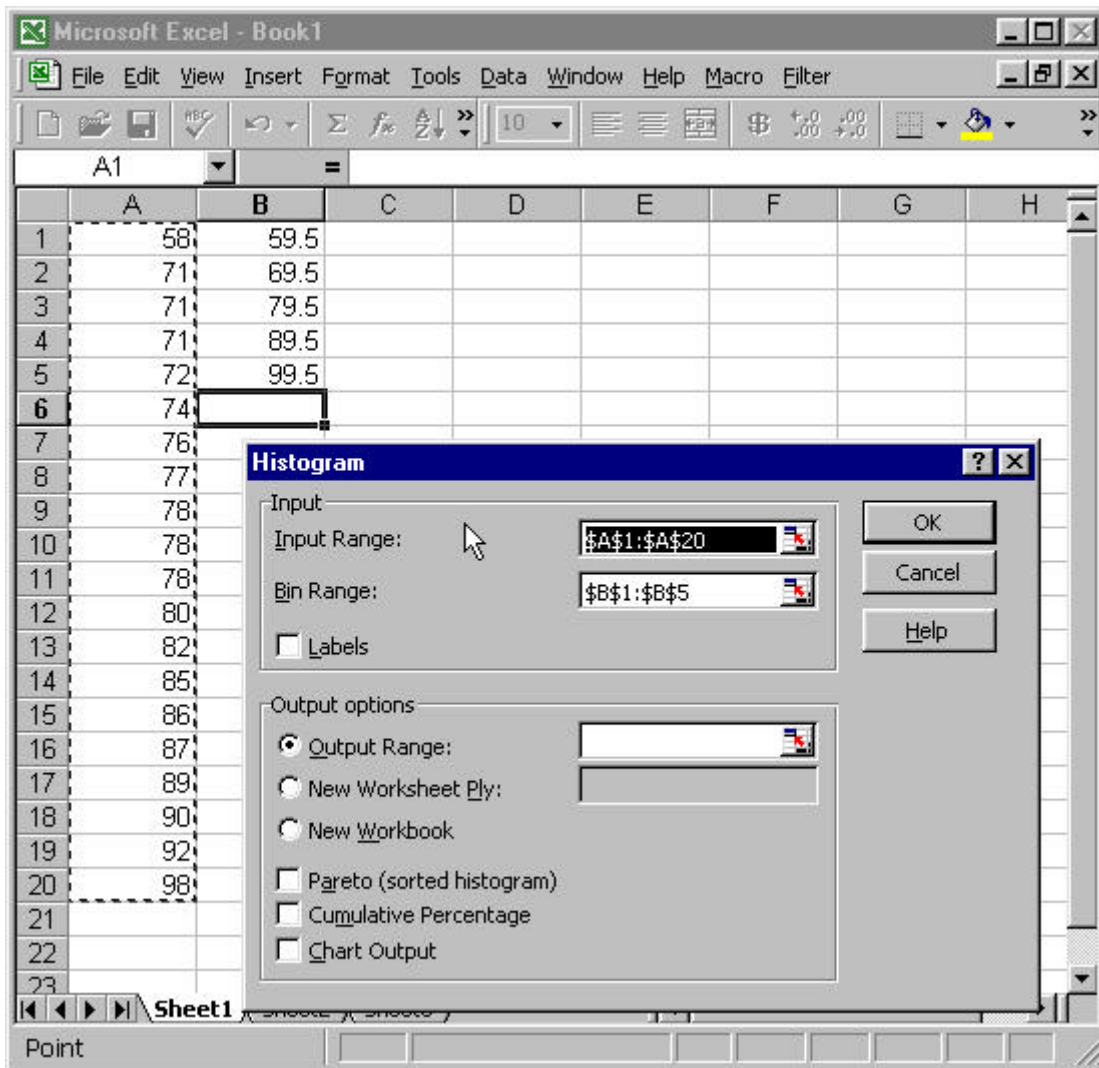With the cursor flashing in the field labeled "Input Range," select your data set as shown below. For the "Bin Range" field, follow the same procedure but select your upper real limit column instead. Next, click on the radio button beside "Output Range" (circled). Once you've done all of this, your screen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

