1

Introduction to SPSS

Watch screencasts of the guided examples in this chapter. **edge.sagepub.com/pollock**

Procedures Covered

File ▶ Open ▶ Data

Edit ▶ Options

Utilities ▶ Variables

Analyze ▶ Descriptive Statistics ▶ Frequencies

File ▶ Print

File ▶ Open ▶ Output

Format ▶ TableLooks (in Table Editing window)

File ▶ New ▶ Syntax

In this chapter, we take readers on a quick tour of the SPSS program. We describe the main windows that students will encounter: the welcome screen, the data editor, and the viewer (equivalent to the console in other statistical analysis programs). For maximum benefit, practice the steps and procedures we discuss here on your own computer.

THE DATA EDITOR

Open the General Social Survey dataset, GSS.sav, to get acquainted with SPSS's Data Editor. To open this dataset, locate the GSS.sav file in the folder where you saved it and double-click it. If you already have SPSS running, select File ▶ Open ▶ Data and find the GSS.sav file.

Recent versions of SPSS will open several windows at once. You may see a welcome screen (Figure 1-1). You can skip the welcome screen in the future by checking the "Don't show this dialog in the future" option in the lower left corner of the window.[1] Click the "Close" button to close the welcome screen.
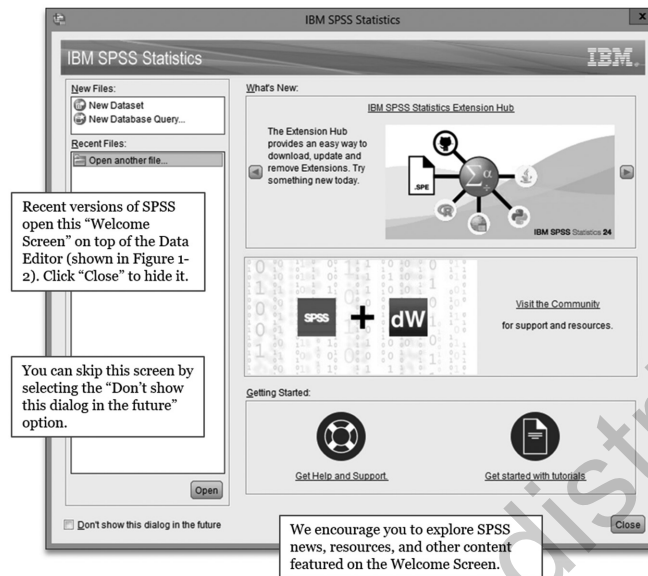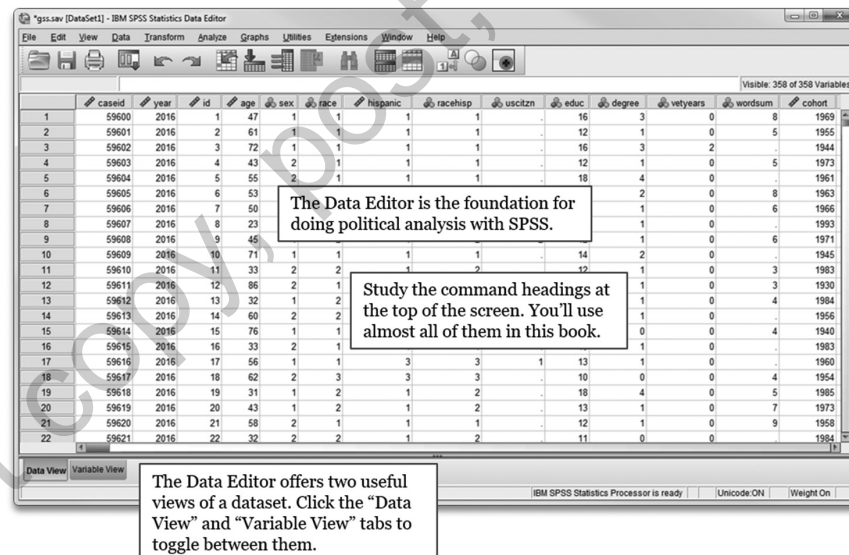
SPSS opens the data file and displays the Data Editor (Figure 1-2).

Screencast

Introduction to SPSS

---

[1] SPSS uses the welcome screen to promote some extensions you can add on to the program along with support resources. We won't delve into these extensions and resources, but they're worth exploring on your own.

FIGURE 1-1   SPSS "Welcome Screen"

Recent versions of SPSS open this "Welcome Screen" on top of the Data Editor (shown in Figure 1-2). Click "Close" to hide it.

You can skip this screen by selecting the "Don't show this dialog in the future" option.

We encourage you to explore SPSS news, resources, and other content featured on the Welcome Screen.

FIGURE 1-2   SPSS Data Editor in Data View

The Data Editor is the foundation for doing political analysis with SPSS.

Study the command headings at the top of the screen. You'll use almost all of them in this book.

The Data Editor offers two useful views of a dataset. Click the "Data View" and "Variable View" tabs to toggle between them.

Notice the two tabs at the bottom of the Data Editor window: Data View and Variable View. The SPSS Data Editor offers two "views" of the dataset. Both views are useful. Select Data View or Variable View by clicking one of the tabs at the bottom of the Data Editor.

