

3: Graphing Data

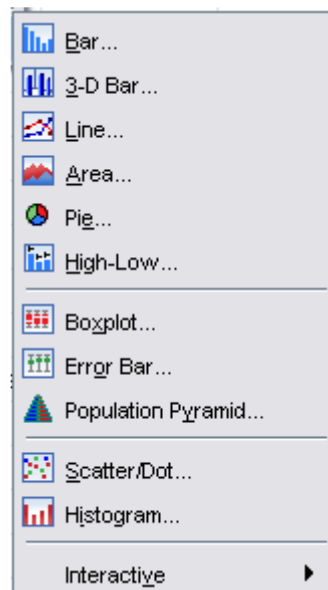
Objectives

- ◆ Create histograms, box plots, stem-and-leaf plots, bar graphs, scatterplots, and line graphs
- ◆ Edit graphs using the Chart Editor
- ◆ Use chart templates

(SPSS graphics has changed considerably since this section was written. SPSS now gives you the option of using Chart Builder or Legacy Dialogs. The latter are closer to previous versions, and we will work with that. Even so, there have been changes in choices over the years, and these will likely continue. It seems to be a fact of life that not all graphic options work the way you think that should. Some things just drive me crazy. But you have nothing to lose by experimenting, and I encourage you to do so.)

SPSS has the capability to create many types of charts as can be seen by clicking on the **Graphs** menu bar. Graphs can also be created by using options available in some dialog boxes for analyses. For example, histograms can be created from the Graphs menu or from **Analyze/Descriptive Statistics/Frequencies**. In this chapter, we will focus on the **Graphs/Legacy** menu. Additional options will be discussed throughout the book. One thing to focus on throughout this chapter is the various options for editing graphs.

We'll begin with frequency distributions. Follow the steps as you read along.

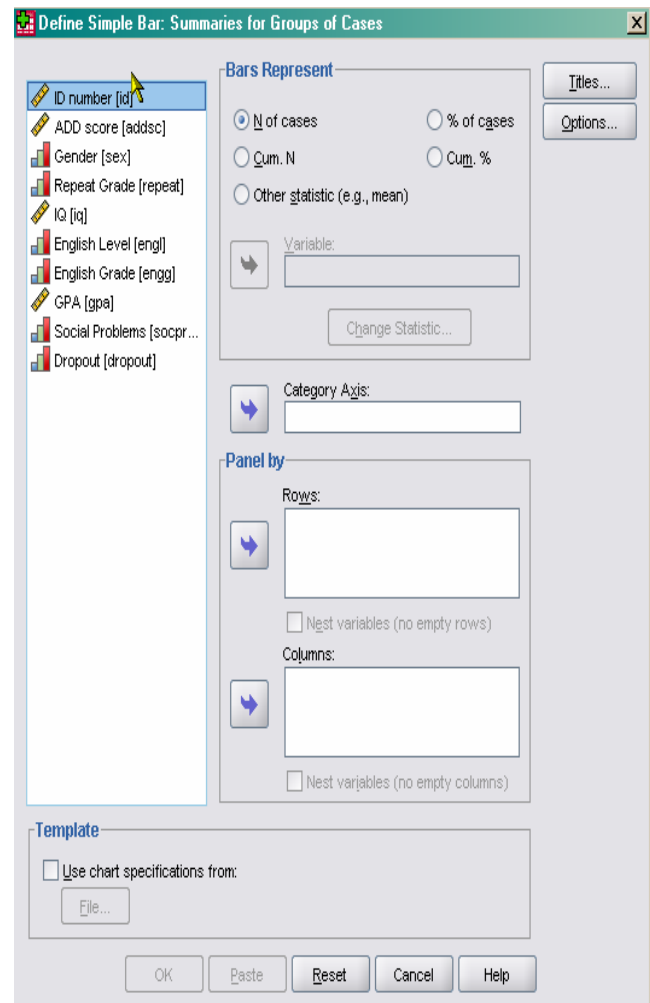


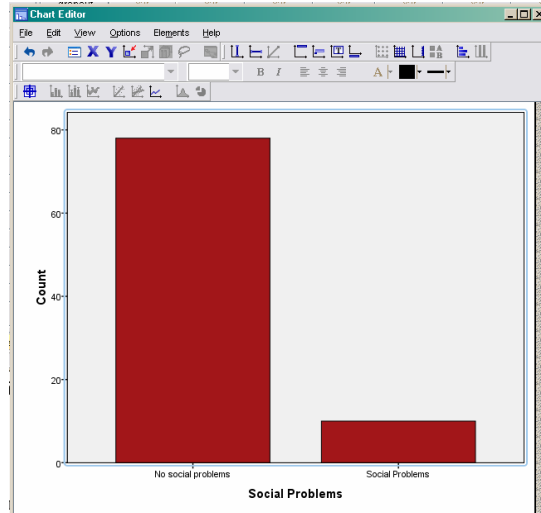
Frequency Distributions

Frequency distributions plot the number of occurrences or counts for each value of a variable. Let's start by looking at the frequency of social problems from the data in Appendix D. Your variable and value labels may be slightly different than mine depending on how you completed set up your data when you read them in. This may make your chart labels different than those displayed in this chapter.