Data Analysis

Analysis Tools

Anova: Single Factor
Anova: Two-Factor With Replication
Anova: Two-Factor Without Replication
Correlation
Covariance
Descriptive Statistics
Exponential Smoothing
F-Test Two-Sample for Variances
Fourier Analysis
Histogram

OK
Cancel
Help

Microsoft Excel - Book1

File   Edit   View   Insert   Format   Tools   Data   Window   Help   Macro   Filter

A1

	A	B	C	D	E	F	G	H
1	58	59.5						
2	71	69.5						
3	71	79.5						
4	71	89.5						
5	72	99.5						
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							
21								
22								
23								

Histogram

Input

Input Range: $A$1:$A$20

Bin Range:

☐ Labels

Output options

○ Output Range:

⦿ New Worksheet Ply:

○ New Workbook

☐ Pareto (sorted histogram)
☐ Cumulative Percentage
☐ Chart Output

OK
Cancel
Help

Sheet1   Sheet2   Sheet3

Point

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.
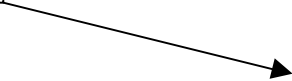
SAMLab Tip Sheet #4
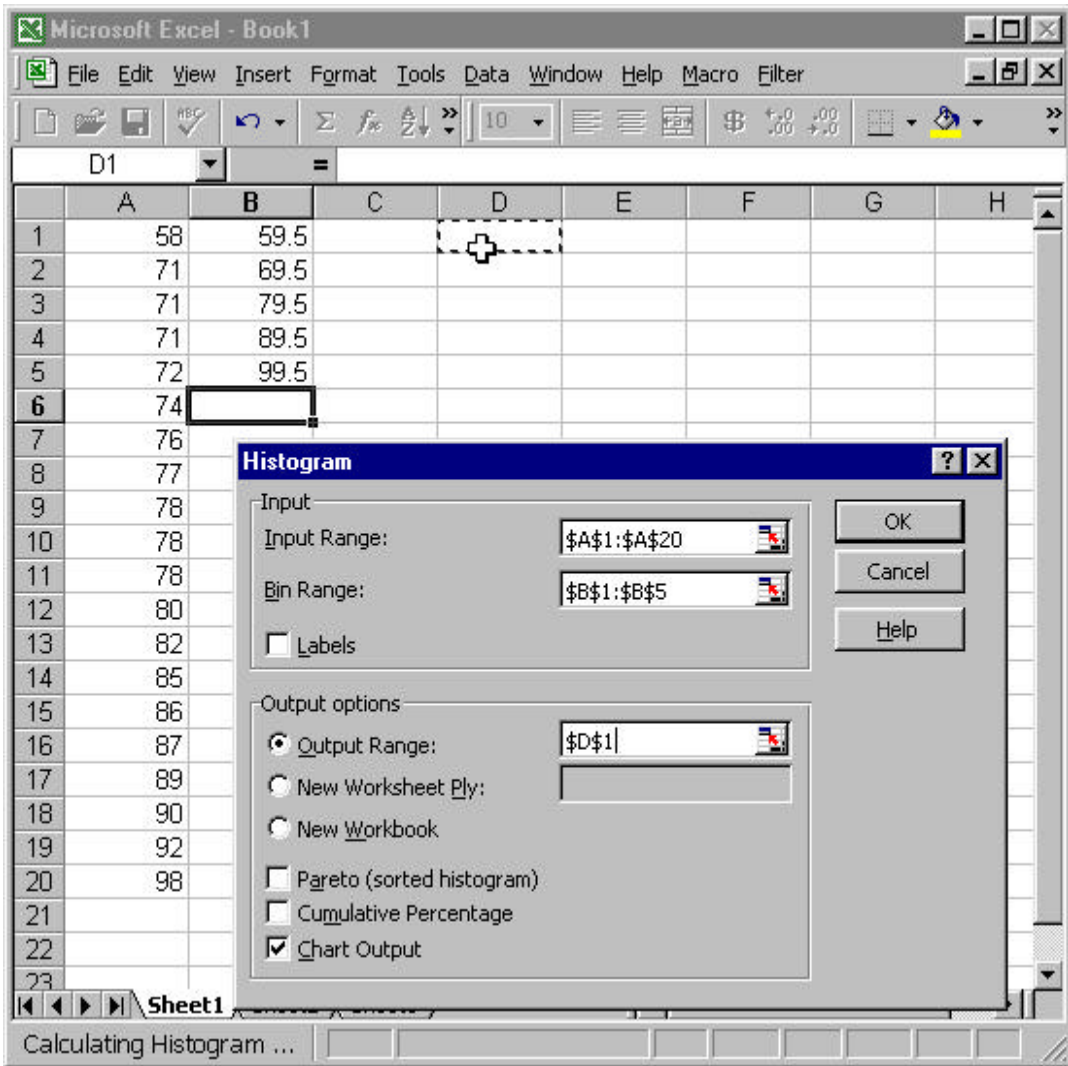
After clicking OK, your spreadsheet should look something like the one on the bottom. You may have to move your chart (the box labeled "Histogram") around so that it is visible. Note that you have also created a **grouped frequency distribution** ("*Bin*" & "*Frequency*" columns) along with your histogram. 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.
