Suppose you were hired by a telephone-polling firm to interview a large number of respondents. Your job is to find out and record three characteristics of each person you interview: age, educational attainment, and ideological leanings. The natural human tendency would be to record these attributes in words. For example, you might describe a respondent this way: “The respondent is 22 years old, has a college degree, and is ideologically moderate.” This would be a good thumbnail description, easily interpreted by another person. To SPSS, though, it would make no sense at all. Whereas people excel at recognizing and manipulating words, SPSS excels at recognizing and manipulating numbers. This is why researchers devise a coding system, a set of numeric identifiers for the different values of a variable. For one of the above variables, age, a coding scheme would be straightforward: Simply record the respondent’s age in number of years, 22. In recording information about education and ideology, however, a different set of rules is needed. For example, the General Social Survey (GSS) applies these codes for education (named educ_4 in the dataset) and ideology (polviews):

<table>
<thead>
<tr>
<th>Variable (GSS2012 variable name)</th>
<th>Response</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education (educ_4)</td>
<td>Less than high school</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>High school</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Some college</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>College degree or higher</td>
<td>4</td>
</tr>
<tr>
<td>Ideological views (polviews)</td>
<td>Extremely liberal</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Liberal</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Slightly liberal</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Slightly conservative</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Conservative</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Extremely conservative</td>
<td>7</td>
</tr>
</tbody>
</table>
Thus, the narrative profile “the respondent is 22 years old, has a college degree, and is moderate” becomes “22 4 4” to SPSS. SPSS doesn’t really care what the numbers stand for. As long as SPSS has numeric data, it will crunch the numbers—telling you the mean age of all respondents or the modal level of educational attainment. It is important, therefore, to provide SPSS with labels for each code so that the software’s analytic work makes sense to the user. Accordingly, the SPSS Data Editor has two “views.” The Data View shows the codes that SPSS recognizes and analyzes. The Variable View, among other useful features, shows the word labels that the researcher has assigned to the numeric codes.

**THE DATA EDITOR**

Open the 2012 General Social Survey, GSS2012, and see how this works. (If you are using Student Version, open GSS2012_Student_A. You will use GSS2012_Student_A through Chapter 5.) Locate GSS2012 in the folder where you saved it. Double-click the file to open it. SPSS opens the data file and displays the Data Editor (Figure 1-1). Notice the two tabs at the bottom of the window: Data View and Variable View. Turn your attention to the Data View. (Make sure the Data View tab is clicked.) This shows how all the cases are organized for analysis. Information for each case occupies a separate row. The variables, given brief yet descriptive names, appear along the columns of the editor. You can tell that the first respondent in the dataset is 22 years old. You can also see that this respondent has a college degree (coded 4 on the variable named educ_4) and is a moderate (coded 4 on polviews). To paint a more complete word-portrait of this respondent, however, you need to see how all the variables are coded. To reveal this information, click the Variable View tab (Figure 1-2). This view shows complete information on the meaning and measurement of each variable in the dataset. (You can adjust the width of a column by clicking, holding, and dragging the column border.)

The most frequently used variable information is contained in Name, Label, Values, and Missing. Name is the brief descriptor recognized by SPSS when it does analysis. Names can be up to 64 characters in length, although they need to begin with a letter (not a number). Plus, names must not contain any special characters, such as dashes or commas, although underscores are okay. You are encouraged to make good use of Label, a long descriptor.
(up to 256 characters are allowed), for each variable name. For example, when SPSS analyzes the variable mobile16, it will look in the Variable View for a label. If it finds one, then it will label the results of its analysis by using Label instead of Name. So mobile16 shows up as “Geographic Mobility Since Age 16”—a bit more descriptive than “mobile16.” Just as Label permits a wordier description for Name, Values attaches word labels to the numeric value codes. To find out the value labels for mobile16, click the mouse anywhere in the Values cell and then click the gray button that appears. A Value Labels window opens, revealing the labels that SPSS will attach to the numeric codes of mobile16 (Figure 1-3). Unless you instruct it to do otherwise, SPSS will apply these labels to its analysis of mobile16. (Click the Cancel button in the Value Labels window to return to the Variable View.)

