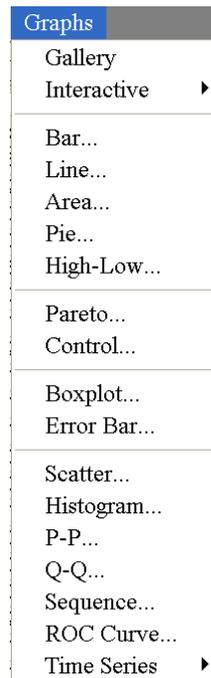# 3: Graphing Data

Objectives

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

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 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 your read along.

Graphs
Gallery
Interactive ▸
Bar...
Line...
Area...
Pie...
High-Low...
Pareto...
Control...
Boxplot...
Error Bar...
Scatter...
Histogram...
P-P...
Q-Q...
Sequence...
ROC Curve...
Time Series ▸

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 Exercise 2.3, which may make your chart labels different from 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.
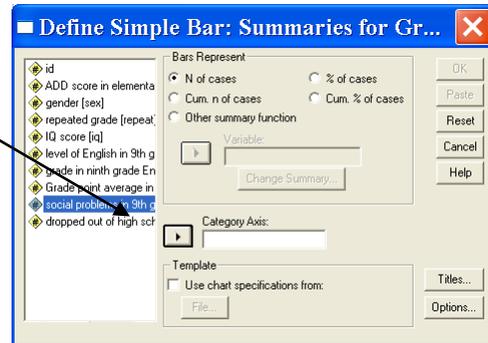
✔ To edit the graph, double click on it and the Chart Editor view will open. The Chart Editor has several Menu bars and Icons. Take a moment to review them by clicking on the menus and reading the commands that appear and putting your cursor over the icons and reading the descriptions that appear to the bottom left hand corner.
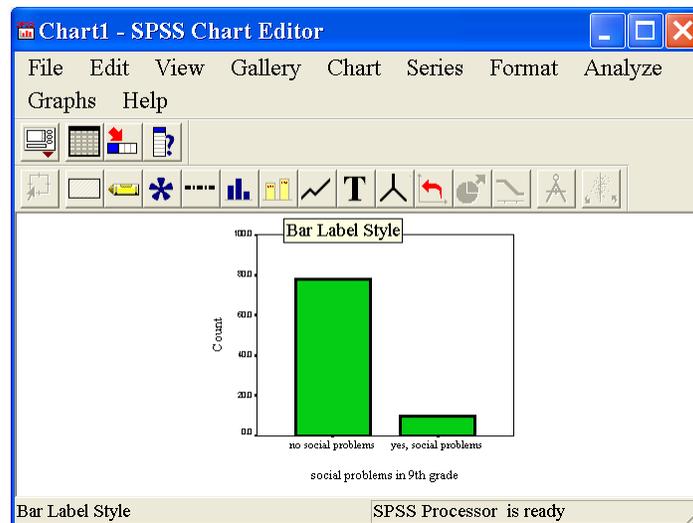
My chart displayed the value labels ("no social problems," and "yes, social problems") rather than the numbers 0 and 1 because we specified this earlier using Edit/Options. To edit virtually anything on a chart, you can double click it and relevant dialog boxes will appear. Let's try it.

✔ Double click the labels under the bars, and a dialog box will appear.

✔ To edit the title, click in the **Axis Title** box and begin typing. I like my labels to be capitalized, so I'll change it to "Social Problems in 9th Grade." I like titles to be centered, so don't change the title justification, but click on the arrow and look at the other options. Then, click **Labels** to change the bar or category labels. Another dialog box will appear.

✔ Notice, you can specify that SPSS show all value labels or intermittent labels.   Select **All labels** since there are only 2 categories. (If you had many categories, showing all labels might make things unreadable.) Then, to edit the Label Text, highlight the label you want to change and then type in the box named Label. I will change the labels to a simple Yes and No since the variable label clarifies what Yes and No mean.   Click change after you edit each one. Also, take a moment to view the options under Orientation. Select whichever one suits you. Then click **Continue**, which returns you to the previous dialog box, where you can click **OK**.

