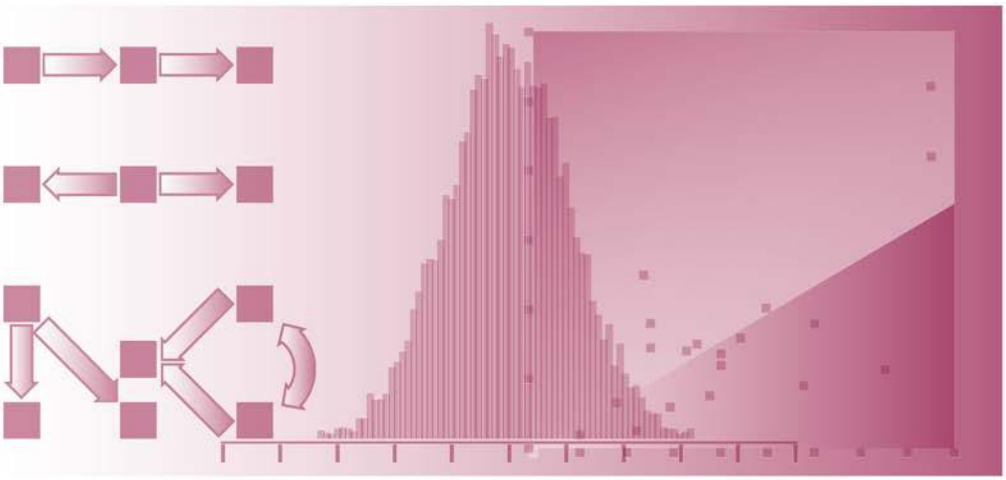
A Simple Guide to IBM SPSS®

for Version 23.0



Lee A. Kirkpatrick
Brooke C. Feeney

A Simple Guide to IBM SPSS® Statistics

for Version 23.0

A Simple Guide to IBM SPSS® Statistics

for Version 23.0

Lee A. Kirkpatrick

The College of William & Mary

Brooke C. Feeney

Carnegie Mellon University

CENGAGE Learning

Australia • Brazil • Mexico • Singapore • United Kingdom • United States

CENGAGE Learning

***A Simple Guide to IBM SPSS® Statistics for Version 23.0*, Fourteenth Edition**
Lee A. Kirkpatrick and Brooke C. Feeney

Product Director: Jon-David Hague
Product Manager: Timothy Matray
Product Assistant: Kimiya Hojjat
Marketing Manager: Melissa Larmon
Art and Cover Direction, Production Management, and Composition: Lumina Datamatics, Inc.
Manufacturing Planner: Karen Hunt
Intellectual Property Analyst: Deanna Ettinger

WCN: 02-200-208

ISBN: 978-1-305-87771-9

Cengage Learning
20 Channel Center Street
Boston, MA 02210
USA

Cengage Learning is a leading provider of customized learning solutions with employees residing in nearly 40 different countries and sales in more than 125 countries around the world. Find your local representative at **www.cengage.com**.

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

To learn more about Cengage Learning Solutions, visit **www.cengage.com**.

Purchase any of our products at your local college store or at our preferred online store **www.cengagebrain.com**.

Printed in the United States of America
Print Number: 01 Print Year: 2015

Contents

Preface

A tremendous variety of computer software options is available for use in introductory statistics and research methods courses. Although programs designed specifically for teaching purposes and/or ease of use have obvious advantages, there are also good reasons to prefer a more advanced, research-oriented program such as SPSS. First, SPSS is in widespread use, so if one moves to another setting to finish college, pursue a graduate education, or work in an applied research setting, the odds are good that one version or another of SPSS will be available at the new location. Second, learning a powerful program such as SPSS in an introductory course prepares the student for data analysis in subsequent, more advanced courses, as well as in "real" research projects later on. It might be a little rough at first, but in the long run the extra investment of time and effort early on will pay off.

This edition, like the previous versions of this book, was written to cover what the first-time or casual user needs to know—and *only* what the user needs to know—to conduct data analyses in SPSS at the level of an introductory statistics course in psychology. Other books and manuals have been published with similar intentions, but we have not found one that we felt quite accomplished the goals as simply and inexpensively as we think should be possible. Some, for example, give equal coverage to one or more computer programs in addition to SPSS; the reader interested only in SPSS must sift through (and pay for!) the irrelevant pages to find what he or she needs. Other books attempt to achieve a level of comprehensiveness paralleling that of the program itself, delving into advanced multivariate techniques, complicated data transformation procedures, and so forth that are of little value to the beginning student. Still other books delve deeply into the theory and mathematics of the procedures, and consequently overlap with—and potentially conflict with—textbook material and classroom instruction. In contrast, our approach is to create a simple anthology of examples to illustrate the kinds of analyses typically covered in an introductory statistics course in psychology, providing just enough explanation of procedures and outputs to permit students to map this material onto their classroom and textbook knowledge.

Back in the 1990's before the Windows operating system was born, the only way to use SPSS was to learn its programming language and type in the appropriate commands. Beginning with the Windows versions, SPSS introduced an interface (which we refer to as the *Point-and-Click Method*) that eliminates the need to learn any syntax or command language. Rather than typing commands into the program, as was required in SPSS/PC+, the user merely point-and-clicks his or her way through a series of windows and dialog boxes to specify the kind of analysis desired, the variables involved, and so forth. The program is generated internally and thus is invisible to the user. The user can then examine the output without ever viewing the program code itself—in fact, users must go out of their way to view the internally generated program code at all—and without having to ever learn any syntax or computer programming skills.