Turn your attention to the Data View. (Make sure the Data View tab is clicked.) The Data View shows how all the cases are organized for analysis. Information for each case occupies a separate row. When you're working with the GSS dataset, each row represents a person who participated in the survey. Numbers in the "id" column record each respondent's case identification number ("caseid"). The variables, given brief yet descriptive names, appear along the columns of the editor.

Scroll right to see information recorded from the first respondent. You can tell that the first respondent in the dataset is 47 years old (see the first value in the "age" column). You can also

FIGURE **1-3**   SPSS Data Editor in Variable View
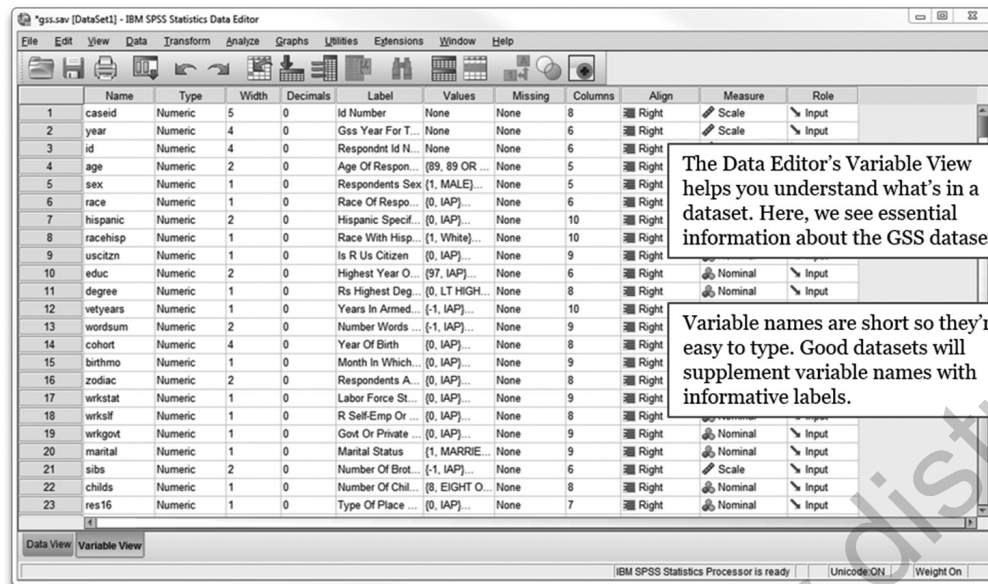
see that this respondent has a 1 in the "sex" column and a 1 in the "race" column. Storing information as numbers is efficient, but it's not immediately clear what these numbers really mean. To paint a more complete word portrait of this respondent, you need to see how all the variables are coded.

To see how the GSS variables are coded, click the Variable View tab (Figure 1-3). The Variable View, among other useful features, shows the word labels that the researcher has assigned to the numeric codes. 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.

Because variable names are short and often abbreviated, researchers will use labels, long descriptors (up to 256 characters are allowed), to provide more detailed information about variables. For example, when SPSS analyzes the GSS 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 descriptive labels to a variable's numeric values. We can examine the value labels for the sex and race variables to find out what it means when they're coded 1. To find out the value labels for the sex variable, find its row in the Variable View, 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 sex. Unless you instruct it to do otherwise, SPSS will apply these labels to its analysis of the sex variable. Repeat this process to find out what value 1 on the race variable signifies, or what the numeric codes of the "marital" variable mean (see Figure 1-4). You can see that respondents who have never been married are coded 5 on the variable marital. Click the Cancel button in the Value Labels window to return to the Variable View.

FIGURE 1-4   Value Labels of a Variable in the Data Editor
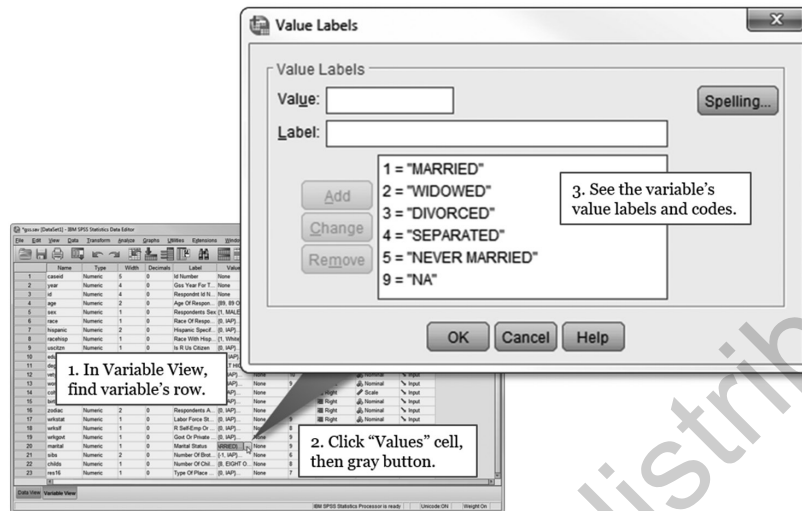
Finally, a word about the Missing column. Sometimes a dataset does not have complete information for all variables and observations. This happens for a variety of reasons; researchers may add or remove questions from the survey, some questions may not apply to everyone, or the response may not be clear. In coding the data, researchers typically give special numeric codes to missing values. In coding mobile16, for example, the GSS coders entered a value of 0 for respondents who were not asked the question ("IAP"), 8 for respondents did not know ("DK"), and 9 when the information was 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.