✔ I'd also like to change the Y axis to read Frequency rather than Count and to change the scale. To do this, double click on "Count" to activate the dialog box.
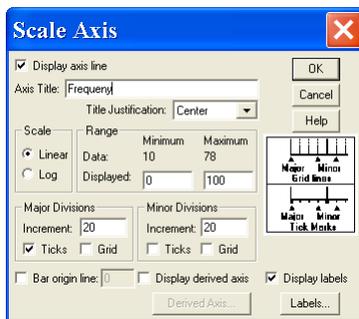
✔ Type Frequency in the **Axis Title** box, and select the type of **Title justification** you want. Then, in the box for **Major Divisions**, change it to 10 to alter the scale. Then click **Labels** and change the number of decimal places to 0. Click **Continue** and then **Ok**.

✔ To label the actual frequency value inside the bars, click on the **Bar Labels Style** Icon and select the style you like. Click **OK**.

✔ To change the color of the bars, single click the mouse when the cursor is over one of the bars, then click the **Color** Icon.

✔ Select **Fill**, then click on the color of your choice. Then, click **Apply**. If you don't like the colors, click **Edit** and create your own. Click **Apply** when done.

✔ To change the font of axis and category labels, click them once to select them, then use the Text icon to select different font styles and sizes as desired.

My final frequency distribution looks like this.

Social Problems in 9th Grade
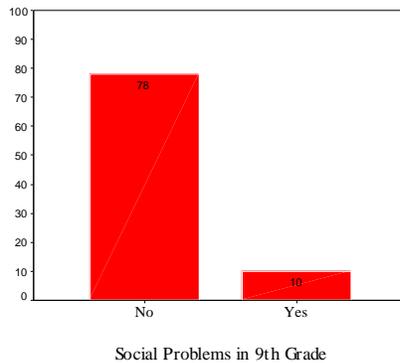
Histograms

As noted in the textbook, 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**.
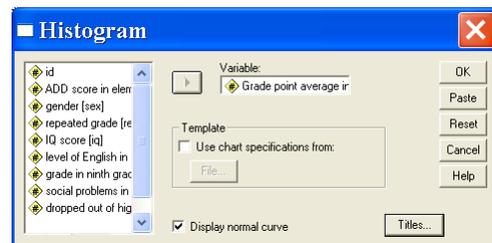
✔ Select grade point average for **Variable**. Check **Display the normal curve** if you want to see if the distributions is approximately normal. Then, click **Ok.**

✔ SPSS will select value ranges for you. If you want to change the ranges, double click on the values along the X axis. I like the way they are.

✔ On the dialog box that appears, select **Custom** under **Intervals** and then click **Define**. In the next box, indicate the number of intervals you would like and click **Continue**. Then, click **Ok**.
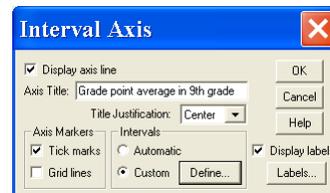
✔ Add a title for the Y axis as you did before. Double click the area, then type Frequency for the **Axis Title**, select the type of **Title justification** you want, and click **Ok**.   The resulting graph will look like this.
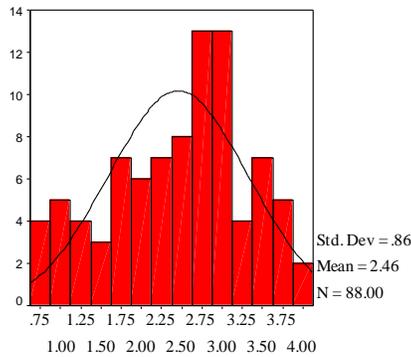
✔ SPSS includes descriptives as a default. If you don't want them on your histogram, double click in that area, unselect **Display Legend** in the dialog box that appears and click **Ok**.

Grade point average in 9th grade

✔ To add a title to your chart, click **Title** under the **Chart Menu**.

✔ Type in the title of your choice. **Title 1** and **Title 2** really mean line 1 and 2 of the title. select your **Title justification**, and click **Ok**. Notice that you could add a subtitle as well.