Although the simplicity of this procedure seems ideal at first, we believe it is at best a two-edged sword. IBM SPSS Statistics also permits the user to type commands the old-fashioned way (which we refer to as the *Syntax Method*) rather than simply pointing-and-clicking. There are several good reasons for learning how to use SPSS in this way rather than relying exclusively on the Point-and-Click Method. First, we think there is pedagogical value in students' learning a little computer programming, if only at the simple level of the SPSS command language, as part of an introductory statistics course. Second, for anyone planning to use SPSS for more advanced purposes than completing a few homework assignments, there is simply no escaping SPSS syntax: You'll probably want to learn it sooner or later. As you move on to more sophisticated uses of the program for real research projects, the advantages of being able to write and edit command syntax increase. For example, if you mis-specify a complex analysis and want to go back and rerun it with minor changes, or if you wish to repeat an analysis multiple times with minor variations, it is often more efficient to write and edit the program code directly than to repeat point-and-click sequences. Finally, and perhaps most important, several elements of the SPSS command language, including several procedures covered in this manual, are available only via the Syntax Method.

The question of which method is better learned first, particularly in introductory statistics courses, is a difficult one on which knowledgeable and well-intentioned instructors can (and do) honestly disagree. Consequently, the book covers both methods. Thus, beginning users and course instructors can choose to focus on one or the other method—or both, in parallel or sequentially.

In this edition, we have updated our latest book, *A Simple Guide to SPSS: For Version 22.0* (Wadsworth, 2014) to include coverage of the most recent release from SPSS, Version 23.0. Users of older versions of SPSS will find one of our previous books appropriate for their respective version. Like the previous editions, this book is designed primarily to serve as an inexpensive companion volume to any standard introductory statistics text and for use in such a course. The types of analyses covered, as well as their sequence of presentation, are designed to parallel the sequence of topics covered in a typical introductory statistics course in psychology. However, the book should prove useful not only in other courses, but to first-time users of SPSS in many other contexts as well.

Good luck, and happy computing!

About This Edition

In the last few years, the name of the program described in this book has gone through a series of changes. After many years in which it was known as *SPSS for Windows*, it briefly was renamed *PASW Statistics* ("Predictive Analytics Software") before changing again to *IBM SPSS Statistics.* The title of the current edition of this book correctly reflects the latest name. Fortunately, though, the operation and appearance of the program itself changed very little during this time—at least as far as the contents of this book are concerned—so users need not be concerned about which name appears on the version of the program they are using in conjunction with the book.

Acknowledgments

We would like to thank the following reviewers of the manuscript for their helpful suggestions: Terry Ackerman of the University of Illinois, John Cope of East Carolina University, William Frederickson of the University of Central Oklahoma, Ira T. Kaplan of Hofstra University, Donald F. Kendrick of Middle Tennessee State University, and Michael Poteat of East Carolina University.

Lee A. Kirkpatrick
Brooke C. Feeney

About the Authors

Lee A. Kirkpatrick is professor of psychology at the College of William & Mary in Williamsburg, Virginia. He received his B.S. from Lynchburg College, his M.A. in general/experimental psychology from the University of Texas at El Paso, and his Ph.D. in social/personality psychology from the University of Denver. For more information, visit his home page at http://faculty.wm.edu/lakirk.

Brooke C. Feeney is professor of psychology at Carnegie Mellon University. She received her B.A. from Salisbury State University, her M.A. in general/experimental psychology from the College of William & Mary, and her Ph.D. in social psychology from the State University of New York at Buffalo. Her research interests include attachment, social support, and close relationships. For more information, visit her home page at http://www.psy.cmu.edu/people/feeney and her lab website at http://www.cmu.edu/dietrich/psychology/relationships-lab.

PART 1

How to Use SPSS for Windows

SPSS for Windows, like earlier versions of SPSS, is designed to be a relatively comprehensive data analysis package for use in research and business. As such, it is designed to do a great many things. And because it does so much, there is a lot to learn in order to use it.

We have divided the book into two parts. In this first part, which all users should read, we explain the general ins and outs of using SPSS—from entering data to specifying the desired analysis to examining and manipulating output. These basic steps are a part of any kind of analysis, and users must understand them before they proceed to Part 2. Chapter 1 is a brief overview of the entire process from start to finish. Chapters 2–5 present the major steps in greater detail. In Part 2, we explain the details for conducting a variety of specific statistical procedures, by way of example. After reading Part 1, you will be ready to skip to the chapter in Part 2 that covers the specific type of analysis you want.

Important: Do not be alarmed if what you see on your screen sometimes differs in minor ways from our illustrations in the book. The exact appearance of many aspects of SPSS will vary depending on the particular installation of the program on the computer you are using, as well as the settings for various options that can be changed by system administrators. For example, in your version of SPSS you might see more or fewer icons across the top of your screen than appear in our illustrations, and some pull-down menus might include more or fewer options than ours. When specifying your analyses, you might find that when SPSS provides a list of variables from which to choose, the variables appear in a different order than in our illustrations. When viewing program outputs -- the results of your analyses -- you might find that some tables produced by SPSS are arranged somewhat differently than in our illustrations (e.g., tables that are too wide to fit on your screen might be broken into two tables, one below the other). Also, you might find that numbers in your output tables are rounded off differently than in our version (i.e., reported with more or fewer decimal places); you might even find that very small numbers are reported in scientific notation. As long as you are aware of these possibilities, none of these differences should have any practical effect on your ability to use this book with your particular installation of SPSS.

Chapter 1

Introduction to SPSS for Windows

In this chapter we provide a brief overview of the program and outline the basic procedures for (1) entering your data and naming your variables, (2) specifying statistical analyses, and (3) examining and manipulating your output. Each of these steps is then discussed in detail in Chapters 2–5.

A Quick Tour of the Screen

The SPSS program is started in different ways depending on the installation. Typically, the program is started from Windows by simply double-clicking on the appropriate icon or by choosing it from a menu of options. After the program is finished "loading," your screen should look like Figure 1.1. If SPSS opens in a small window rather than filling the entire screen, click on the *maximize* button in the far upper-right corner to enlarge it.