Finally, a word about Missing. Sometimes a dataset does not have complete information for some variables on a number of cases. In coding the data, researchers typically give a special numeric code to these missing values. In coding mobile16, for example, the GSS coders entered a value of 0, 8, or 9 for respondents who were not asked the question (“IAP”), did not know (“DK”), or for whom the information is otherwise not available (“NA”). Because these numeric codes have been set to missing (and thus appear in the Missing column), SPSS does not recognize them as valid codes and will not include them in an analysis of mobile16. In many cases, the author has set most missing values in the datasets to system-missing, which SPSS automatically removes from the analysis. However, when you use an existing variable to create a new variable, SPSS may not automatically transfer missing values on the existing variable to missing values on the new variable. Later in this volume, we discuss how to handle such situations.

**A MUST-DO: SETTING OPTIONS FOR VARIABLE LISTS**

Now you have a feel for the number-oriented side and the word-oriented side of SPSS. Before looking at how SPSS produces and handles output, you must do one more thing. To ensure that all the examples in this workbook correspond to what you see on your screen, you will need to follow the steps given in this section when you open each dataset for the first time.
DO THIS NOW: In the main menu bar of the Data Editor, click Edit → Options. Make sure that the General tab is clicked. (See Figure 1-4.) If the radio button Display names and the radio button Alphabetical were already selected when you opened the Options menu, you are set to go. Click Cancel. If, however, Display names and/or Alphabetical were not already selected when you opened the Options menu, select them (as in Figure 1-4). Click Apply. Click OK, returning to the Data Editor. When you open a new dataset for the first time, go to Edit → Options and ensure that Display names/Alphabetical are selected and applied.

THE VIEWER

We will run through a quick analysis and see how SPSS handles variables and output. On the main menu bar, click Analyze → Descriptive Statistics → Frequencies. The Frequencies window appears (Figure 1-5). There are two panels. On the right is the (currently empty) Variable(s) panel. This is the panel where you enter the variables you want to analyze. On the left you see the names of all the variables in GSS2012 in alphabetical order, just as you specified in the Options menu. Although the names are not terribly informative, complete coding information is just a (right) mouse click away. Suppose you want to analyze educ_4. Scroll the alphabetized list until you find educ_4. (Hint: Select any variable in the variable list. Type “e” on the keyboard. SPSS will go to the first e’s in the list.) Put the mouse pointer on the variable, educ_4, and right-click. Then click on Variable Information. As shown in Figure 1-6, SPSS retrieves and displays the label (Education: 4 Cats), name (educ_4), and, most usefully, the value labels for the numeric codes. (To see all the codes, click the drop-down arrow in the Value Labels box.)

Return the mouse to the Frequencies window and click educ_4 into the Variable(s) panel. (Click on educ_4 and then click the arrow between the panels.) Click OK. SPSS runs the analysis and displays the results in the Viewer (Figure 1-7). The Viewer has two panes. In the Outline pane, SPSS keeps a running log of the analyses you are performing. The Outline pane references each element in the Contents pane, which reports the results of your analyses. In this book we are interested exclusively in the Contents pane. Reduce the size of the Outline pane.

1. Click in the Values cell, and then click the gray button.

2. SPSS shows the word labels for each numeric value code. For example, respondents who live in a different state are coded 3 on mobile16.
1. Click Edit → Options.

2. In the Variable Lists panel, select Display names and select Alphabetical.

3. Click Apply. Click OK.

pane by first placing the cursor on the Pane divider. Click and hold the left button of the mouse and then move the Pane divider over to the left-hand border of the Viewer. The Viewer should now look like Figure 1-8. The output for educ_4 shows you the frequency distribution, with value codes labeled. In Chapter 2 we discuss frequency analysis in more detail. Our immediate purpose is to become familiar with SPSS output.