✔ Review your final histogram. Continue to edit it to fit your style (e.g., change the color, font, axes labels, etc).
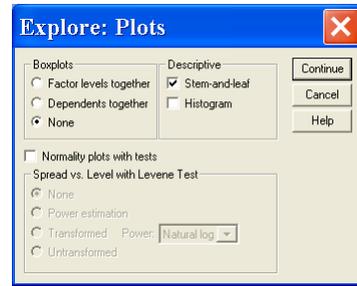
Stem-and-Leaf Plot

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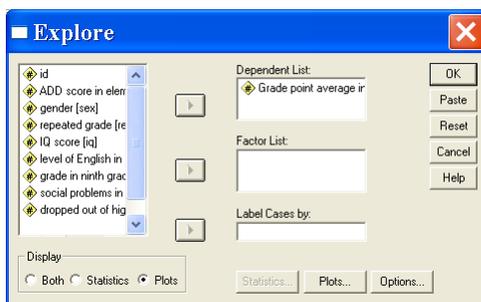
✔ Put Grade point average in the **Dependent list**. Select **Plots** under display since we are not calculating statistics yet, then click **Plots.**

✔ In the next dialog box, select **None** for **Boxplots**. Select **Stem-and-Leaf** under **Descriptives**. 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.

```
Grade point average in 9th grade Stem-and-Leaf Plot
     Frequency     Stem &  Leaf
      4.00          0 .  6667
      9.00          1 .  000002333
     10.00          1 .  5556677777
     14.00          2 .  00000022222234
     20.00          2 .  55555566667777777777
     18.00          3 .  000000000000022234
     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.
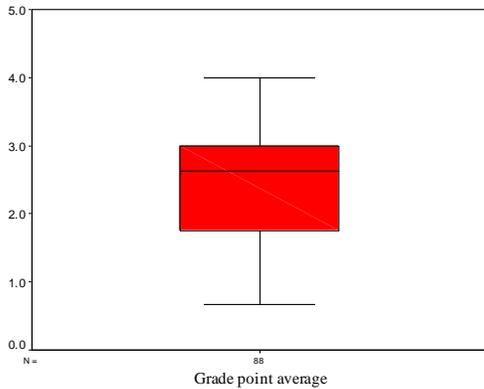
Boxplots

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.

✔ Select **Boxplot** under **Graphs.**

✔ Select **Simple** for style and click **Define**.

✔ Select grade point average for **Boxes Represent** and then click **Ok**. The resulting boxplot is below.
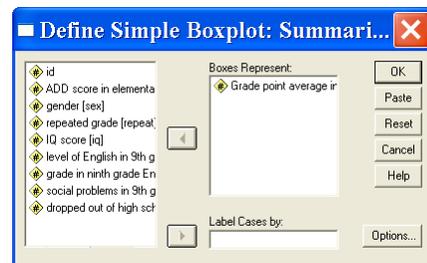
✔ Id like to change the scale on the Y axis to .5 units to make the median value more clear. Can you figure out how to do this on your own? [Double click the area, then change the **Major Divisions** as you did above.]

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**.
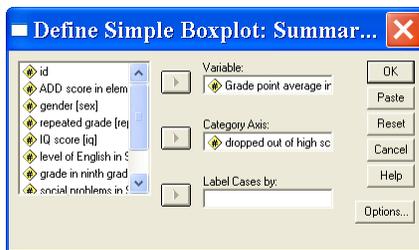
✔ Select grade point average as the **Variable** and dropped out of high school as the **Category Axis**. Then, click **Ok**. The resulting boxplot is below. Notice the difference between the two groups.