- ✓ **Open** *appendixd.sav*.

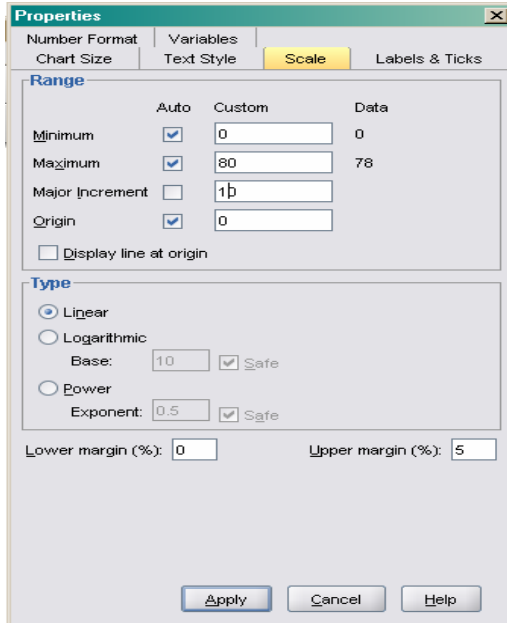
- ✓ Under **Graphs**, select **Bar**.
- ✓ At the next dialog box, select **Simple** and **Summaries for groups of cases**. Then click **Define**.
- ✓ In the next dialog box, select the social problems variable, and put it into **Category Axis** by using the arrow. Make sure **N of cases** (meaning number of cases) is selected for **Bars Represent**, then click **Ok**. The graph will appear in the Output Viewer window.
- ✓ To edit the graph, double click on it and the Chart Editor view will open. The Chart Editor has several Menu bars and Icons. Take a moment to review them by clicking on the menus and reading the commands that appear and putting your cursor over the icons and reading the descriptions that appear to the bottom left hand corner.





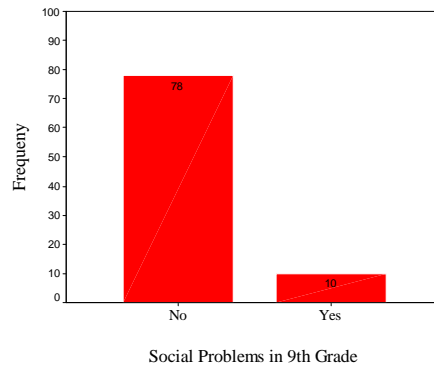
My chart displayed the value labels (“no social problems,” and “yes, social problems”) rather than the numbers 0 and 1 because we specified this earlier using Edit/Options. To edit virtually anything on a chart, you can double click it and relevant dialog boxes will appear. Let’s try it.

- ✓ Double click the labels under the bars, and a dialog box will appear.
- ✓ To edit the title, click on the axis title box (social problems”) and begin typing. I like my labels to be capitalized, so I’ll change it to “Social Problems in 9th Grade.” My graph came out a sickly yellowish brown, and we can’t have that. So double click on the bar, select Fill and Border from the menu, and pick a nicer color.
- ✓ You can specify that SPSS show all value labels or intermittent labels. As long as your dialog box with the graph says “Chart Editor,” you can simply click on the label you want to change and then type a new one. I changed to “No” and “Yes.”
- ✓ I’d also like to change the Y axis to read Frequency rather than Count and to change the scale. First I will click on “Count” and replace that with “Frequency.” Then I will double click on “Frequency,” which will open a dialog box , in which I will select **Scale**. To do this, double click on “Count” to activate the dialog box.



- ✓ Then, in the box for **Major Divisions**, change it to 10 to alter the scale. Click **Apply** and then **Close**.

There are several other parameters that you can change, but I'll leave it to you to play around and see what you can accomplish. My final graph looks like