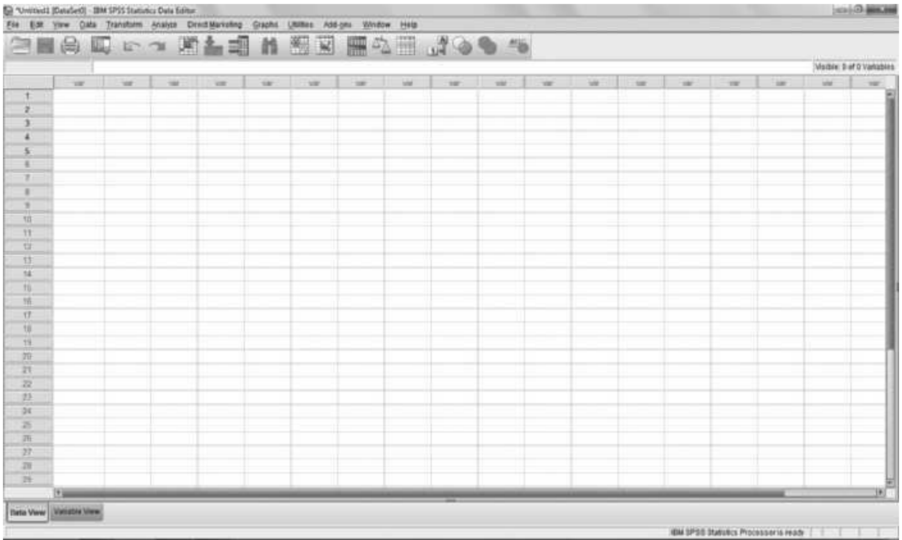


Figure 1.1

Note: Don't be surprised if when SPSS first starts up, the screen shown in Figure 1.1 is partially obscured by a large window containing sections labeled ''New Files,'' ''What's New,'' and so forth. If this happens, just click on the *close* button in the upper-right corner, or on the **Cancel** button in the lower-right corner, to make it go away. You may also be given an option for using Unicode Encoding or Locale Encoding when SPSS first starts up. If you will need to open or use your data files in an earlier version of SPSS (version 16 or earlier), you will want to select Locale Encoding. Otherwise, go ahead and use Unicode Encoding.

Let's take a quick tour of this screen. At the very top is a program title ("Untitled—IBM SPSS Statistics Data Editor"), as well as *minimize*, *restore*, and *close* buttons. (If this window does not occupy the entire screen, a *maximize* button will appear instead of *restore.)*

The second line, which contains the words **File**, **Edit**, and so on, is called the *menu bar*. Clicking on any of these words produces a pull-down menu of options from which to choose in order to accomplish certain tasks. We use several of these pull-down menus in this book.

The third line, containing a row of little pictures or icons, is known as the *tool bar*. These buttons provide shortcuts for tasks otherwise accomplished via the pull-down menus. For example, the button with the little printer on it does the same thing as choosing **File**, then **Print**, from the menu bar. The particular buttons appearing on the tool bar can vary widely. The functions of these buttons are not necessarily self-evident from the little pictures, but you can find out what a particular button does by resting the cursor (using the mouse) on top of it but *not* clicking on it. When you rest the cursor on one of these buttons, a brief phrase appears summarizing the button's function. We don't use the tool bar much in this manual, but you might want to experiment with it on your own.

The remainder (and majority) of the screen is occupied by one or more windows for entering data, writing syntax, displaying output, and so forth. When you first start the program, the window in the foreground (or, alternatively, the only window visible on the screen) is the *Data Editor*.

Using SPSS for Windows: Overview

The Data Editor is typically the starting point for everything you do in SPSS for Windows. It is a spreadsheet-style interface in which you enter your data and specify the names of your variables. Once you have done this, you proceed to the subsequent steps of specifying the analysis you want and examining your results.

Step 1: Data Entry

The first thing you need to do, of course, is enter your data into SPSS and tell the program what these data represent. The easiest way to do this is to use the Data Editor, in which you will (1) enter the data into the rows and columns of the Data Editor, and (2) name your variables. These names will be used to specify the variables in your analyses. Chapter 2 explains these steps in detail.

Step 2: Specifying the Statistical Analysis

Once the data have been entered, the next step is to instruct SPSS to conduct the specific kind of analysis you want. There are two different ways of doing this step; each has advantages and disadvantages.

In the *Point-and-Click Method*, the desired analyses are specified by using the mouse to open menus and dialog boxes and to choose options from them. This is generally the simpler of the two methods because it does not require learning any programming language (syntax). Your computer program is, in effect, written for you, behind the scenes and out of sight, based on your choices. This method should be comfortable and familiar to experienced Windows users, as the interface is designed to be similar to most other Windows programs.

The second method, which we call the *Syntax Method*, involves using SPSS for Windows in a more "old-fashioned" way—much as the various pre-Windows versions of SPSS were used. In this method, you begin by first opening a new window, called a *Syntax Editor*, and then typing commands in the SPSS programming language that specify one or more analyses. Of course, this method requires you to learn a little of the SPSS programming language, which may sound a bit daunting at first. However, there are many advantages to learning this method, not the least of which is that (as you'll see in subsequent chapters) you can do some things in the Syntax Method that you cannot do using the Point-and-Click Method.

Chapter 3 explains the Point-and-Click Method in detail, and Chapter 4 explains the Syntax Method. In Part 2, we explain *both* ways of specifying analyses for each of the statistical procedures covered. You can use one method exclusively, or go back and forth between them whenever you wish.

Step 3: Examining and Manipulating Your Output

Once the data have been entered and the analysis has been specified using one of the two methods just described, a new window appears containing the results of your analyses. You may want to print these results and/or save them for future reference. You might also want to edit the output somewhat (for example, by deleting unneeded parts) before printing or saving. We explain how to accomplish these tasks in Chapter 5.

Chapter 2

Entering Data and Naming Variables

When you enter SPSS to analyze data, the first thing you need to do is type in your data and tell SPSS what the data represent. You can do this most easily using the Data Editor (as illustrated in Figure 1.1). You may want to first expand this window to fill the entire screen (by clicking the *maximize* button in the window's upper-right corner) if it doesn't already do so.