SETTING OPTIONS FOR VARIABLE LISTS

Now you have a feel for how data are organized and stored in 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, select Edit ▶ Options. Make sure that the General tab is clicked (see Figure 1-5). If the Variables Lists options for Display names and Alphabetical are not already selected, select them (as in Figure 1-5). Click Apply and then 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. This will help you find variables to analyze more efficiently. (If the radio button Display names *and* the radio button Alphabetical are already selected when you opened the Options menu, you are set to go and can click the Cancel button.)

FIGURE 1-5   Setting Options for Variable Lists
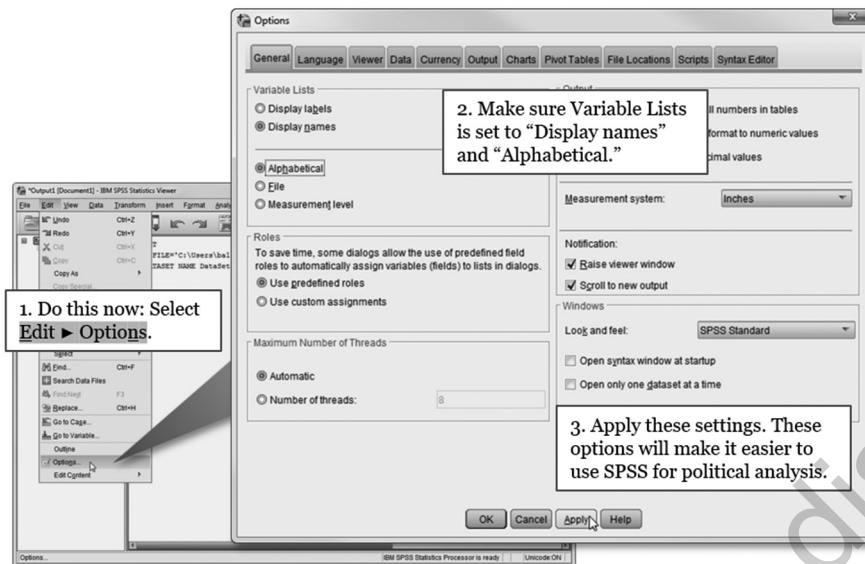
A CLOSER LOOK: VARIABLES UTILITY

Although the names of GSS variables are not terribly informative, SPSS makes it easy to view complete coding information. In the text, we show how you can access information about variables in the Data Editor's Variable View. You can also view detailed information about the variables in a dataset using Utilities ▶ Variables. This selection will yield the Variables window (Figure 1-6).

Suppose you want to view detailed information about the GSS variable marital to better understand how the dataset records respondents' marital statuses. Scroll through the alphabetical list of variables on the left side of the Variables window until you find "marital" and select it. You'll then see some essential information about the variable, like the text label associated with it ("Marital Status"), along with a breakdown of how different marital statuses are encoded.

FIGURE 1-6   Retrieving Coding Information
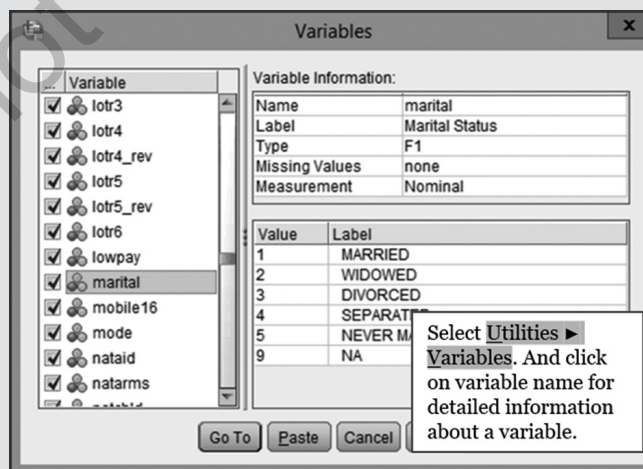
THE VIEWER

We will run through a brief example to show how SPSS analyzes variables and generates output. You execute most SPSS commands using a graphical user interface (GUI). SPSS's methods of analyzing variables are organized under the "Analyze" tab. You'll start most of your data analysis by clicking the Analyze tab and selecting the type of analysis you wish to perform from its menu of options. For this example, select Analyze ▶ Descriptive Statistics ▶ Frequencies. The Frequencies window appears (Figure 1-7).

You'll notice that the Frequencies window has 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 GSS in alphabetical order, just as you specified in the Options menu.[2]

Scroll down the alphabetized list of GSS variables window until you find marital and add "marital" to the Variable(s) panel. (*Hint*: Click anywhere on the variable list and type "m" on the keyboard. SPSS will jump to the first m's in the list.) To add marital to the Variable(s) panel, click on marital and then click the arrow between the panels or drag and drop marital from the alphabetical list to the Variable(s) panel.[3] Click OK. SPSS runs the analysis and displays the results in the Viewer (Figure 1-8).

A frequency distribution table for the marital statuses of GSS respondents appears in the Viewer. We'll have more to say about frequency distribution tables in the next chapter and discuss many different types of tables in this book. In the future, we'll focus on the tables SPSS generates and won't show the entire Viewer as we do in Figure 1-8.