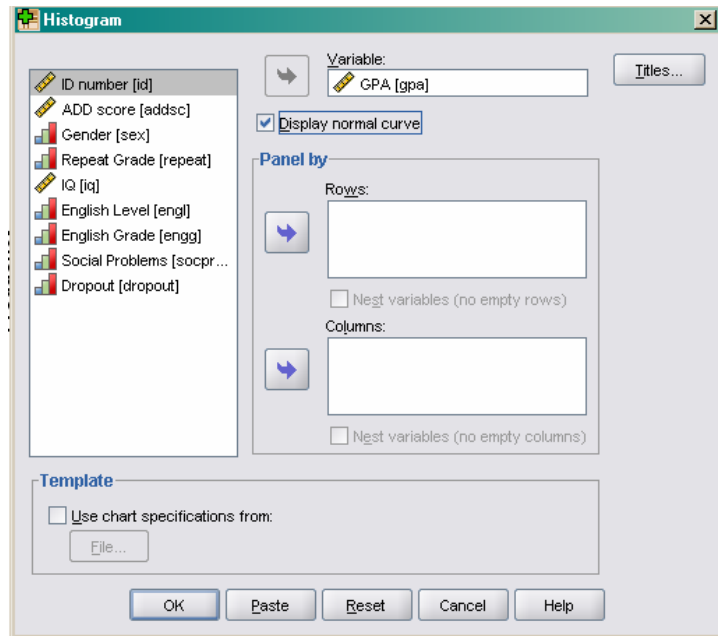


Histograms

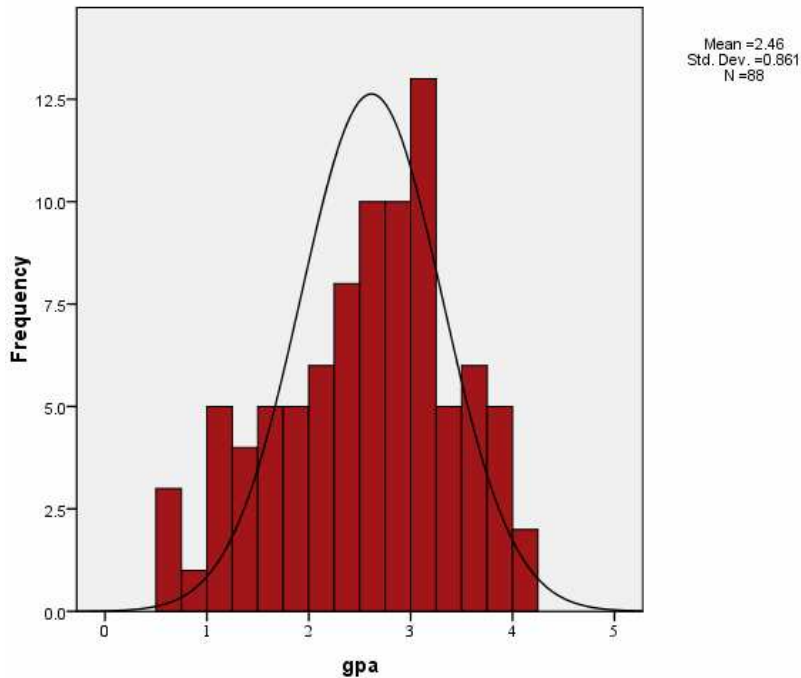
As noted in the text, histograms are essentially a frequency distribution for ranges of values rather than individual values. These are ideal when there are several values and relatively low frequencies for each. A histogram is a perfect way to display the GPA data from Appendix D. Let's try it together.

- ✓ Select **Histogram** under **Graphs/Legacy**.

- ✓ Select grade point average for **Variable**. Check **Display the normal curve** if you want to see if the distributions is approximately normal. Then, click **Ok**.



- ✓ SPSS will select value ranges for you. If you want to change the ranges, double click on the graph to move to Chart Editor, and then double click on the values along the X axis and select **Scale** from the dialog box. I like the way they are.
- ✓ On the dialog box that appears, change the scale and anything else that you think needs changing and then click close. For example, I changed the font to a serif font, the color to red, and the values on the X axis to have no decimal places. The result follows.



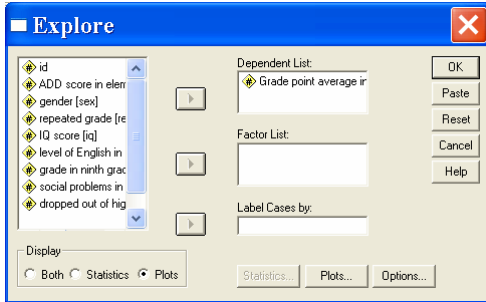
Notice that by default SPSS prints the mean, standard deviation, and sample size next to the graph. If you don't want that, you can click on the output and delete it.

Whenever you are done editing a graphic, you will need to close a dialog box, if one is open, and then close the Chart Editor. That will take you back to your original printout with all of your changes.

Stem-and-Leaf Plot

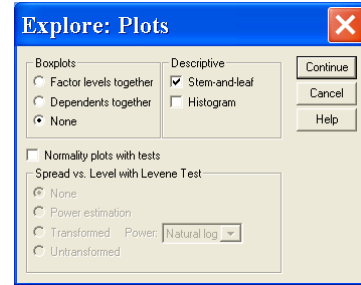
As noted in the text, a stem-and-leaf plot is another type of histogram. This would be another appropriate way to display the GPA data. To create a stem and leaf plot, we will not use the Graphs menu. Rather, the stem-and-leaf plot is an option in the Analyze menu. Let's try one together.

- ✓ Select **Analyze/Descriptive Statistics/Explore**.



- ✓ Put Grade point average in the **Dependent list**. Select **Plots** under display since we are not calculating statistics yet, then click **Plots**.

- ✓ In the next dialog box, select **None** for **Boxplots**. Select **Stem-and-Leaf** under **Descriptive**. Note that you could create boxplots and histograms from here, although we won't right now. Then, click **Continue** and then **Ok**. The following stem-and-leaf plot will appear.



gpa Stem-and-Leaf Plot

Frequency	Stem & Leaf
4.00	0 . 6667
9.00	1 . 000002333
10.00	1 . 5556677777
14.00	2 . 0000002222234
20.00	2 . 5555556666777777777
18.00	3 . 00000000000000022234
11.00	3 . 55555577777
2.00	4 . 00

Stem width: 1.00
Each leaf: 1 case(s)

- ✓ Notice the style of this chart is not as elegant as the others. To edit it, double click on it. This does not activate the Chart Editor since the stem-and-leaf plot is essentially text. It does activate the Output Editor though, which allows you to change font, alignment, color, etc. Edit the stem-and-leaf plot to suit your style.