As in other spreadsheet programs, data are entered into a matrix in which *rows* represent individuals or participants (or whatever entities were measured) and *columns* represent different variables (that is, things about those entities that have been measured). One particular cell of the matrix is the "currently selected" one, and is highlighted by a dark border. Initially, this is the upper-leftmost cell (that is, the intersection of row 1 and column 1), as seen in Figure 1.1. You select a new cell by using the arrow keys on the keyboard or clicking the mouse on the desired cell.

To illustrate data entry and naming variables, consider an example in which we have collected data on midterm exam scores for five students. Each student has a value on each of three variables: (1) an identification number, ranging from 1 to 5; (2) sex, coded 1 for males and 2 for females; and (3) midterm exam score. The data are as follows (note that these are the first five cases from the data in Chapter 6):

Student	*Sex*	*Score*
1	1	87
2	1	53
3	1	92
4	1	70
5	1	78

The following two steps—entering the data and naming the variables—can be done in either order.

Entering the Data

Entering data into the Data Editor works just as you would expect. In the upper-left cell (that is, row 1, column 1), type the number **1** to represent the first student's identification number, and then hit the *Enter* key. Now use the right-arrow key to move one cell to the right (or use the mouse to click on the cell to the right), and enter **1** again, this time representing the first student's score on the second variable, sex. Next, move right again, and enter into the cell the first student's exam score of **87**. Now that the first row is completed, use the arrow keys or mouse

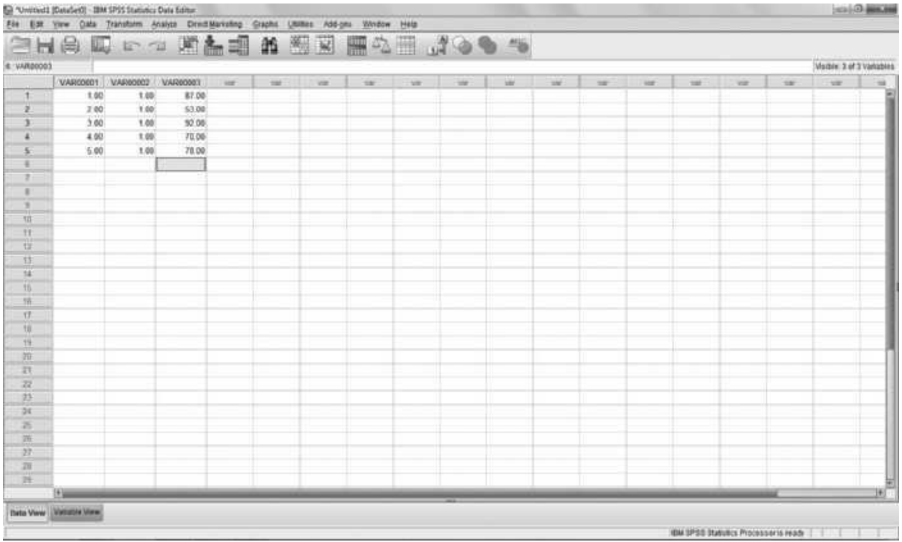


Figure 2.1

to move to the beginning of the second row, and enter the values for the second student in the appropriate columns. Repeat the procedure until all five rows of data have been entered. When you are finished, the screen will look like Figure 2.1.

Types of Data

In this manual we deal only with numerical data—that is, data composed solely of numbers. For various reasons, even categorical variables such as sex (male, female) or group assignment in an experiment (for example, control vs. experimental) are best "coded" (entered) as numbers rather than as letters or words. The most obvious system for coding categories of a variable is simply to assign the numbers 1, 2, 3, and so on to the categories. For the variable *sex* in our example, we assigned the value 1 to men and 2 to women. Of course, this assignment is arbitrary, and we might just as well have done it the other way around. The important thing is that you remember which is which.

If your data contain decimals, just type the decimal point where appropriate using a period. You may enter as many or as few decimal places as necessary.

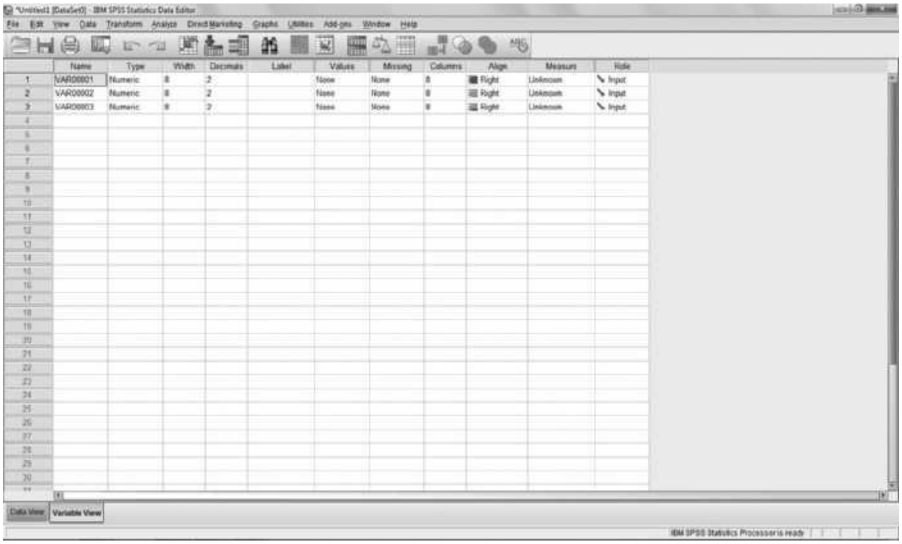


Figure 2.2

Naming the Variables

Whenever you enter data into SPSS, you need to decide what names to assign to your variables and tell SPSS what these names are so you can refer to them later in specifying the analyses you want. Use words or abbreviations that will help you to remember what the variables are. For example, *gender* or *sex* is obviously a good name for a variable that distinguishes men from women. It really makes no difference to SPSS, though, and as far as SPSS is concerned, you could just as well call this variable *cabbage* or *acijd5*. This won't bother SPSS in the least, but it is likely to confuse *you* at some point and should probably be avoided. The important thing is that you assign each of your variables a distinct name and that you remember what these names represent.

