INTRODUCTION TO SOFTWARE, DATA SETS, AND VARIABLES

End-of-chapter practice problems have been organized into three sections: SPSS, Excel, and Chapter Exercise calculation and interpretation problems. Chapter Exercises do not require the use of computer software. SPSS Problems are based on the program IBM SPSS Version 25 or IBM SPSS Statistics Base Student Edition 24. Excel Problems use Microsoft Excel for MacBook Version 16.16.10.

Before attempting the SPSS and/or Excel Problems, you will find demonstrations that we strongly encourage you to work through. The demonstrations and related problems are organized by software and labeled accordingly: SPSS Demonstrations are followed by SPSS Problems. The same format follows for Excel. Excel Demonstrations are followed by Excel Problems.

For all SPSS and Excel Problems in this textbook, we will be working with 2018 General Social Survey (GSS) data. The GSS has been conducted biennially since 1972. Conducted by the NORC at the University of Chicago, with principal funding from the National Science Foundation, the GSS is designed to provide social science researchers with a readily accessible database of socially relevant attitudes, behaviors, and attributes of a cross section of the U.S. population. Next to the U.S. Census data, the GSS is the most frequently analyzed source of social science information by educators, legislators, and media outlets. From the GSS, we've created two data sets for use with SPSS and titled them: GSS18SSDS-A and GSS18SSDS-B. They each contain a selection of 50 variables[11] and 1,500 cases.

We also created one data set for use with Excel and titled it GSS18SSDS-E. It contains 22 variables, 135 cases, and two sheets (Data View and Variable View). We locked both Excel sheets to avoid any changes that might accidentally be introduced as you click around the GSS18SSDS-E data set. All this means is that you cannot change any information in any of the cells. You can easily unlock an Excel sheet by clicking on *Home → Format → Unprotect Sheet*. If you have not yet installed Excel's Analysis ToolPak (an available add-in), you can do so by selecting *Tools → Excel Add-ins → Analysis ToolPak → OK* in the main toolbar. After adding the Analysis ToolPak, you will find the Data Analysis option off to the right in the Excel Data Tab.

SPSS DEMONSTRATION [GSS18SSDS-A]

The SPSS appendix found on this text's study site explains the basic operation and procedures for SPSS for Windows Student Version. We strongly recommend that you refer to this appendix before beginning the SPSS exercises.

When you begin using a data set, you should take the time to review your variables. What are the variables called? What do they measure? What do they mean? There are several ways to do this.

To review your data, you must first open the data file. Files are opened in SPSS by clicking on *File*, then *Open*, and then *Data*. After switching directories and drives to the appropriate location of the files (which may be on a hard disk or on a ZIP drive), you select one data file and click on *Open*. This routine is the same each time you open a data file. SPSS automatically opens each data file in the SPSS Data Editor window labeled Data View. We'll use GSS18SSDS-A for this demonstration.

One way to review the complete list of variables in a file is to click on the *Utilities* choice from the main menu, then on *Variables* in the list of submenu choices. The SPSS variable names, which are limited to eight characters or less, are listed in the scroll box (refer to Figure 1.4). When a variable name is highlighted, the descriptive label for that variable is listed, along with any missing values and, if available, the value labels for each variable category. (As you use this feature, please note that sometimes SPSS mislabels the variable's measurement level. Always confirm that the reported SPSS measurement level is correct.) SPSS allows you to display data in alphabetical order (based on the variable name) or in the order presented in the file (which may not be alphabetical).