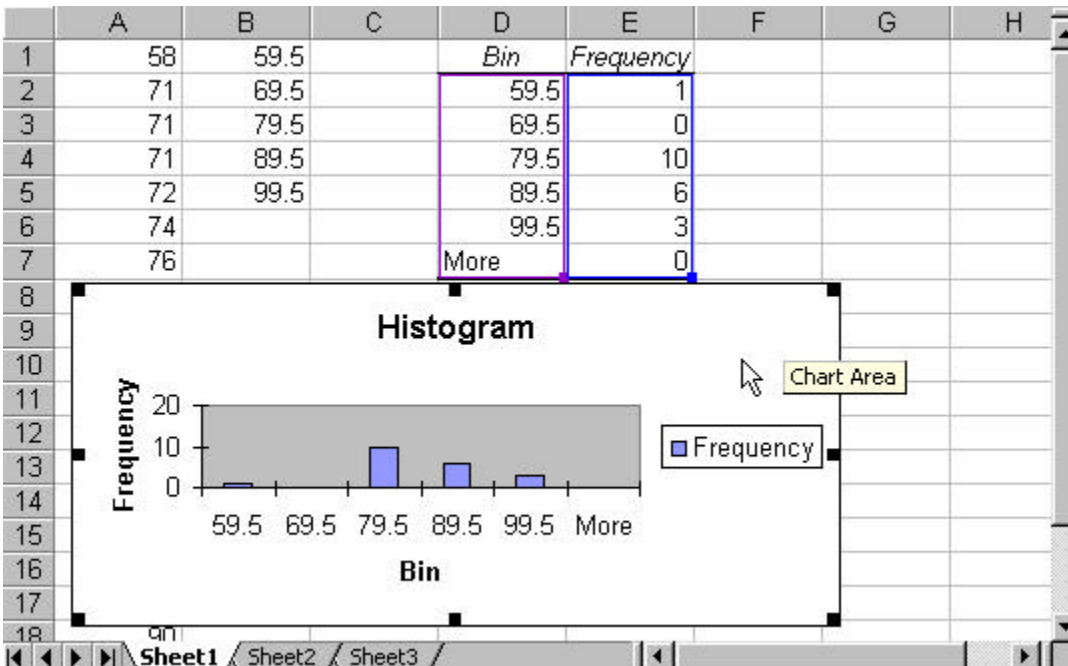
SAMLab Tip Sheet #4
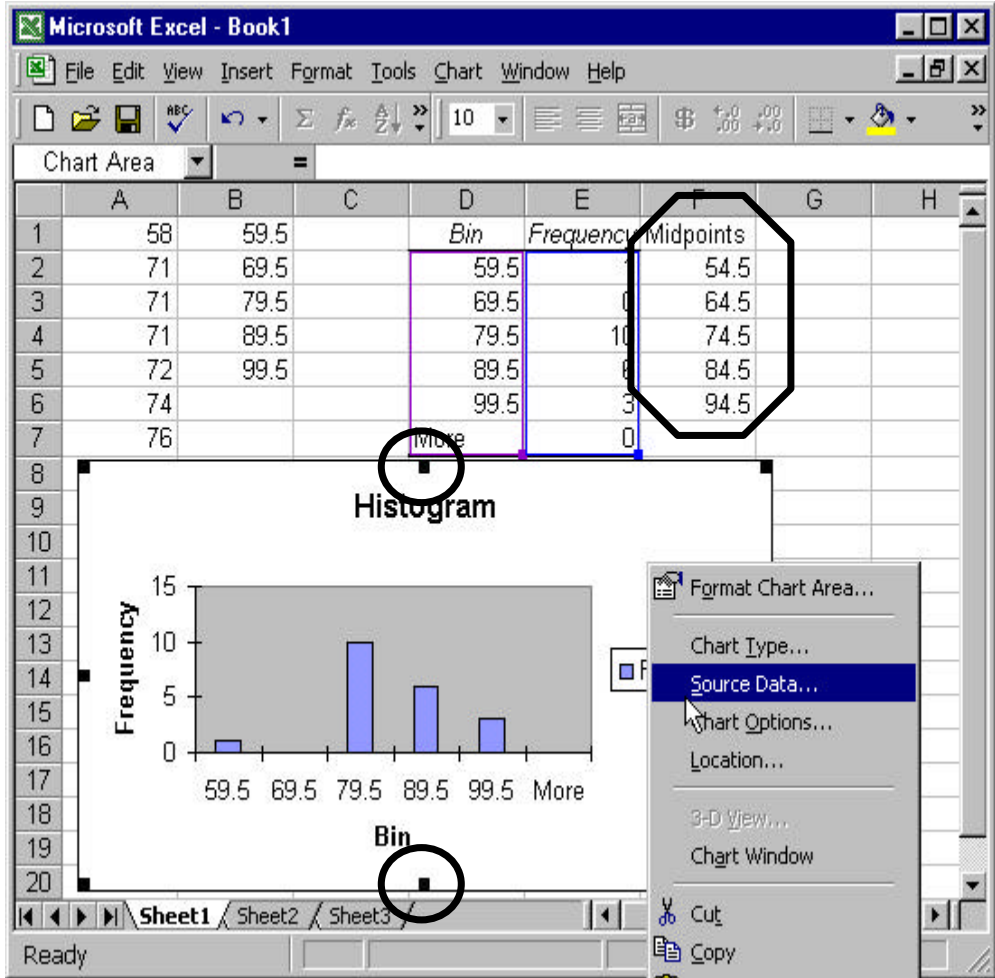
Our first step is to make our chart taller by dragging one of the central little black boxes (circled) away from the center of the chart. Next, we need to make a new column of midpoints (inside octagon) so that we can center the bars of our histogram over the midpoints of the interval later on. Finally, **right click** on the chart and select "Source Data" from the menu. Alternatively, you can select "Source Data" from the Chart menu displayed after clicking on the white portion of the chart.
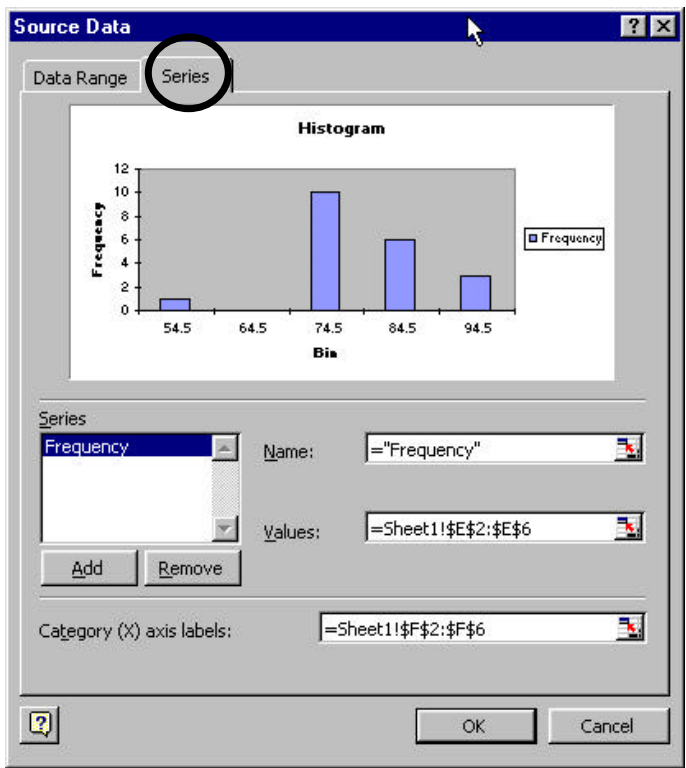