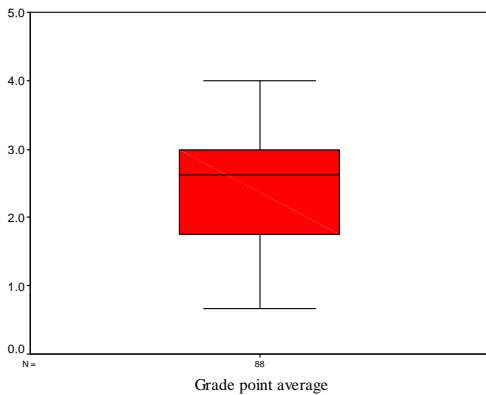
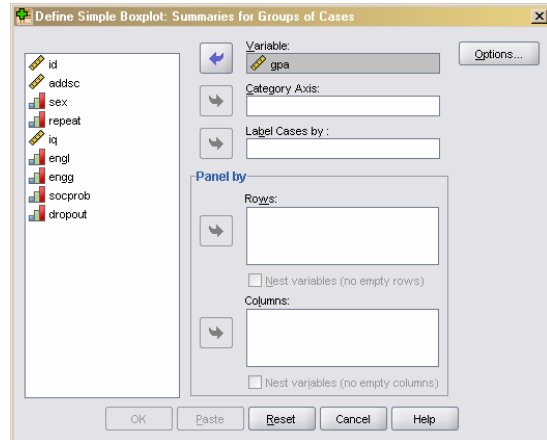
Boxplots

Boxplots are useful to illustrate the dispersion of data. Let's create a boxplot together using the same data. We'll begin with a simple example. We could create a boxplot just as we did the stem-and-leaf display, but instead we will go back to the **Graph** menu.

✓ Select **Boxplot** under **Graphs/Legacy**.

✓ Select **Simple** for style, click **Summaries for Separate Variables**, and then click **Define**.

✓ Select grade point average for **Boxes Represent** and then click **Ok**. The resulting boxplot is below. Notice that the dialog box would also let you add a category axis to break down the data by another variable, perhaps Gender, and then Panels to further divide the results, perhaps by SocProb. You can try these options.

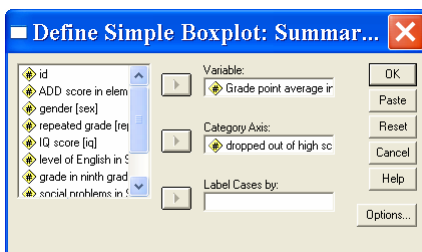


✓ I'd like to change the scale on the Y axis to .5 units to make the median value more clear. Can you figure out how to do this on your own? [Double click the Y axis units, then change the **Major Increment** as you did above.]

Now, let's try a more complicated boxplot. I'd like to see the boxplot of GPA in ninth grade for those who did and did not ultimately drop out of high school.

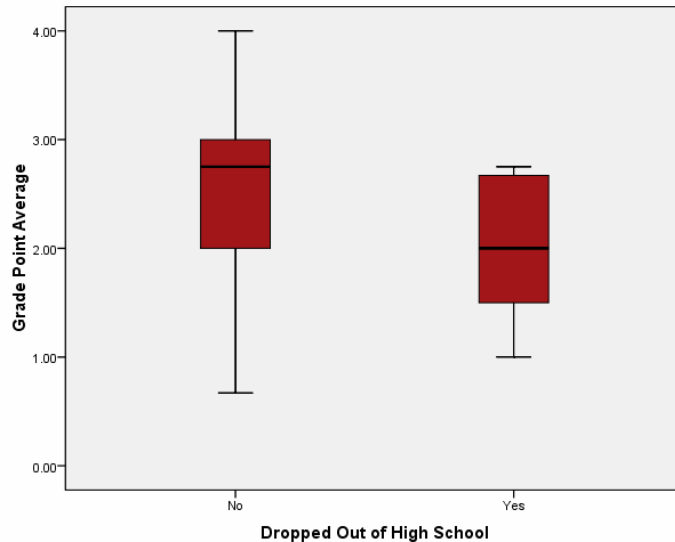
✓ Click on **Boxplot** in the **Graphs** menu.

✓ Select **Simple** and **Summaries for Groups of cases**, then click **Define**.



✓ Select grade point average as the **Variable** and Dropout as the **Category Axis**. Then, click **Ok**. The resulting boxplot is below. Notice the difference between the two groups.

- ✓ Using what you learned in the previous examples, edit the boxplot to suit your style (e.g., edit the labels, change the color, change the scale, etc).



You can also create box plots for more than one variable. You would typically only do so if the variables are on a similar scale, which is not the case in the example we are using. This would be appropriate if we had GPA data from multiple points in time, for example. To create a box plot of this type, you would select **Simple** and **Summaries** of separate variables.