When you execute a procedure using the GUI, SPSS temporarily stores the information you inputted so you can return to the same window and adjust your selections. This is particularly useful

FIGURE 1-7   Analyzing Frequencies

---
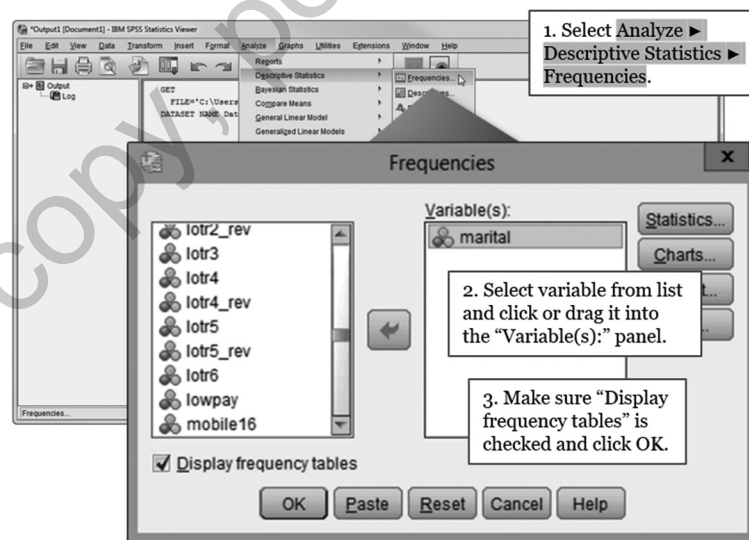
[2] If you don't see an alphabetized list of variable names in the Frequencies window, follow the DO THIS NOW instructions in the "Setting Options for Variable Lists" section above. Setting correct options for variable lists will make it easier for you to execute SPSS commands.

[3] You can also access variable information within this dialog. Put the mouse pointer on the variable, marital, and right-click. Then click on Variable Information. SPSS retrieves and displays the variable's name (marital), label (Marital Status), and, most usefully, the value labels for the marital variable's numeric codes. (To see all the codes, click the drop-down arrow in the Value Labels box.)

**FIGURE 1-8**   Sample Table Output in the SPSS Viewer
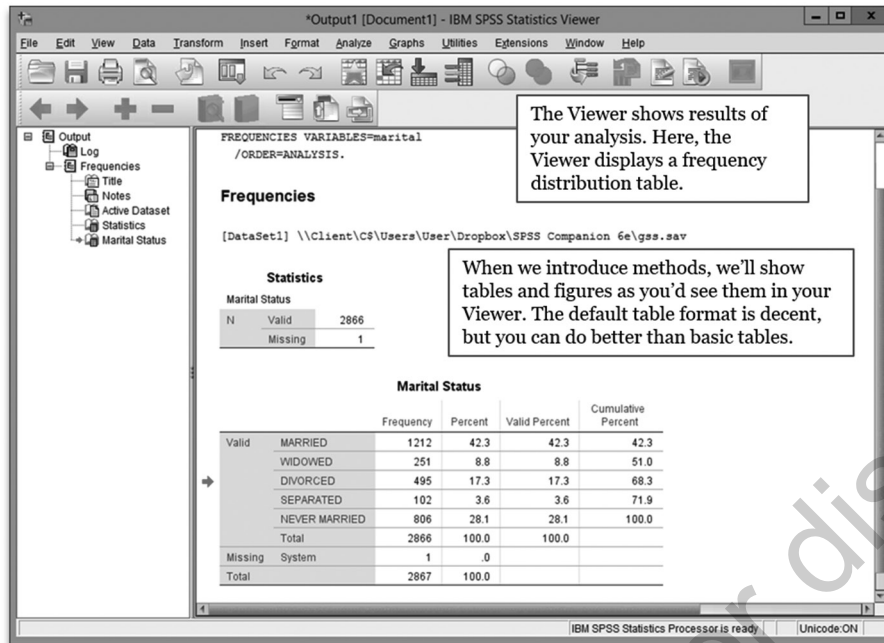
when you're executing complex commands for the first time, making graphics, or performing the same operation repeatedly on different variables. After running the frequency analysis illustrated in Figures 1-7 and 1-8, for example, you could select Analyze ▶ Descriptive Statistics ▶ Frequencies again and find marital still on the variables list, making it easy to change the command settings or analyze a different variable.

A CLOSER LOOK: KEYBOARD SHORTCUTS

Some SPSS users may find navigating the GUI cumbersome after a while. Those who prefer the keyboard to the mouse will be happy to know that there is an easy way to navigate the GUI using keyboard shortcuts. To get to the Frequencies window, hold down the "Alt" key and press "A", "E", then "F" (you don't need to capitalize the letters or use quotation marks). You can navigate the SPSS menu by pressing Alt followed by the letter(s) corresponding to different branches of the menu. If you look closely at the command notations above, you'll see that we've underlined the letters A, e, and F to show the keyboard shortcuts and we follow this convention throughout the book when we introduce new procedures.

Notice that the SPSS Viewer has two panes. In the Outline pane, SPSS keeps a running list 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.

You can minimize the Viewer's Outline 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 1-9. The Contents pane shows you the frequency distribution of the marital variable with value codes labeled. In Chapter 2 we discuss frequency analysis in more detail. Our immediate purpose is to become familiar with SPSS output.