In earlier versions of SPSS for Windows (prior to version 12.0), variable names were restricted to no more than 8 characters in length. Although much longer names are now permitted, we still recommend using short, simple variable names for most purposes, as we have done throughout this book.

In our example there are three columns of data representing the students' identification number, sex, and exam score. First, click on the tab labeled "Variable View" in the lower-left corner of the screen (see Figure 2.1). This will reveal a new screen that looks like Figure 2.2. In this screen, each variable is represented by a row containing various bits of information about it. For our current purposes, we wish only to change the names of the variables.

In the first column of the table, you can see that your first variable is named "VAR00001"; this is the name SPSS assigns to the variable by default. This is okay as long as you are able to remember what it represents, but a better idea is to name the variable something more meaningful and easier to remember. To do so, simply click on **VAR00001,** and edit the contents of this box so it contains the name you wish to assign to the variable. The easiest way to do this is to press the *Backspace* key, which deletes the contents of the box, and then type in the variable name you prefer—in this case, "student."

Now use the arrow keys or mouse to move down to the second row to change "VAR0002" to "sex," and then to the third row to change "VAR00003" to "score." Finally, click on the "Data View" tab in the lower-left corner of the screen to return to the screen containing the actual data values.

Once the data have been entered and variables named, it is a good idea to save your work. See Appendix A for details on how to do this.

Chapter 3

Specifying Analyses Using the Point-and-Click Method

Once the data have been entered into the Data Editor, you're ready to tell SPSS what you would like it to do with those data—that is, what kind of statistical analysis you wish to conduct. In Part 2 of this manual we explain exactly how to do this, using both the Point-and-Click Method and the Syntax Method, for a variety of statistical procedures. In this short chapter we briefly preview the Point-and-Click Method; in Chapter 4 we preview the Syntax Method.

To specify analyses using the Point-and-Click Method, you usually begin by clicking on **Analyze** on the menu bar, which produces a pull-down menu listing various categories of statistical procedures from which to choose. Choosing any of these options by clicking on it produces yet another menu of options. For example, clicking on **Compare Means** on the **Analyze** menu produces a list of options, including **One-Sample T Test . . .** and **One-Way ANOVA . . .** , as illustrated in Figure 3.1. (As will be the case for most illustrations in this book, some of the details may differ slightly, depending on your particular installation.) Finally, clicking on one of *these* options produces a

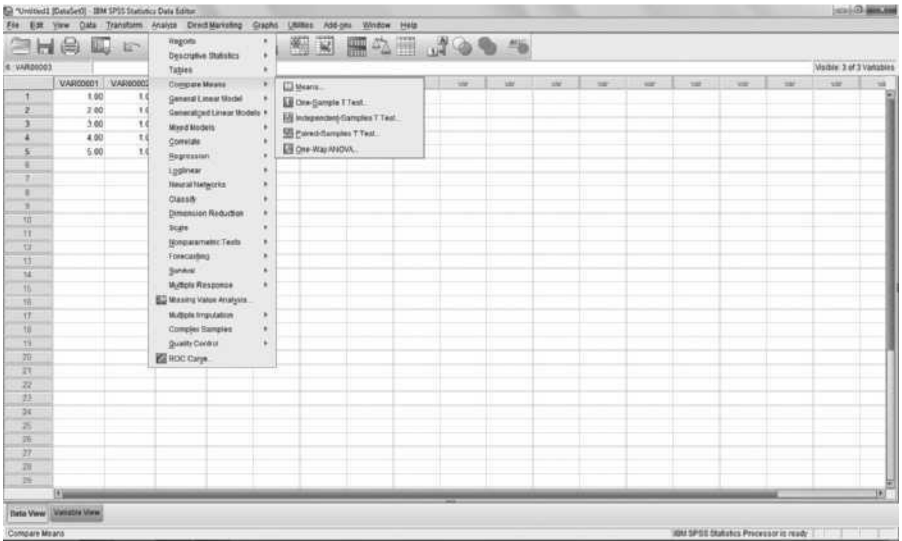


Figure 3.1

dialog box in which you specify the details for the analysis you have selected. In this box you specify such things as the names of the variables to be used in the analysis, details concerning the manner in which the analysis is to be conducted, and choices concerning what information is to be included in the output. In Part 2 we explain the details for different kinds of analyses. Once you have made all your choices, click on a button labeled **OK** and SPSS will spring into action. That's all there is to it.

Although we have described the Point-and-Click and Syntax Methods as mutually exclusive alternatives, a kind of "hybrid" method is also available. After specifying an analysis using the Point-and-Click Method, you can click on a button labeled **Paste** rather than the one labeled **OK**. This tells SPSS to show you the command (syntax) corresponding to the analysis you just specified by pointing-and-clicking—that is, what you would have to type in order to use the Syntax Method. This command is "pasted" into the Syntax Editor, where you can examine it and then run it as if you had typed it in yourself (see Chapter 4). This "hybrid" method can be very helpful in learning how to write SPSS command language.

If you plan to use the Point-and-Click Method exclusively, you can skip the next chapter and proceed to Chapter 5 on viewing and manipulating output. However, we recommend going ahead and learning the Syntax Method, too. As you will see in several chapters in Part 2, for some kinds of analyses the Syntax Method is simpler, or makes it possible to do things that cannot be done with the Point-and-Click Method. It will be worthwhile in the long run to become familiar with both methods.

Chapter 4

Specifying Analyses Using the Syntax Method

To use the Syntax Method, you need a *Syntax Editor* in which to type and edit SPSS program commands ("syntax") to specify the desired analysis. To open the syntax editor, (1) click on **File** on the menu bar, then (2) click on **New** in the resulting pull-down menu, and finally (3) click on **Syntax** in the resulting submenu. The Syntax Editor now appears in the foreground. You may expand it to fill the screen by using the *maximize* button.