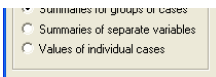
Bar Graphs – Again

Another way to visually compare the data from different groups is a bar chart. Let's create one from the same example as above so we can compare them.

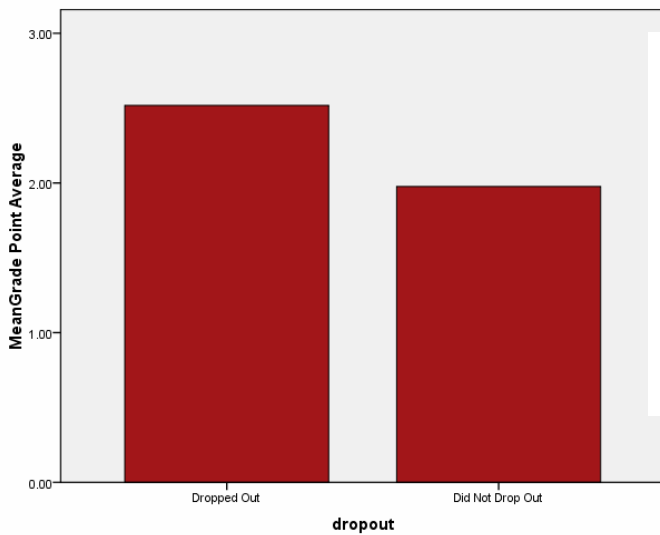
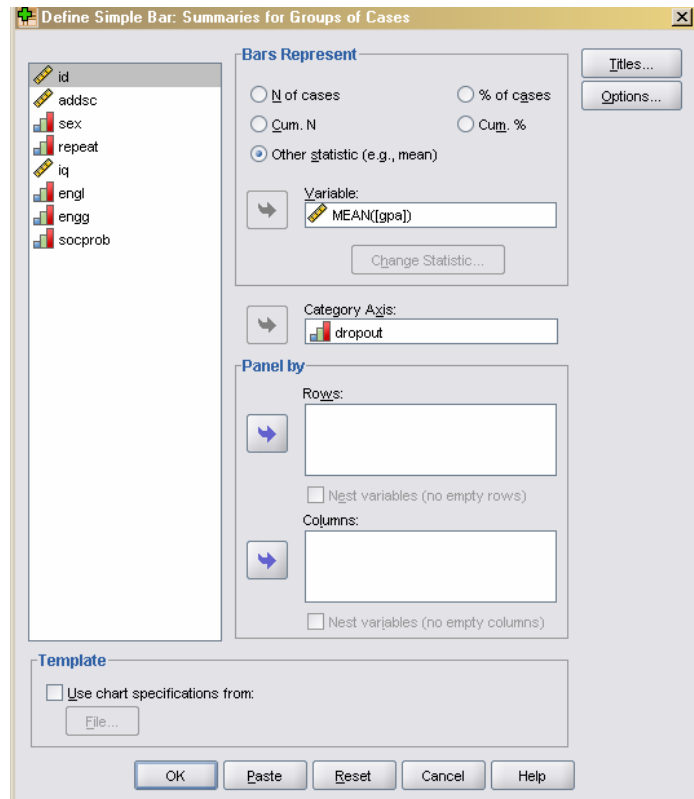
- ✓ Select **Bar** under **Graphs/Legacy**.



- ✓ Select **Simple** and **Summaries for groups of cases**, then click **Define**.



- ✓ Under **Bars Represent**, select **Other statistic**. Add GPA to the variable box and then click **Change Statistic**. Select **Mean of values**, but take a moment to review the other options, then click **Continue**. Put dropped



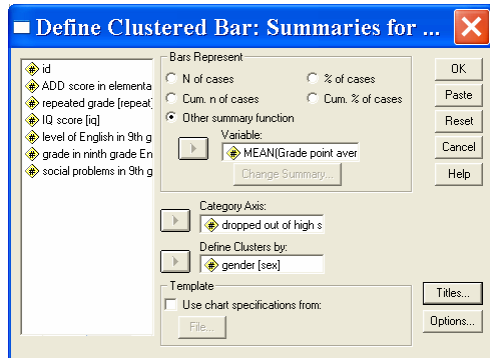
- ✓ Use what you learned above to edit the graph (e.g., change the text style, add a title, etc).

Notice the biggest difference between this and the boxplot on the previous page is that the boxplot gives you a sense of variability and central tendency. A bar graph of means does not illustrate variability.