Once the "Source Data" window appears, click on the "Series" tab (circled) to make your window resemble the one on the left. Next, highlight the "Values" field and in the spreadsheet select the cells under the "*Frequency*" label but DON'T select the cell next to "More." Now we need to highlight the "Category (X) axis labels:" field and select our column of midpoints (only the numbers). Your screen should now look exactly like the one on the left. Once it does, Click OK.
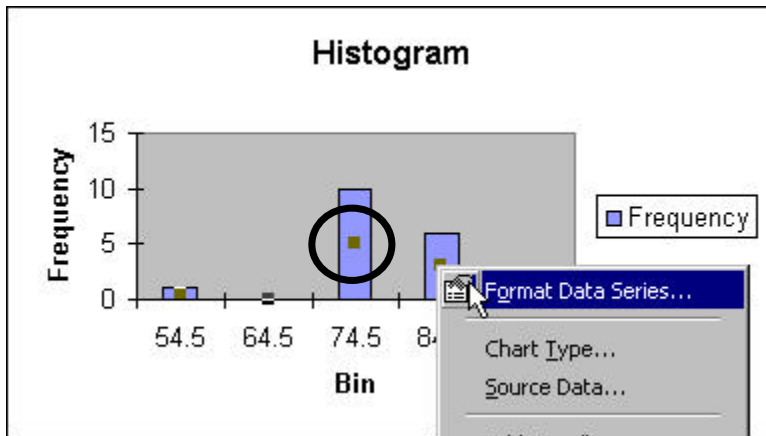
Next, click on one of the bars in the chart so that little boxes (circled) appear in them. Right click one of these boxes and select "Format Data Series…" from the menu.
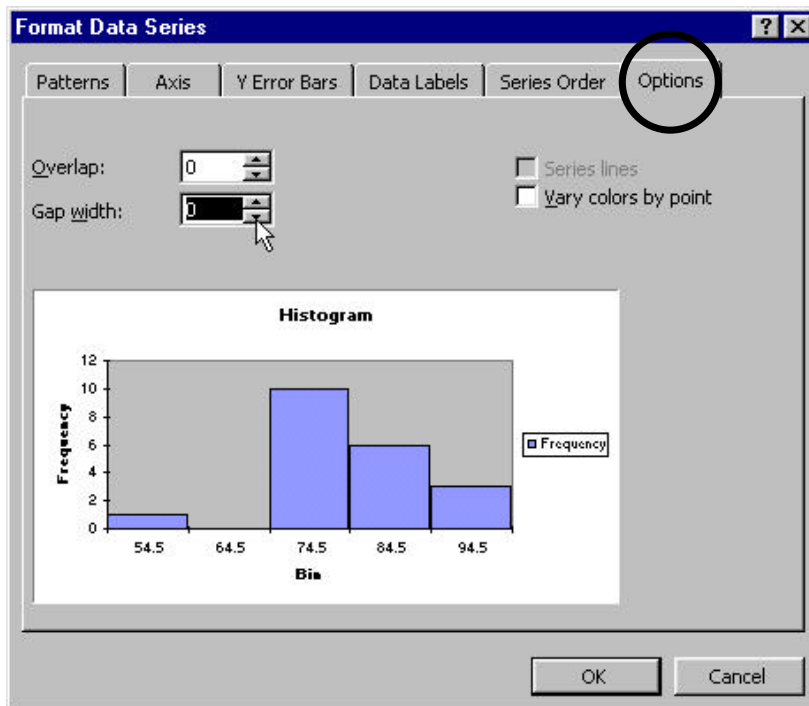
Once the "Format Data Series" window appears, select the "Options" tab (circled) so that your window resembles the one to the left. Next, click on the bottom arrow button in the "Gap width:" field (touched by the mouse pointer) until the value decreases to 0. Once your window looks just like the one to the left, click OK.
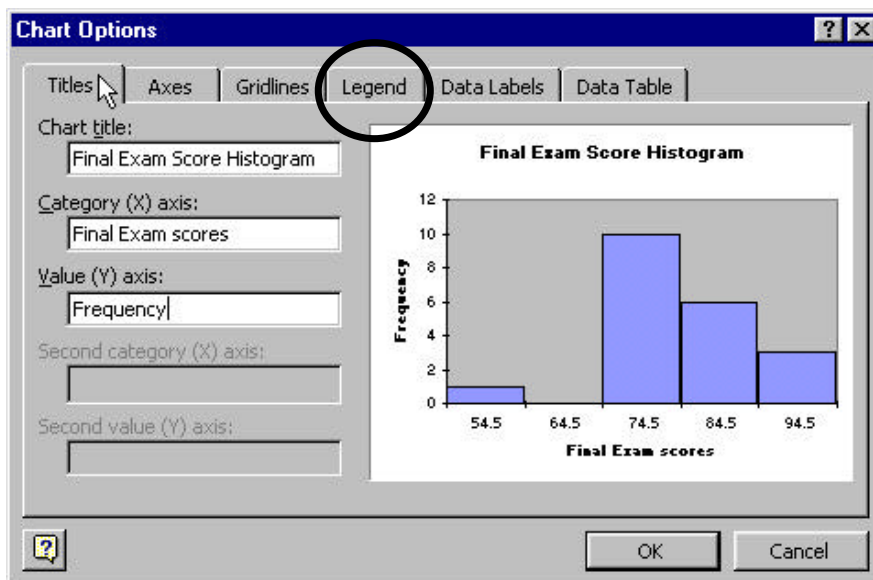
Next, right click in the white space of the chart and select "Chart Options" from the menu. The window that appears should resemble the one on the left. First, click the "Legend" tab (circled) and uncheck the "Show Legend" box. Now, click the "Titles" tab (touched by the mouse pointer) and change both the "Chart Title" and the "Category (X) axis:" fields to represent your data. Once you've done all of this, click OK.