✔ 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).
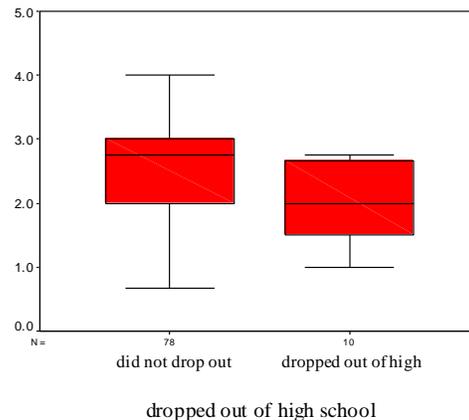

You can also create box plots for more than one variable. You would typically only do so if the variables are on a similar scale, which is not the case in the example we are using. This would be appropriate if we had GPA data from multiple points in time, for example. To create a box plot of this type,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

Bar Graphs

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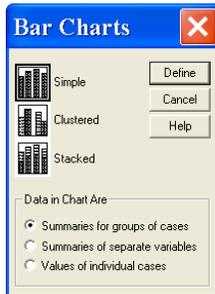
✔ Select **Simple** and **Summaries for groups of cases**, then click **Define**.

✔ Under **Bars Represent**, select **Other summary function**. Click **Change Summary**. Select **Mean of values,** but take a moment to review the other options, then click **Continue**. Put dropped out of high school in the **Category Axis**, then click **Ok**. The resulting graph is below.
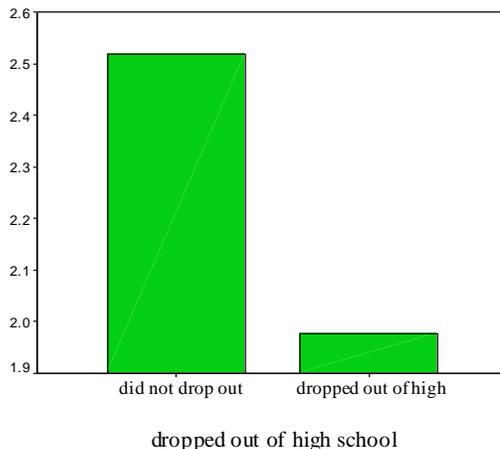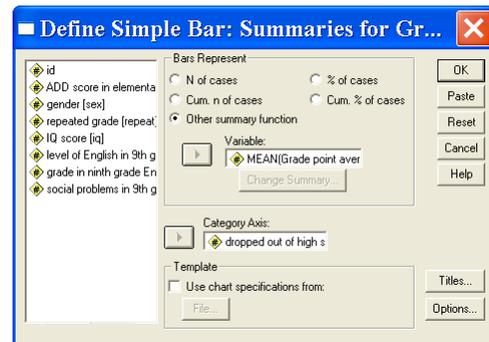
✔ Use what you learned above to edit the graph (e.g., change the text style, add a title, etc).

dropped out of high school

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** menu.

✔ Select **Clustered** and **Summaries for groups of cases,** then click **Define.**
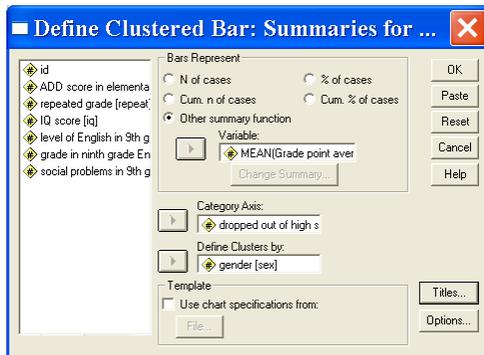
✔ Select **Bars Represent Other summary function**, and select GPA. Mean will be the default summary. Select dropped out of high school for the **Category Axis** and gender for the **Cluster**. Select **Titles** from this dialog box and type "GPA by Gender and Dropout Status." Then click **Ok**.

✔ 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, click **Transpose Data** in the **Series** menu as illustrated below.