Now, let's create a more complicated bar chart in which there are two independent variables. This is an ideal way to display main effects and interaction effects for factorial designs, which are discussed in Chapter 17. This time, we'll graph the mean

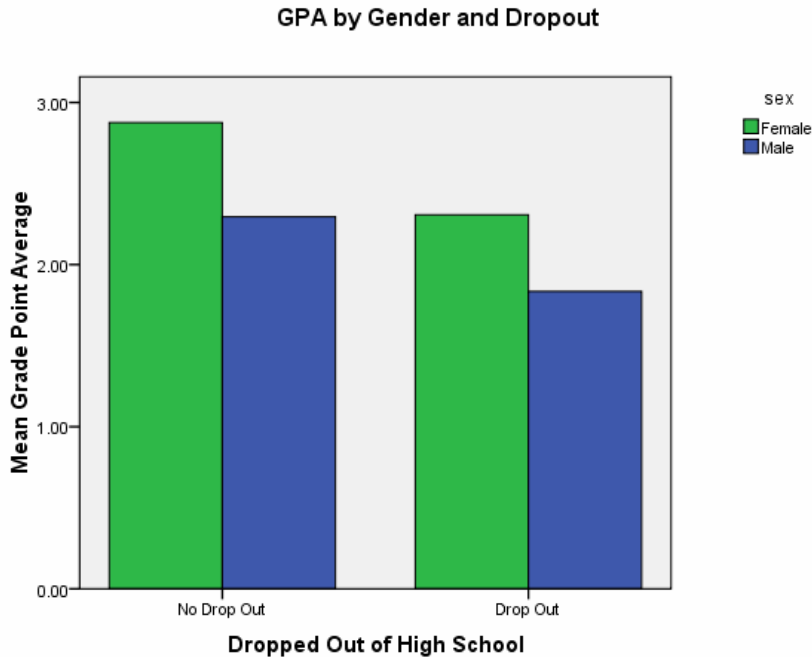
grade point average based on both gender and whether or not the student ultimately dropped out.

- ✓ Select **Bar** from the **Graphs/Legacy** menu.
- ✓ Select **Clustered** and **Summaries for groups of cases**, then click **Define**.



- ✓ Select **Bars Represent Other statistic**, and select GPA. Mean will be the default summary. Select dropped out of high school for the **Category Axis** and gender for the **Cluster**. Select **Titles** from this dialog box and type “GPA by Gender and Dropout Status.” Then click **Continue** and **Ok**.

- ✓ Double click on the graph to activate the Chart Editor. When there are multiple independent variables, you have to decide which one should be displayed on the x axis and which one should be used as the categories. Fortunately, you can shift them back and forth and decide which view is better. To do so, double click on the Chart Editor, select **Categories** from the dialog box, highlight 2(2) and click the little up arrow to the right. This will reorder 1(1) and 2(2) to 2(2) and 1(1). When you click apply this will reverse the variables in the graph.

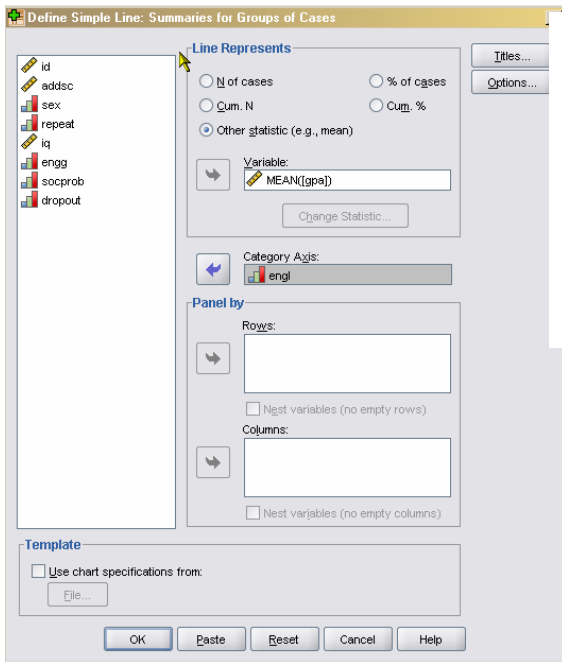


- ✓ Now, edit the labels and axes as you learned above so the graph makes sense and looks polished.
- ✓ It's very important that the gender groups are visually distinct from one another. If I were printing to a black and white printer, I would apply patterns to the bars to make them stand out from each other. To do so, click on one of the bars. You will see each of the related bars selected. Then click the **Fill Pattern** icon and select the fill of your choice and click **Apply**.

Line Graphs

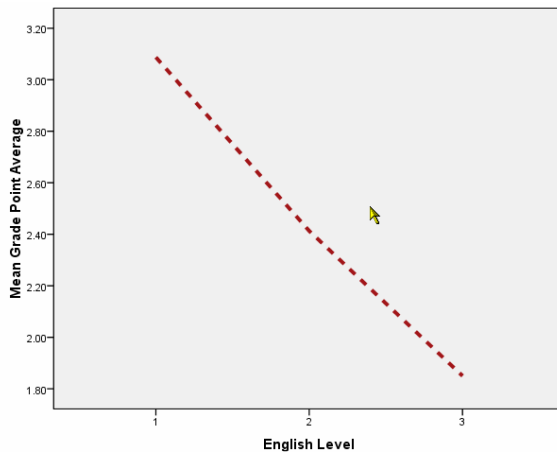
Line graphs can be used also to display mean differences. Line graphs are most commonly used to display mean differences over time or conditions. They can also be used to display mean differences between groups. Let's make a simple line graph with one independent variable. Most journals in Psychology would prefer that people used bar graphs for simple graphs unless the X axis is a continuous one such as time. I like line graphs, and since I am getting old, I can ignore the APA.