## Final Exam Score Histogram
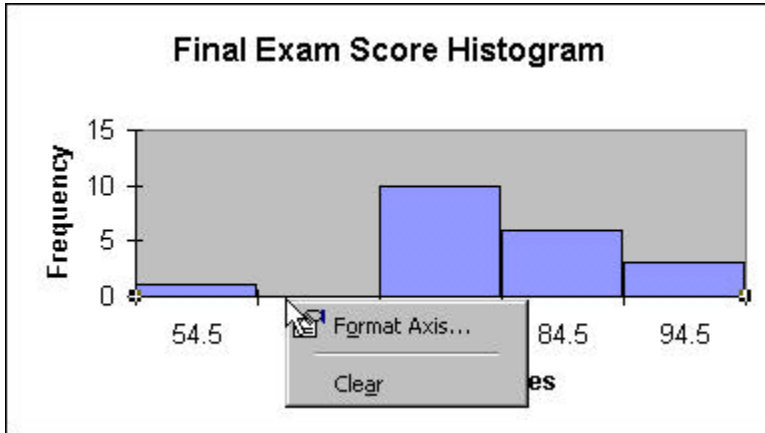
Next (we're almost done), we want the tick marks on the x-axis to correspond to the numbers. Right click on the x-axis of your chart and select "Format Axis…" from the menu.