FIGURE 1-9 SPSS Viewer with the Outline Pane Minimized
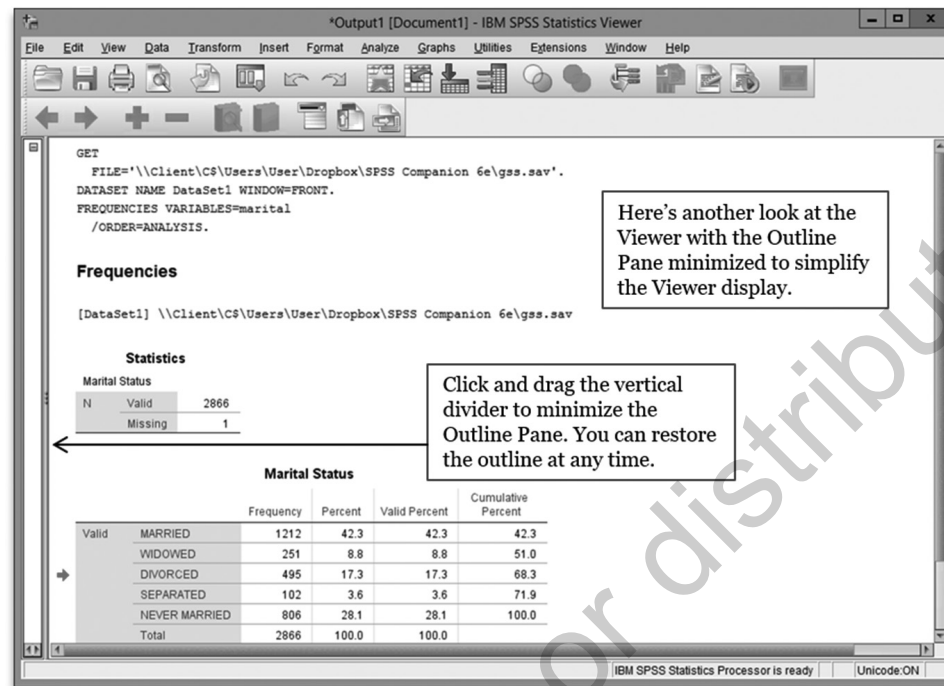
Here are some key points about the Viewer to keep in mind:

- The Viewer is a separate file, created by you during your analysis of the data. The Viewer file, a log of your analysis and output, 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 reopened later in SPSS. You can't open a *.spv file in another program, like a word processor (e.g., Microsoft Word); if you want to use SPSS output in a document, follow the directions below for exporting graphics and copying tables.

- Output from each succeeding analysis does not overwrite the Viewer's *.spv 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.

- 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.

- The "Analyze" tab appears at the top of the Viewer and the Data Editor so you can start your data analysis from either SPSS window. In fact, all tabs in the Data Editor window appear in the Viewer window (along with two tabs that appear only in the Viewer: Insert and Format). Because the results of your analysis appear in the Viewer, it makes sense to start your analysis there, but you can get the same results starting from the Data Editor.

As we discuss in the next two sections, you may select any part of the output file, format it, print it, or copy and paste it into a word processing program.

SELECTING, PRINTING, AND SAVING OUTPUT

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.

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 marital frequency distribution table produced in the preceding section, place the cursor on the frequency table and click. A red arrow appears in the left-hand margin next to the table (Figure 1-10). Click the Printer icon on the Viewer menu bar or select File ▶ Print. 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 copy your Viewer output to your computer's clipboard to paste into another document, simply select the table(s) you want to copy, right-click, and select the "Copy" option (see Figure 1-10). Recent versions of SPSS have greatly improved table formatting over prior versions and the table copied from the Viewer now looks decent in a document.

To save your Viewer output, simply click the familiar Save icon on the Viewer menu bar (refer to Figure 1-10 again). 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 GSS.dta) into the Data Editor. In the main menu bar of the Data Editor, you select File ▶ Open ▶ Output, find your .spv file, and open it. Then you can pick up where you left off.

FIGURE **1-10**   Selecting, Printing, and Saving Output
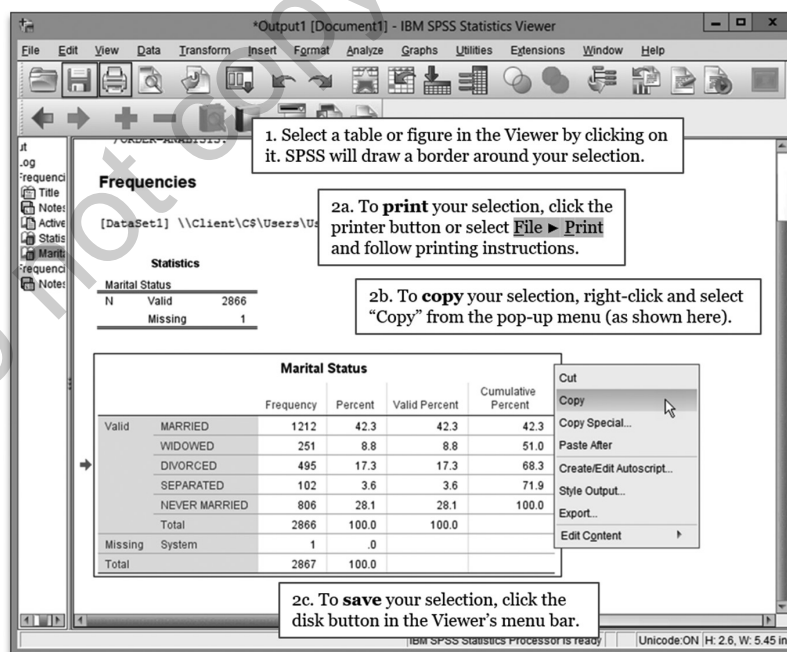
If you want to save a table or graphic that appears in the Viewer but don't want to save all of your Viewer output, select the Viewer object you want to save and right-click it. Select "Export . . . " from the pop-up dialog window (see Figure 1-11). If you've selected a table, you can export it to a variety of document types, such as a .pdf document. If you've selected a graphic, you can create a variety of different types of image files.