- ✓ Select Line from the Graphs menu.
- ✓ Then select **Simple** and **Summaries for groups of cases**. Click **Define**.



- ✓ Select **Other summary function** under **Line Represents**, then select GPA as the **Variable**. Mean is the default option. Select English Level for the **Category Axis**, because that has more than two levels. Then, click **Ok**.

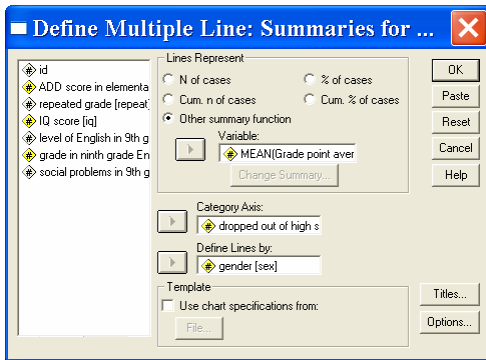
- ✓ Double click on the graph to activate the Chart Editor. Then double click on the line and select a weight and style from the dialog box that comes up. Change the style and weight of the line to suit your style and click **Apply**. My line graph follows.



- ✓ Edit your graph (e.g., add a title, change the labels and scale, etc) as illustrated previously.

As with the bar graph, a line graph can also be used to display interactive effects. Let's create the same graph we did above illustrating mean GPA based on both gender and drop out status.

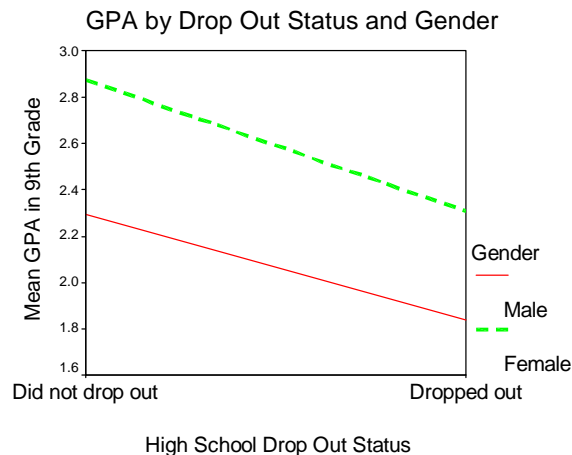
- ✓ Select **Line** from the **Graphs/Legacy** menu.
- ✓ Select **Multiple** and **Summaries for groups of cases** and click **Define**.



- ✓ Much like before, select **Other statistic**, Mean GPA for **Lines Represent**. Specify dropped out as the **Category Axis** and Gender for the **Lines**. Click **Titles**, name your graph, and click **Continue**. Then, click **Ok**.

- ✓ Double click the graph to activate the Chart Editor. As before, you can change colors, fonts, etc. I will leave mine pretty much as it is.
- ✓ Now, I need the lines that represent gender to be distinct from one another. Click on a line, then click on the **Line Style** Icon. I am going to increase the weight of the line and make it dashed. Click **Apply**. I'd like the other line to be thicker as well. Click the other line, then the **Line Style** icon, select the heavier weight and click **Apply**.
- ✓ Edit the labels and axes as desired.

My final graph appears on below. Take a moment to interpret the graph and compare it to the bar graph we made of the same data. Which do you prefer? One thing that's nice about the line graph is that you can see the lines representing gender are parallel. In the language of factorial designs discussed in Chapter 17, I would guess there are main effects of gender and drop out status, but no interaction effect. Of course, we would need to calculate a Factorial ANOVA to be certain.

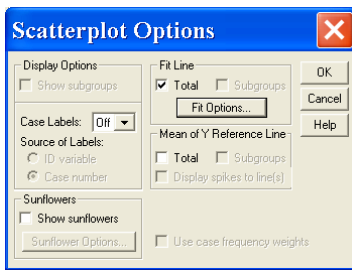
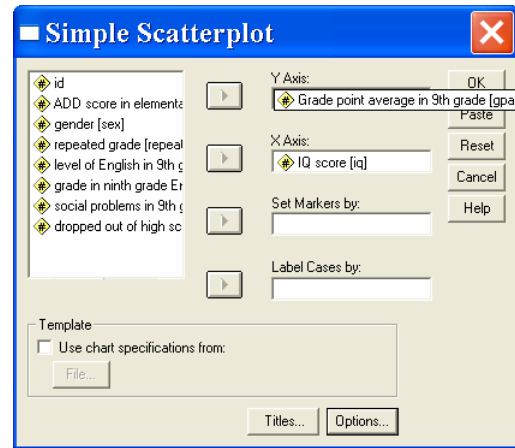


Scatterplots

Scatterplots are typically used to visualize relationships between continuous variables. In Appendix D, it makes sense to expect a positive association between IQ and GPA. Let's create this scatterplot together.