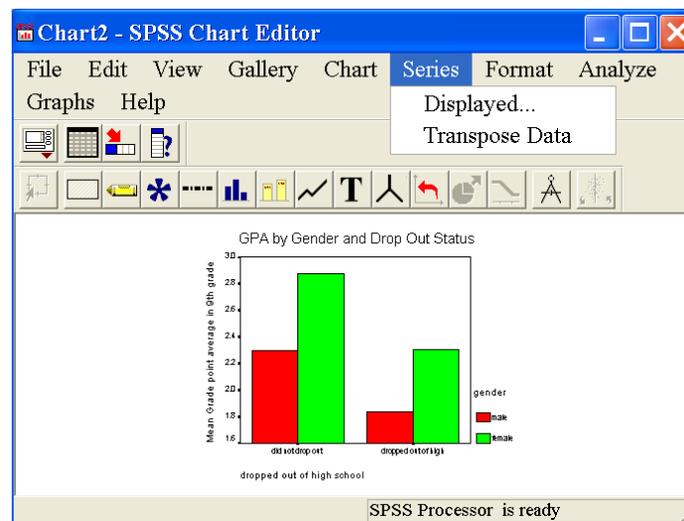
✔ Take a moment to view the new layout illustrated on the following page. Personally, I preferred the first graph with drop out status on the X axis. So, click **Series/Transpose** again to return it to the first layout.

GPA by Gender and Drop Out Status
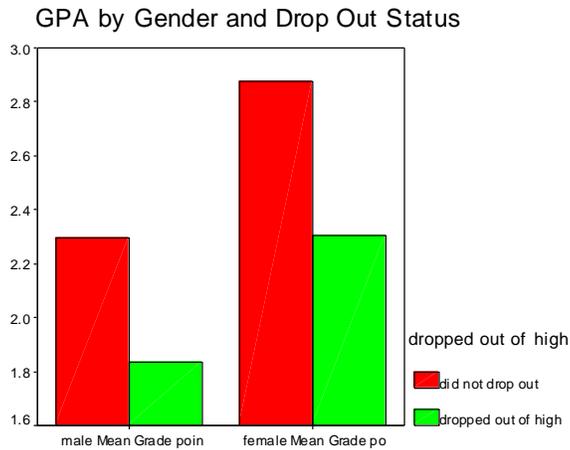
- Now, edit the labels and axes as you learned above so the graph makes sense and looks polished.

- It's very important that the gender groups are visually distinct from one another. Since I am printing to a black and white printer, I will 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**. I'll leave the other bar a solid color. My final graph appears below.

GPA by Gender and Drop Out Status
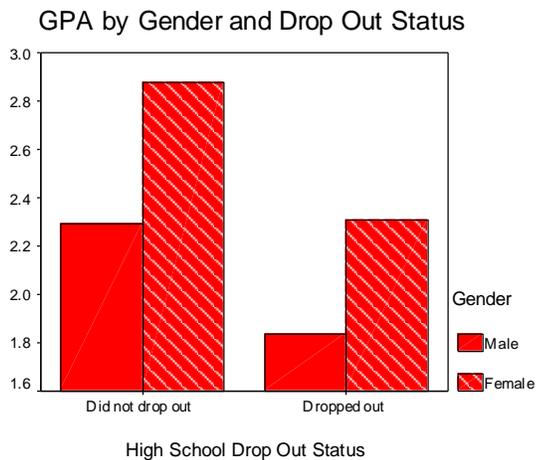
Line Graphs

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.

✔ Select Line from the Graphs menu.

✔ Then select **Simple** and **Summaries for groups of cases**. Click **Define**.
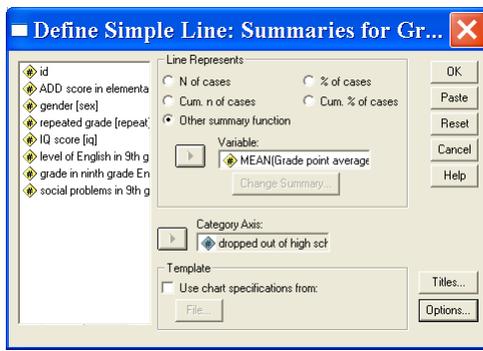
✔ Select **Other summary function** under **Line Represents**, then select GPA as the **Variable**. Mean is the default option. Select dropped out of school for the **Category Axis**. Then, click **Ok**.

✔ Double click on the graph to activate the Chart Editor. Then click on the **Line Style** icon. Change the style and weight of the line to suit your style and click **Apply**. My line graph follows.


dropped out of high school
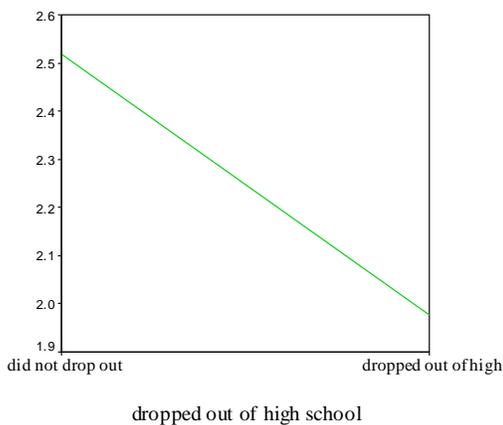
✔ Edit your graph (e.g., add a title, change the labels and scale, etc) as illustrated previously.

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** menu.

✔ Select **Multiple** and **Summaries for groups of cases** and click **Define**.
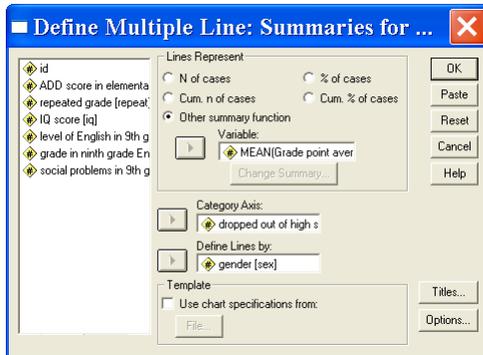