At this point, the Syntax Editor window is blank, and a cursor appears in its upper-left corner. If you begin typing, what you type will appear in the window at the point where the cursor is located. As you'll see, this window operates very much like a (very simple) word-processing program, and you can use the *Delete* (or *DEL*) key, the *Insert* (or *INS*) key, the *Backspace* key, and so forth to edit your commands and correct your errors. In the Syntax Method sections of the chapters in Part 2, we explain exactly what to type into this window to specify the desired analysis. Figure 4.1 illustrates the syntax window after a command (from Chapter 6) has been typed in.

The narrow window on the left side of the screen is designed to serve as a table of contents for your syntax, and SPSS will automatically add items to this section as you type commands in the main window. Also, you will notice that as you type commands, the text might suddenly change color and little drop-down menus might sometimes pop up. SPSS is just trying to help you type the correct commands, a feature which you may feel free to ignore or use as you wish.

Executing ("Running") the Analysis

Once your commands have been typed into the Syntax Editor, you need to tell SPSS to "execute" or "run" them—that is, to actually carry out the instructions you have written. First, make sure that the cursor is located somewhere—anywhere—on the command you want to execute by using the mouse and clicking somewhere on the command. Now locate the *Run* button on the tool bar near the top of the screen. It is the button with a right-pointing arrow on it. You can see it in Figure 4.1 next to the icon that looks like binoculars. Click on this button and SPSS will leap into action.

If you typed more than one command into the Syntax Editor, you can execute several commands at once by first highlighting the commands you wish to run. To do this, (1) use the mouse to position the cursor just to the left of the first command, and (2) click (*but do not release*!) the left mouse button; (3) while continuing to hold this button down, use the mouse to reposition the cursor to the end of the last command to be executed; and

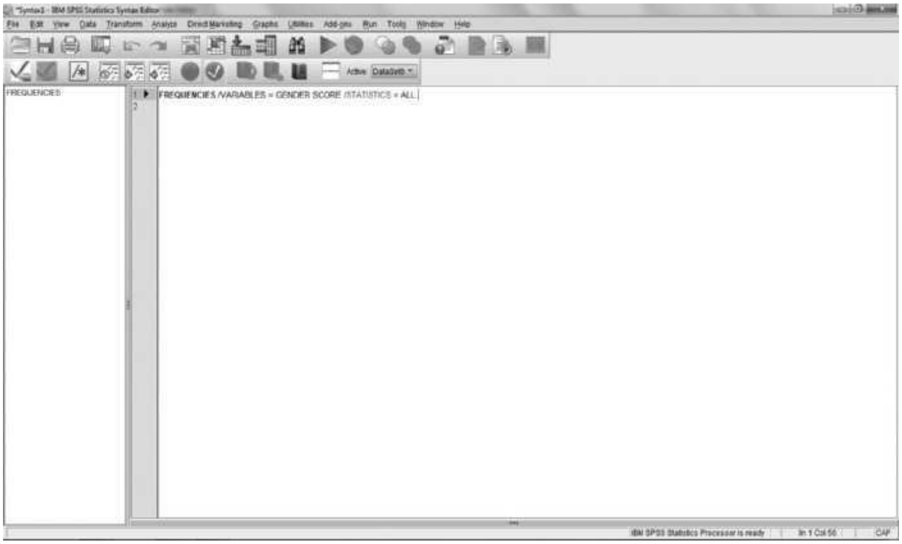


Figure 4.1

(4) release the mouse button. In Windows jargon, this is referred to as "clicking-and-dragging." Then click on the *Run* button to execute the highlighted commands.

Yet another option is available as well. The Syntax Editor has an extra item on its menu bar labeled with the word **Run**. Click on **Run**, and a pull-down menu presents you with several options, including **All** (which runs all the commands appearing in the Syntax Editor without your having to highlight them), **Selection** (which runs only the commands highlighted as just described), and **To End** (which runs all commands beginning at cursor). A final option for running analyses is to hold down the Control key while pressing the letter R on your keyboard.

Some General Notes on SPSS Commands

If you follow our examples carefully and don't improvise too much, you shouldn't have too many problems. However, a few general features of SPSS commands are worth knowing:

1. All SPSS commands should begin in the first column (that is, don't indent them) and must end with a period (**.**). If you ever get an error message (see the following section), the first thing to check is whether you left out a period at the end of a command. No matter what the error message, a

missing period may very well be the culprit. This is an easy mistake to make and, believe us, we both do it all the time.

2. Between the beginning of the command and the final period, there is considerable flexibility in format. For example, wherever a space is required (between words, for example), you can always use two or more spaces if you prefer. In addition, commas and spaces are usually interchangeable.
3. If a command is too long to fit on one line, just hit *Enter* at some point to move to the next line and continue typing the command. Use as many lines as you need or want: SPSS will just keep reading until it finds a period (don't forget that period!) signaling the end of the command. These continuation lines do not have to begin in the first column; in fact, we personally prefer to indent continuation lines a few spaces to make the program look nice and enhance readability.
4. As you will see, many SPSS commands include one or more subcommands that specify details of the analysis. Subcommands are usually separated from each other with slashes (**/**). (Note that these are forward slashes, not backslashes.) Because subcommands are part of the same command, it doesn't matter if you put extra spaces between them, or if you begin subcommands on a new line, or whatever. Again, though, we sometimes like to start a new (indented) line for each subcommand for aesthetic purposes, but this is by no means necessary.

Errors

One of the drawbacks of using the Syntax Method is that you can easily make mistakes in entering data or specifying SPSS commands. This doesn't happen in the Point-and-Click Method, because SPSS writes your commands for you (behind the scenes) based on the choices you specify in menus and dialog boxes—and it knows what it's doing.

Because mistakes are easy to make, when you execute commands from the Syntax Editor, the analysis may not run. As usual, however, SPSS will put the output window in the foreground so you can see your results; in this case, your results will include *error messages*. These are clearly identifiable; they generally say "Error:" or "Warning:" followed by some (usually cryptic) message intended to help you identify the problem. Explaining all of these messages is far beyond the scope of this manual, but if you've followed our examples and instructions carefully, the only kinds of errors you're likely to encounter will involve typographical errors, missing periods at the ends of commands, and so forth. After reading (and trying to decipher) the error messages, get the Syntax Editor back into the foreground (click on **Window** on the menu bar and choose **Syntax1 - SPSS Statistics Syntax Editor** from the resulting menu), and try to figure out what you did wrong. After making any necessary repairs, run the command(s) again as just described.