HOW TO FORMAT AN SPSS TABLE

When you analyze political science data, you'll create a lot of tables of results. You want the tables you create to communicate the results of your analysis effectively. No one wants to try to decipher results from a mess of numbers. Fortunately, SPSS offers some easy-to-use options for formatting tables.

FIGURE **1-11** Exporting Viewer Output
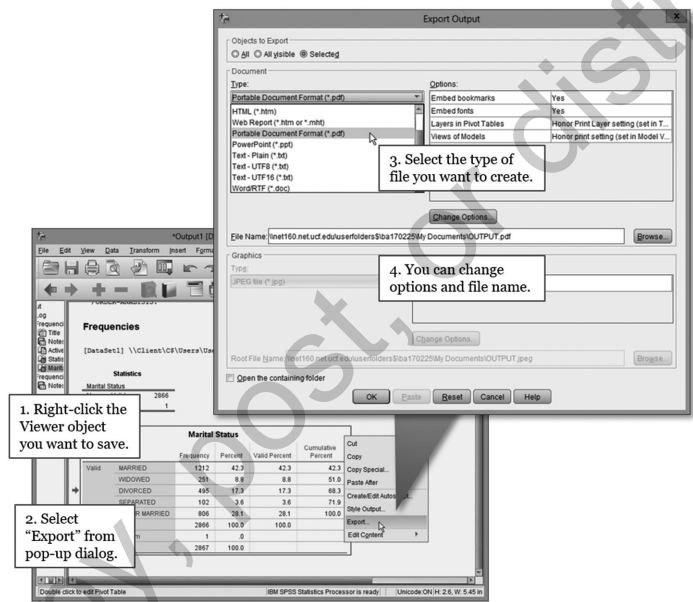

FIGURE **1-12** Formatting a Table
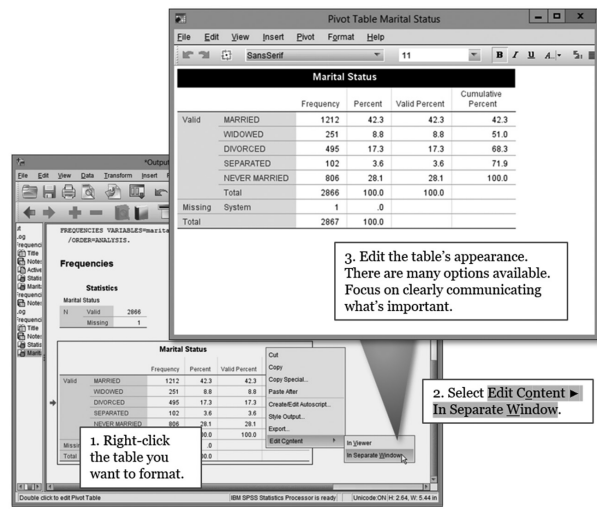
To access the table formatting tools, select the table you want to format and right-click it (like you would to copy or export it). Now, select the "Edit Content ▶ In Separate Window" option. You'll see your table in a new, separate window with formatting tools (see Figure 1-12). This table editing window allows you to change the look and feel of your table; you can change the colors, shading, fonts, alignments, and more. For example, you can widen the far-right column of the marital status frequency distribution table so the heading "Cumulative Percent" stays on one line. Keep in mind what you're attempting to communicate. If you're conducting serious analysis for an academic paper, a tropical color theme isn't the best choice.

Here's a suggestion to help you quickly create professional-style tables. The Academic-style tables are particularly good. In the separate table editing window, select Format ▶ TableLooks (this procedure is only available in the table editing window). TableLooks are a pre-defined set of table styles. There are many styles to choose from, but the "Academic" style is a solid choice for most of your political analysis. Select "Academic" from the list of TableLook files and click OK.

You'll see your table with your TableLook selection applied in the separate table editing window. You can close the separate table editing window by clicking the X button in the upper-right corner. Now you should see your freshly formatted table in the Viewer. If you selected the Academic look for the marital variable's frequency distribution table, your table should look like this:

Marital Status

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	MARRIED	1212	42.3	42.3	42.3
	WIDOWED	251	8.8	8.8	51.0
	DIVORCED	495	17.3	17.3	68.3
	SEPARATED	102	3.6	3.6	71.9
	NEVER MARRIED	806	28.1	28.1	100.0
	Total	2866	100.0	100.0	
Missing	System	1	.0		
Total		2867	100.0		