✔ Much like before, select **Other summary function**, Mean GPA for **Lines Represent**. Specify dropped out as the **Category Axis** and Gender for the **Lines**. Click **Titles**, name your graph, and click **Continue**. Then, click **Ok**.

✔ Double click the graph to activate the Chart Editor. As before, you can **Transpose the data** through the **Series** menu if you want, but I like the present layout.

✔ 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 the following page. 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.

GPA by Drop Out Status and Gender

Gender

Male

Female

High School Drop Out Status
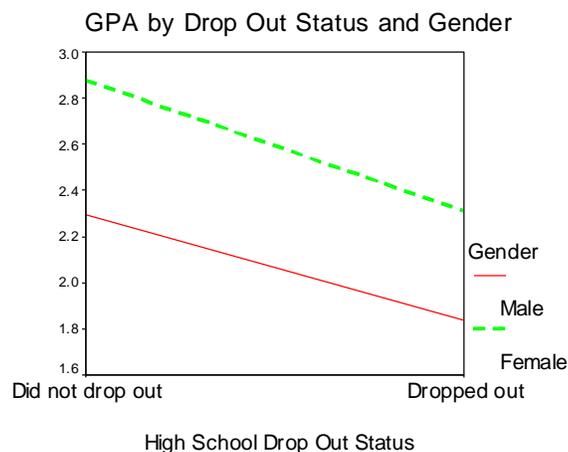
Scatterplots

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** under **Graphs**.

✔ 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.
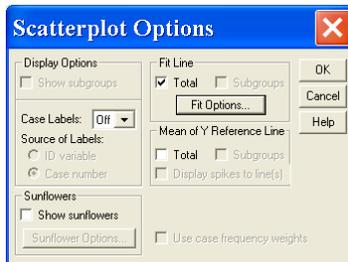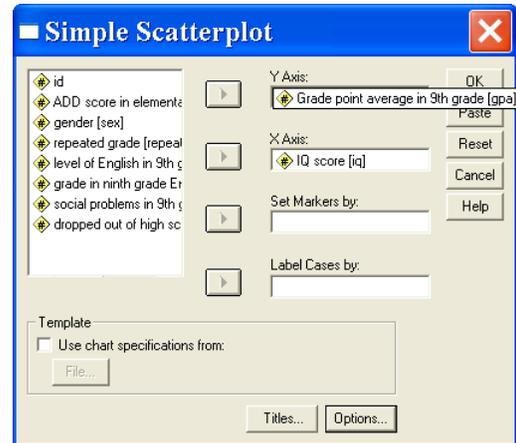
✔ Next, we can add a line of best fit to better illustrate the relationship. Double click on the chart to activate the Chart Editor. Then, select **Chart/Options**. Select **Total** under Fit Line. Then click **Fit Options** and select **Linear Regression**. Note the other options, then click **Continue**. Finally click **Ok**. The scatterplot follows.

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

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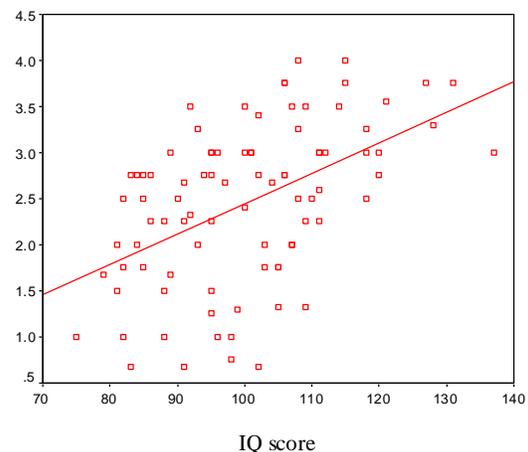
Based on the graph, it appears that GPA and IQ are positively correlated. Of course, we would calculate the correlation coefficient to confirm this statistically.

Pie Charts