Once you have written your syntax and confirmed that it works, it is a good idea to save your work. See Appendix A for details on how to do this.

Viewing and Manipulating Output

When you run an analysis using either the Syntax Method or the Point-and-Click Method, your output appears in a new kind of window called the *Statistics Viewer*. This window contains SPSS's report on the results of your analysis, and might look something like Figure 5.1. (This particular screen resulted from the sample problem in Chapter 6.) On the left side of the screen is an outline or table of contents listing the sections of your output; to the right (and taking up most of the screen) is the output itself.

Note: Depending on your installation, you may see some extra text at the top of your output that does not appear in our illustration. For example, if you used the Syntax Method, the SPSS commands you typed into the Syntax Editor might be reprinted at the beginning of the output. Also, if you used the Syntax Method to enter your data, the beginning of the output might include a few lines of text containing information about the format in which

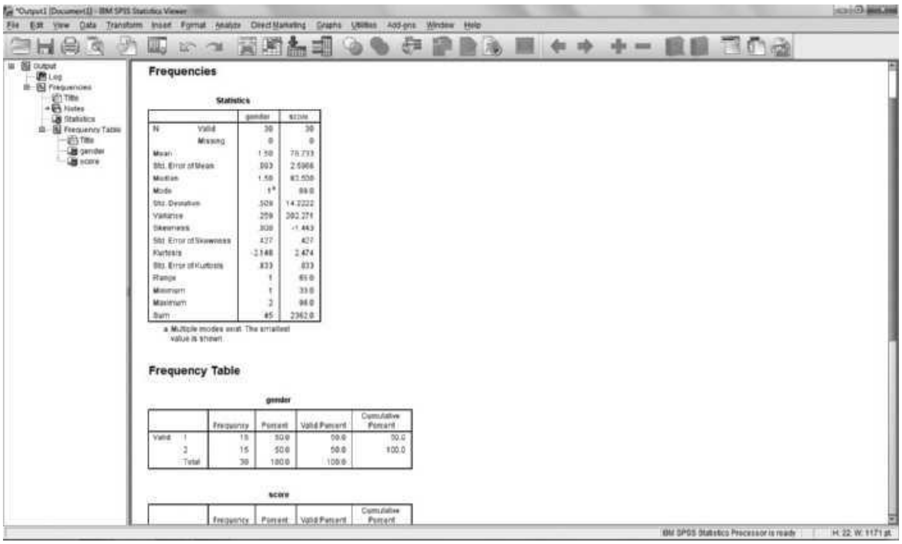


Figure 5.1

the data were read by SPSS. In Figure 5.1, such information would have appeared in the right side of the screen above the title "Frequencies."

Viewing Output

Obviously, the first thing you want to do is examine the results. You can move around within the window using the arrow keys on the keyboard, or use the mouse to manipulate the scroll bars along the window's right and bottom edges. For a different view, try clicking on **File** on the menu bar at the top of the screen, then click on **Print Preview** on the pull-down menu. This gives you a full-screen view of an entire page of output, just as it would appear if you printed it. Experiment with the **Zoom In** and **Zoom Out** buttons on this new screen to change the view to a size you like, and with the **Next Page** and **Prev Page** buttons if the output comprises more than one page. Scroll bars appear on the right and/or bottom edges of the screen when you need them. To return to the main Statistics Viewer screen (Figure 5.1) at any time, click on **Close**.

Editing Output

Before saving and/or printing the output file, you might want to modify it somewhat. Editing the contents of the Statistics Viewer window seems very complicated to us, so for beginning users we recommend not doing much editing of output files.

The one editing task that is relatively simple, and often desirable, is deleting sections of unwanted output. In many procedures, SPSS prints out tables of information you don't really need, so you might want to delete these sections to save paper and unclutter your printout. This is where the outline on the left side of the Statistics Viewer window comes into play. If you click on an item in that outline—say, **Title** or **Statistics**—two things happen: (1) the word you clicked on is selected (and appears highlighted) and (2) the specified section of the output appears in the large window (if it didn't already) surrounded by a box identifying what is in that section. You can delete this entire section by simply hitting the *Delete* key on the keyboard.

Printing Output

Printing the contents of the output window is simple. From the main Statistics Viewer window (Figure 5.1), click on **File** on the menu bar and choose **Print** from the pull-down menu. A dialog box appears somewhere on your screen. First, be sure the box labeled "Copies" contains the number of copies you want printed (usually **1**). Then choose between "All visible output" and "Selected output" by clicking on the circle to the immediate left of the one you choose.

"All visible output" means to print everything displayed in the output window: This is probably the one you want. If you want to print only part of the output, we recommend editing the output in SPSS (see the preceding section) to remove the unwanted portions, and then printing "All visible output." You can also print from the **Print Preview** window by simply clicking on the **Print** button at the top of this screen, and then following the instructions just given.

Saving Output

It is usually a good idea to save your output for future reference, particularly if you don't print out a hard copy. See Appendix A for details on how to do this.

Once you have finished examining, saving, and/or printing your output, you may wish to return to another window to run more procedures, modify your data, and so on. Just click on the *minimize* button in the upper-right corner of the screen, or click on **Window** on the menu bar and choose the window to which you wish to return.

As with the rest of SPSS for Windows, it should be obvious that many other options and features are available in the Statistics Viewer to control the display, editing, and printing of the output, but they are beyond the scope of this manual. Once you feel comfortable with the basics of SPSS, you might want to explore some of these fancier features on your own.