Format Axis

Patterns | Scale | Font | Number | Alignment

Lines
- Automatic
- None
- Custom

Style:

Color: Automatic

Weight:

Sample

Major tick mark type
- None
- Inside
- Outside
- Cross

Minor tick mark type
- None
- Inside
- Outside
- Cross

Tick mark labels
- None
- Low
- High
- Next to axis

OK | Cancel
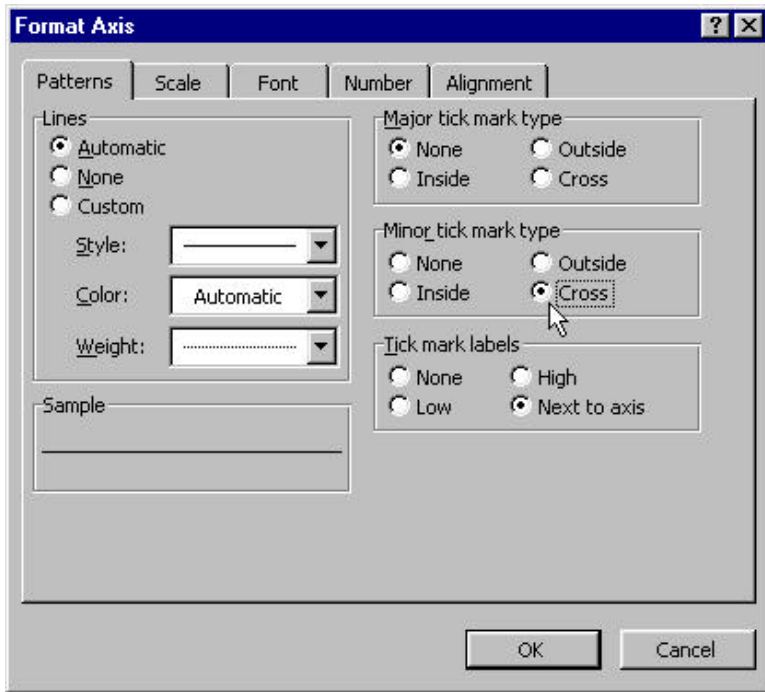
The window to the left should now appear and the "Patterns" tab should already be selected. In the "Major tick mark type" box click on the radio button next to "None" and in the "Minor tick mark type" box click on the radio button next to "Cross". Click OK.

Final Exam Score Histogram

Final Exam scores
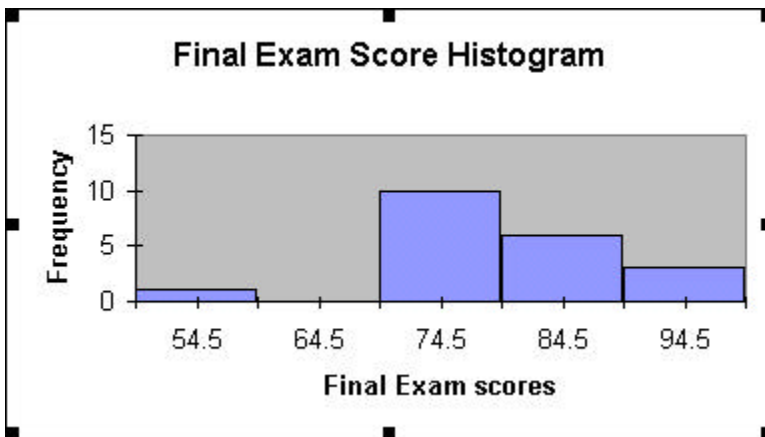
Your histogram should look very similar to if not exactly like the one on the left. The histogram is now technically correct (the bars are centered on the midpoints of the intervals and each bar covers an entire interval). You certainly can still make changes by resizing the chart or fiddling with the 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.

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

[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.

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

To make "Data Analysis" appear in the Tools menu, select "Add-ins" from the Tools menu (shown on the left). The window on the right should then appear. Check the box next to "Analysis ToolPak" and click OK. "Data Analysis" should now appear in the Tools menu.