FIGURE **1-13**   Creating Academic-style Tables
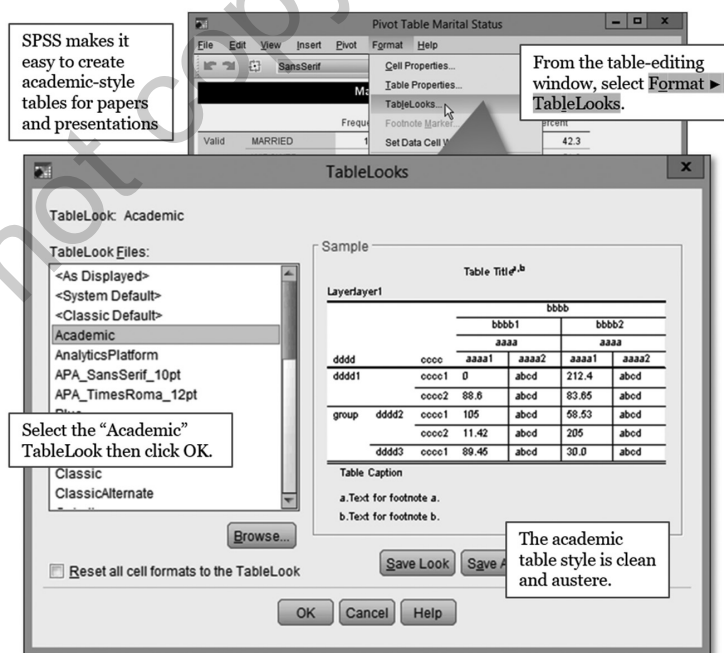
In this book, we'll show tables in the SPSS default style to make it easier to follow our examples, but you can make the Academic-style table your default format by selecting Edit ▶ Options and then the Pivot Tables tab, select Academic from the TableLook options, and then click OK. Tables in this pre-defined SPSS table format look a lot like the tables one sees in top political science journals. It's a good look for your political analysis.

SAVING COMMANDS IN SYNTAX FILES

In this book, we show how to implement essential political science research methods using SPSS's graphical user interface. SPSS's GUI offers a straightforward and consistent framework for analyzing data. In some situations, however, you may want to document your data analysis to enable others to see what you did and replicate your analysis. These situations call for making a syntax file.

A syntax file is a text document with the *.sps file extension that records the series of commands used to perform some data analysis. The procedures you execute using SPSS's GUI can also be stated as text commands. If you look closely at Figure 1-9, you'll see that SPSS displays the text equivalent of the frequency analysis of the marital variable in the Viewer:

FREQUENCIES VARIABLES=marital
/ORDER=ANALYSIS.

This text-command equivalent of the frequency analysis demonstrated above, executed from a syntax file, would yield the same results as using the GUI, enabling someone else to replicate the analysis.

This book demonstrates the essentials of political analysis without using the esoteric SPSS command language. For most data analysis tasks, the GUI works fine and will allow you to start applying core concepts much sooner than encoding commands. Thankfully, SPSS makes it easy for users to "reverse engineer" a syntax file for replication purposes.

To create a syntax file, select File ▶ New ▶ Syntax. This selection will call up a new window, the Syntax Editor (see Figure 1-14, which shows commands executed in this chapter). As we've seen, SPSS prints the text-command equivalent of procedures implemented through its GUI in the Viewer. You can copy and paste these text-equivalent commands into a syntax file. To create a complete

FIGURE 1-14   The Syntax Editor
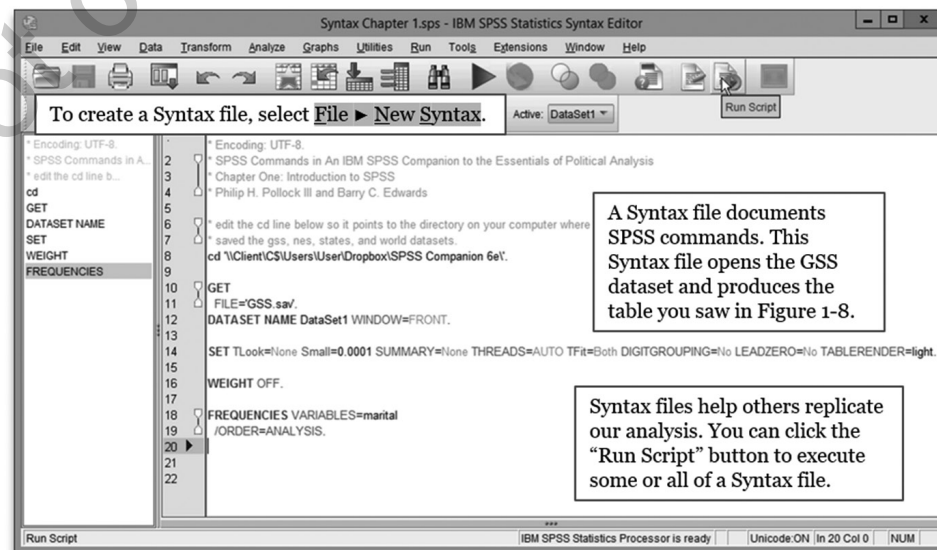
syntax file, you can also copy and paste the commands to get the GSS dataset, set viewing options, and execute the analysis. The grayed-out lines that start with * are comments (lines in the syntax file to be read by human users rather than executed by SPSS).

Another way to save your commands in a syntax file is the "Paste" button that appears in GUI windows that execute commands. Take another look at Figure 1-7, the procedure we used to generate a frequency distribution table for the marital variable. Next to the OK button, you'll see the Paste button. If you press this button, SPSS will paste the syntax for the procedure at the end of your syntax file (if you don't have a syntax file open, SPSS will open a new one). For completeness, you may want to add the commands to open the dataset and some user-friendly comments, but it's a very convenient feature for generating syntax files that replicate your analysis.