A second way to review all variables is through the Variable View window. Notice on the bottom of your screen that there are two tabs, one for *Data View* and the other for *Variable View*. Click on *Variable View*, and you'll see all the variables listed in the order in which they appear in the Data View window (as depicted in Figure 1.5). Each column provides specific information about the variables. The columns labeled "Label" and "Values" provide the variable label (a brief label of what it's measuring) and value labels (for each variable category).
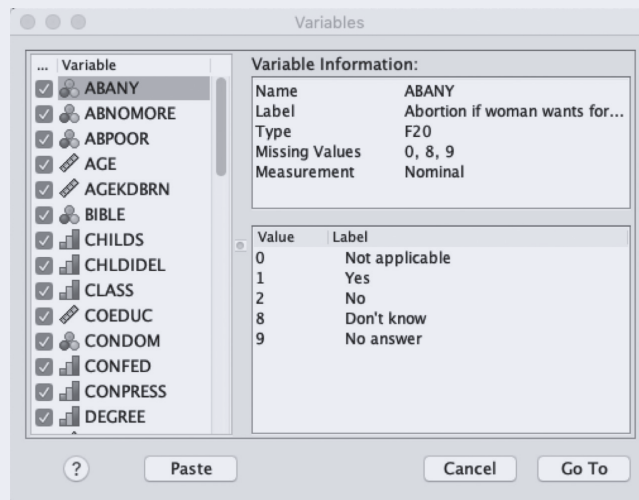
Figure 1.4

Figure 1.5

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ABANY	Numeric	20	0	Abortion if wo...	{0, Not appl...	0, 8, 9	9	≡ Right	⚃ Nominal	⟋ Input
2	ABNOMORE	Numeric	20	0	Married--want...	{0, Not appl...	0, 8, 9	9	≡ Right	⚃ Nominal	⟋ Input
3	ABPOOR	Numeric	20	0	Low income--c...	{0, Not appl...	0, 8, 9	9	≡ Right	⚃ Nominal	⟋ Input
4	AGE	Numeric	20	0	Age of respond...	{89, 89 or ol...	99	9	≡ Right	⟋ Scale	⟋ Input
5	AGEKDBRN	Numeric	20	0	R's age when 1...	{0, Not appl...	0, 98, 99	9	≡ Right	⟋ Scale	⟋ Input

SPSS PROBLEM [GSS18SSDS-A]

S1. Based on the *Utilities-Variables* option, review the variables from the GSS18SSDS-A. Can you identify three nominal variables, three ordinal variables, and at least one interval-ratio variable? Based on the information in the dialog box or Variable View window, you should be able to identify the variable name, variable label, and category values.

EXCEL DEMONSTRATION [GSS18SSDS-E]

Throughout our discussion of the Excel program in this book, we focus on how to create tables and produce summaries of data to enhance your learning of statistics. We do not discuss how to enter data into Excel, for we will be working with an existing data set that we've created especially for this textbook (GSS18SSDS-E). For many students, entering data into Excel is the easy part. It's using the program to summarize the data that most PEOPLE find challenging. If you would like a greater discussion of the full range of capabilities Excel offers, we suggest you review any number of more exhaustive Excel guides that are available.

Microsoft Office is a relatively affordable software bundle that includes Microsoft Word, Microsoft PowerPoint, Microsoft Excel, and more. If you are taking your statistics class near the end of your undergraduate career, you are likely very proficient with Word and PowerPoint. While we are sure you've heard of Excel, our experience teaching statistics over the years suggests you are probably not very proficient using it to analyze data. Given how readily available Excel is to most computer owners, we believe it is underused in undergraduate statistics courses. Let's change that.

Open up GSS18SSDS-E and examine its contents. The cells shaded in light gray indicate missing data. Also, notice how we've created two Excel sheets: *Data View* and *Variable View*. Figure 1.6 shows the Data View sheet. The first row (not pictured) is a list of all of the variables in the data set beginning with ABANY, which stands for "Abortion if Woman Wants for Any Reason." Beginning with row 2, each row represents an individual respondent. As pictured in Figure 1.6, the last respondent in the data set is in row 136. Because our respondents begin on row 2, we know that there are a total of 135 (136 – 1) respondents in our data set.

Figure 1.6

Figure 1.7

If you click on the "Variable View" tab at the bottom of the file, you will move to the second sheet that contains a list of 22 variables in GSS18SSDS-E. The first row is for information purposes only and is not a variable. See Figure 1.7. This sheet offers more information about each variable, including each variable's label.

EXCEL PROBLEM [GSS18SSDS-E]

E1. Closely examine the Variable View tab of GSS18SSDS-E and identify two nominal variables, two ordinal variables, and two interval-ratio variables.