Here are some key facts about the Viewer. First, the Viewer is a separate file, created by you during your analysis of the data. It is completely distinct from the data file. Whereas SPSS data files all have the file extension *.sav, Viewer files have the file extension *.spv. The output can be saved, under a name that you choose, and then
Chapter 1

1. Click Analyze ➔ Descriptive Statistics ➔ Frequencies.

2. SPSS opens the Frequencies window.

Figure 1-5 Requesting Frequencies

1. Right-click on educ_4, and then click Variable Information.

2. Click the drop-down to see all the codes.

Figure 1-6 Retrieving Coding Information

reopened later. Second, the output from each succeeding analysis does not overwrite the file. Rather, it appends new results to the Viewer file. If you were to run another analysis for a different variable, SPSS would dump the results in the Viewer below the analysis you just performed. Third, the quickest way to return to the Data Editor is to click the starred icon on the menu bar, as shown in Figure 1-8. And, of course, Windows accumulates icons for all open files along the bottom Taskbar. Finally, you may select any part of the output file, print it, or copy and paste it into a word processing program.
Clicking the Go to Data icon returns you to the Data Editor.

Many of the exercises in this workbook will ask you to print the results of your SPSS analyses, so let’s cover the print procedure. We’ll also address a routine necessity: saving output.

**Selecting, Printing, and Saving Output**

Printing desired results requires, first, that you select the output or portion of output you want to print. A quick and easy way to select a single table or chart is to place the cursor anywhere on the desired object and click once.
For example, if you want to print the educ_4 frequency distribution, place the cursor on the frequency table and click. A red arrow appears in the left-hand margin next to the table (Figure 1-9). Now click the Printer icon on the Viewer menu bar. The Print window opens. In the window's Print Range panel, the radio button next to “Selected output” should already be clicked. Clicking OK would send the frequency table to the printer. To select more than one table or graph, hold down the Control key (Ctrl) while selecting the desired output with the mouse. Thus, if you wanted to print the frequency table and the statistics table, first click on one of the desired tables. While holding down the Ctrl key, click on the other table. SPSS will select both tables.

To save your output, simply click the familiar Save icon on the Viewer menu bar (refer to Figure 1-9). Browse for an appropriate location. Invent a file name (but preserve the .spv extension), such as “chap1.spv,” and click Save. SPSS saves all of the information in the Viewer to the file chap1.spv. Saving your output protects your work. Plus, the output file can always be reopened later. Suppose you are in the middle of a series of SPSS analyses and you want to stop and return later. You can save the Viewer file, as described here, and exit SPSS. When you return, you start SPSS and load a data file (like GSS2012) into the Data Editor. In the main menu bar of the Data Editor, you click File → Open → Output, find your .spv file, and open it. Then you can pick up where you left off.

**Figure 1-9 Selecting, Printing, and Saving Output**

**EXERCISES**

1. (Dataset: GSS2012. Variables: income06, attend.) Earlier we spent some time using the Data View and the Variable View to describe the first respondent in the GSS2012 dataset. In this exercise you will use your familiarity with the Data Editor to find out this respondent's income (income06) and how often this respondent attends religious services (attend).

   A. With GSS2012 open, go to the Data View. What numeric code does the first respondent have on income06? A code of (fill in the blank) ____________. Go to the Variable View. Just as you did earlier in this chapter, find income06 and click in the Values cell. What is this respondent's income? (circle one)

      
      $25,000 to $29,999   $60,000 to $74,999   $150,000 or over
B. Return to the Data View. What is this respondent's code on the variable attend? A code of (fill in the blank) _____________. Go to the Variable View. How often does this respondent attend religious services? (circle one)

Never       Once a year       2–3 times a month

2. Suppose that you have just opened the World, States, or NES2012 dataset for the first time. The first thing you do is to click Edit → Options and consider the Variable Lists panel of the General tab. You must make sure that which two choices are selected and applied? (check two)

- Display labels
- Display names
- Alphabetical
- File
- Measurement level