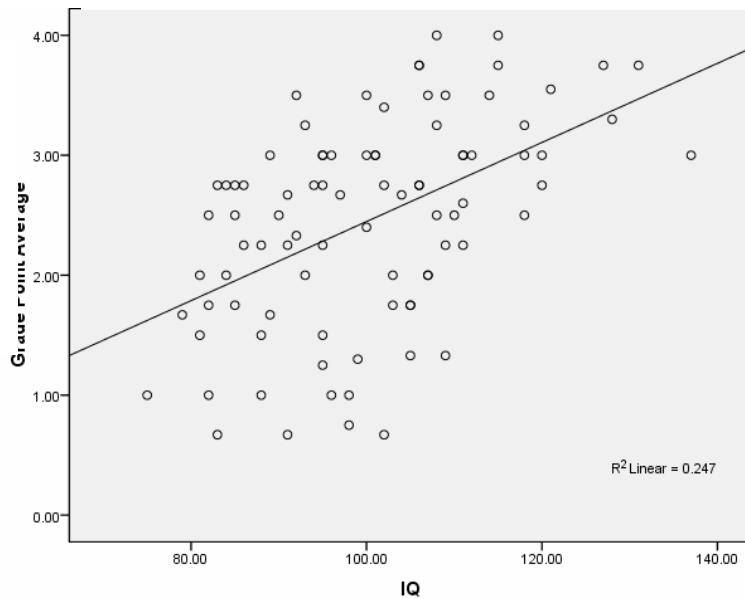
- ✓ Select **Scatter/Dot** under **Graphs/Legacy**.
- ✓ Select **Simple** and click **Define**.

- ✓ Typically, the dependent variable is placed on the Y axis and the independent on the X axis. Although we cannot prove causality in this instance, theoretically it makes more sense that IQ contributes to GPA rather than the reverse. So, select grade point average for the **Y Axis** and IQ score for the **X Axis**, then click **Ok**. The scatterplot will appear.



- ✓ Next, we can add a line of best fit to better illustrate the relationship. Double click on the chart to activate the Chart Editor. Then, select **Fit line at total** from the **Elements** menu. Then click **Fit Line** and select **Linear** under **Fit Method**. Note the other options, then click **Apply** if the line has not been drawn. Finally click **Ok**. The scatterplot follows.

- ✓ Edit the graph to suit your style (e.g., change the font, colors, labels etc).



Based on the graph, it appears that GPA and IQ are positively correlated. Of course, we would calculate the correlation coefficient to confirm this statistically.

Pie Charts

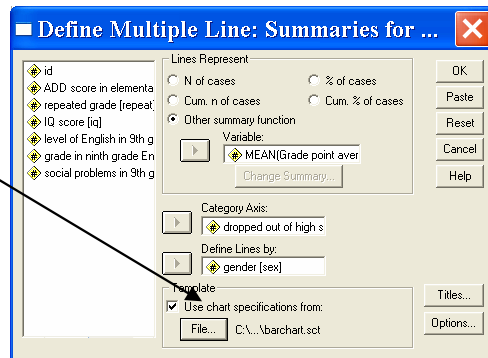
Although not discussed in the text (because they are lowly, useless, misleading, and despicable things), pie charts can be a simple way to describe categorical or nominal variables by displaying the frequency or percent of cases that fit each category. I have decided that I am not even going to give them space here (Sorry, Esther), but you can figure them out if you need to.

Chart Templates

If you do a lot of graphing and develop a style you particularly like, you can save it as a template for future graphs. This can save a lot of time editing.

- ✓ When you have created a graph you really like, in Chart Editor select **File/Save chart template** and save the file with a name that makes sense to you (e.g., interaction bar graph, simple bar graph, etc). The file extension for chart templates is .sct.

- ✓ The next time you are creating a similar graph, select **Use chart specification from**, then click **File** in the Define dialog box. Select the appropriate template, **click Open**, continue defining your graph and click **Ok**. The templates are stored, at least on my machine, in `c:/Program Files/SPSSInc/Statistics17/Looks`.



- ✓ **Exit SPSS**. There is no need to save the Data file since we have not changed it. It is up to you to decide if you would like to save the output file for future reference. If you save it, make sure you give it a name that makes sense!

As you can see, SPSS is capable of creating many different chart types, and each type has many options. In addition, charts can be created from the Graphs menu and from the Analyze menu. Further, charts can be edited tremendously after they are created using the Chart Editor. I encourage you to complete the following exercises and try out several different options along the way—even those that were not covered in this chapter. This will increase your comfort with SPSS and clarify your own graphing preferences.

Exercises

Each of the following exercises is based on the data in *appendixd.sav*.

1. Create a histogram for ADDSC. [Note, this is exercise ***** in the text book.]
2. Create a boxplot for ADDSC. Then, create ADDSC boxplots for those with and without social problems. How would you describe the data?
3. Create a scatter plot illustrating the relationship between ADD symptoms and GPA. Include the line of best fit. Does there appear to be a relationship between these variables?
4. Create a pie chart illustrating the percent of individuals in the various types of English classes.
5. Create a bar chart to illustrate the mean GPA of students in each of the 3 types of English Classes. Do they appear to be similar or different?

Create a bar chart and a line graph illustrating mean differences in GPA based on both gender and level of English class. Do you think there are main effects of gender and type of class? Do you think there is an interaction effect? Which type of graph do you prefer and why?