Although not discussed in the text (because they are lowly, useless, misleading, and despicable things), pie charts can be an silly way to describe categorical or nominal variables by displaying the frequency or percent of cases that fit each category. Pie Charts are attractive because they are easily understood by most people. Let's try one to illustrate the gender composition of the sample in Appendix D. (Please, God, don't let her try to show them how to make 3-dimensional ones!!)

✔ Select **Pie** in the **Graphs** menu.

✔ Select **Summaries for groups of cases**, then click **Define**.
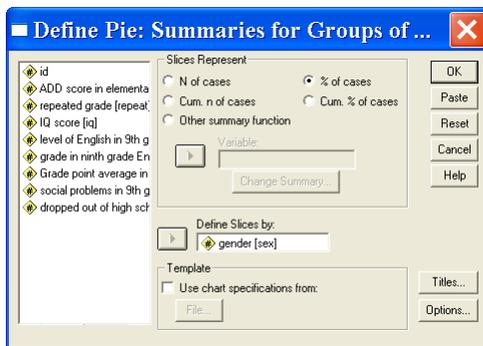
✔ For **Slices Represent** select **% of cases**. For **Define Slices by**, select gender.   Then click **Ok**.

✔ The pie chart will appear. Double click it to activate the Chart Editor.

✔ First, let's add labels that show the percent of males and females. To do this, select **Chart/Options**.
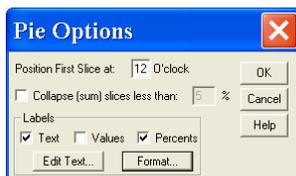
✔ Under labels, keep **Text** selected if you want it to say Male and Female. Click **Edit text** if you want to change the labels. Select **Percents**, then click **Format**. Then view the various positions and select the one of your choice. I will select **Outside**. Then click **Continue** and **Ok**.

✔ Edit the colors of the pie slices. To do this, single click on the male area of the pie, then click the **Fill Pattern** icon. Select the pattern of your choice and click **Apply**. Follow the same steps for the female portion of the pie.
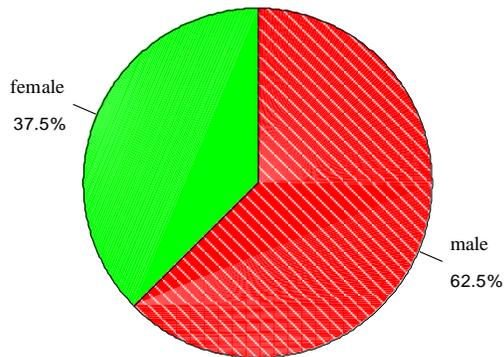
✔ Add a title to your chart using **Chart/Options/Title** as above.

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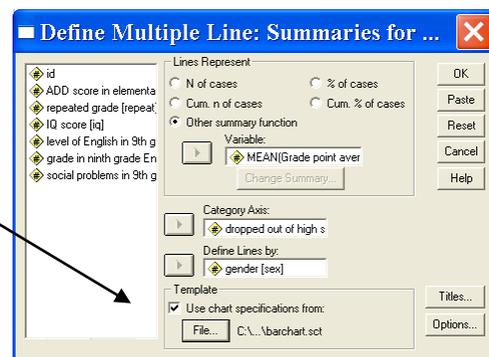
✔ **Exit** SPSS. There is no need to save the Data file since we have not changed it. It is up to you to decide if you would like to save the output file for future reference. If you save it, make sure you give it a name that makes sense!

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?