GETTING HELP

To view the formal how-to manual for any SPSS procedure, you can click the "Help" button from the GUI window that executes that procedure. For example, if you want to see detailed instructions on the frequency analysis procedure you used earlier in this chapter, you could click the Help button in the Frequencies window (see Figure 1-7). SPSS retrieves the technical manual information and displays it in a web browser (Figure 1-15).

You can also get help by watching screencasts the authors of this book produced to show students how to execute the procedures discussed in this book. We've included links to these screencasts throughout this book. You can find a complete list of screencasts on the SAGE Edge website for this book: **edge.sagepub.com/pollock**.

FIGURE **1-15**   SPSS Help Manual
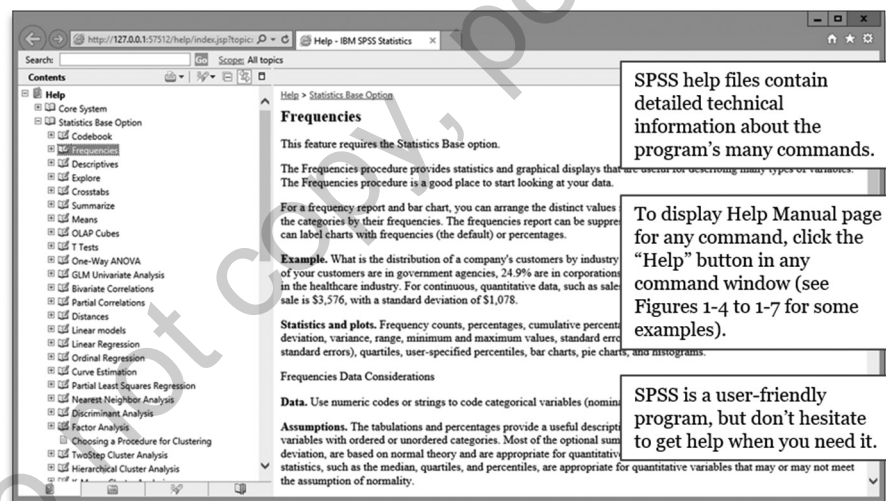
Name: _____   Date: _____

E-mail: _____   Section: _____

CHAPTER 1 EXERCISES

1. (Dataset: gss.sav. Variables: income16, attend.) In this chapter, we spent some time using the Data View and the Variable View to describe the first respondent in the GSS dataset. In this exercise, you will use your familiarity with the Data Editor to find out this respondent's income (income16) and how often this respondent attends religious services (attend).

   A. With the GSS dataset open, go to the Data View. What numeric code does the first respondent have on income16? A code of (fill in the blank) _____. Go to the Variable View. Just as you did earlier in this chapter, find income16 and click in the Values cell. What is this respondent's income? (circle one)

   $25,000 to $29,999 $60,000 to $74,999
   $170,000 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 NES dataset for the first time. The first thing you do is select 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

3. The GSS dataset contains the variable happy, which asks respondents how happy they are. Find the variable happy in the Data Editor's Variable View and answer the following questions.

   A. What is the descriptive label for the variable happy? _____

   B. Respondents who say they are "very happy" have which numeric code? (circle one)

   1   2   3

4. Generate a frequency distribution table for the variable happy in the GSS dataset using the Analyze ▶ Descriptive Statistics ▶ Frequencies procedure.

   A. Print the table.

   B. Copy/paste the table into a blank document. Edit the table for appearance and readability, and then print it.

5. SPSS works with different types of files to analyze data and produce results. These file types include datasets that record information about sample observations, output files of SPSS results, and syntax files used to replicate SPSS commands. Each of these different file types has a unique, three-character file name extension. Complete the table below to help you remember the file name extension for each file type.

SPSS File Type	File Name Extension
Dataset	?
Output	?
Syntax	?

6. A political scientist wants to analyze civic culture in the United States. Civic culture is an important concept but is difficult to measure empirically. The researcher could consider several different variables.

   A. Which variable in the States dataset records the number of years of social studies that states require students to take to graduate high school? _____

   B. Which variable in the States dataset records the percentage of the voting age population that turned out to vote in the most recent federal election (for which data are available)? _____

   C. Which variable in the States dataset records the percentage of state residents who frequently attend religious services? _____

D. Which variable in the States dataset records the percentage of state residents who do voluntary community service? _____

7. Chapter 1 showed you how to view detailed information about a variable named marital in the GSS dataset. This variable identifies the survey respondents' marital statuses with numeric codes and assigns a label to each numeric value. The NES dataset contains a similar variable, also named marital, that records the marital statuses of its respondents. The numeric coding for marital in the NES is a little different than its GSS counterpart, however. It's important to pay close attention to how variables are coded. Fill in the following table to identify the labels that correspond to numeric codes of the NES marital variable.

NES Variable "marital"	
Numeric Code	Value Label
1	?
2	?
3	?
4	?
5	?
6	?

8. A political scientist wants to study health outcomes in countries around the world. Good government may improve health outcomes; a healthy country may also be more politically stable. To study the health-politics relationship, the researcher could consider several variables in the World dataset that measure health in countries around the world. For each variable, briefly describe what it measures in your own words.

A. fertility _____

_____

_____

_____

_____

B. hiv_percent _____

_____

_____

_____

_____

C. infant_mortality _____

_____

_____

_____

_____

D. spendhealth _____

_____

_____

_____

E. unnoncom _____

_____

_____

_____

_____