Exiting SPSS for Windows

Once you are finished analyzing your data, reviewing your output, and so forth, you'll be ready to exit SPSS for Windows. This can be done in various ways. One way is to choose **File** from the menu bar and then **Exit** from the pull-down menu; another is to click the *close* button in the very upper-right corner of the screen (the button with the on it). If you haven't saved your work, you will be prompted about whether you wish to save the contents of each of the windows (data, output, and perhaps syntax) that are currently open. If you have already saved these items before exiting, as we suggested (and as detailed in Appendix A), there is no need to do so again. If you forgot to save something, you can click on **Cancel** when asked about saving one of these files, then go back to the appropriate window and save it properly before exiting.

PART 2

Procedures

The preceding chapters provided an overview of the various steps involved in conducting data analysis in SPSS, from data entry to output. In Part 2, we fill in the details for a variety of common statistical procedures, showing how to conduct each analysis using both the Point-and-Click Method and the Syntax Method. For each type of analysis, we present a sample problem, show how to conduct the analysis using each of the two methods, and then explain the output produced by the procedure.

In most cases, the two methods produce exactly the same results. However, in some cases, certain useful options are available in the Syntax Method that are not available in the Point-and-Click Method. In some of these cases, extra steps are required in the Point-and-Click Method to obtain comparable results. The format of the output differs somewhat depending on which method you used, but they should be sufficiently similar that you should have no trouble understanding your output if it looks a bit different from ours.

Important: Remember that, as we noted at the beginning of Part 1, what you see on your screen may sometimes differ in minor ways from our illustrations in the book. (See p. 1 for some examples of the kinds of differences you might encounter.) Again, any such differences should not cause you any problems as long as you know to expect them.

Chapter 6

Frequency Distributions and Descriptive Statistics

Sample Problem

Fifteen men and fifteen women in an introductory psychological statistics course have taken their midterm exams. Their scores are listed here. (For gender, 1 = male and 2 = female.)

Student	*Gender*	*Score*	*Student*	*Gender*	*Score*
1	1	87	16	2	89
2	1	53	17	2	73
3	1	92	18	2	91
4	1	70	19	2	76
5	1	78	20	2	75
6	1	73	21	2	89
7	1	91	22	2	81
8	1	60	23	2	83
9	1	77	24	2	68
10	1	82	25	2	86
11	1	85	26	2	55
12	1	33	27	2	89
13	1	88	28	2	89
14	1	98	29	2	70
15	1	88	30	2	93

In this problem we wish to construct frequency distributions and obtain some basic descriptive statistics for the variables *gender* and *score.* Refer back to Figure 2.1 (page 6) to see what the screen will look like after you have entered the first five lines of these data.

Analysis

Following the procedure outlined in Chapter 2, enter the data into the first three columns of the Data Editor and label the variables **Student**, **Gender**, and **Score**.

Point-and-Click Method

To begin, click on **Analyze** on the menu bar. This produces a pull-down menu containing a variety of options, such as **Compare Means**, **Correlate**, and so on. Now click on **Descriptive Statistics**. This produces yet another

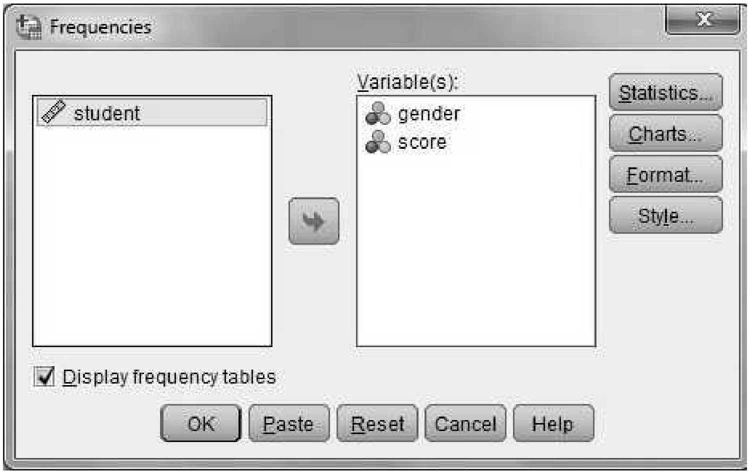


Figure 6.1

menu containing items such as **Frequencies . . .** , **Descriptives . . .** , and **Explore . . . .** Choose **Frequencies . . .** to specify that you want frequency distributions. This produces a dialog box somewhere on your screen, on top of the other windows, that looks like Figure 6.1.

Many of the dialog boxes for specifying procedures in SPSS for Windows are similar to this dialog box in several respects. On the left is a box containing the names of all your variables. To the right is another (empty) box labeled "Variable(s)." The goal is to move from the left box to the right box the names of the variable(s) for which you want frequency distributions. Do this by clicking on one of the variable names to highlight it, and then clicking on the right-arrow button between the boxes to move it. The variable name disappears from the left box and appears in the right box. (Alternatively, you can just double-click on the variable name and it will move to the other box immediately.) Simply repeat this procedure for each variable desired. If you make a mistake or change your mind, you can single-click on a variable in the right-hand box, and then click on the middle arrow (which will have switched to a left-pointing arrow when you clicked on the variable name) to remove it from the "Variable(s)" list. In this example, we want to choose **Gender** and **Score**, the two variables of interest. You may now click on **OK** to run the analysis. However, if you would like to have SPSS print out descriptive statistics or graphs in addition to the frequency distribution, do not click on **OK** but instead keep reading.

Descriptive Statistics. At the bottom of this dialog box, you'll see three other buttons labeled **Statistics . . .** , **Charts . . .** , and **Format. . . .** If you wish to

Figure 6.2

obtain descriptive statistics along with your frequency table, click on **Statistics . . .** to produce a new dialog box, as illustrated in Figure 6.2.

Apart from the section labeled "Percentile Values" and the line "Values are group midpoints," the remaining sections of the box list a variety of descriptive statistics. Simply click on the little box to the left of each statistic you wish SPSS to calculate for you; a check mark will appear in each selected box. (Note that if you wish to unselect a box you've already selected, simply click on it again.) The sample output shown in the next main section, titled Output, is the result of requesting all available statistics by checking all six boxes under "Dispersion," all four boxes under "Central Tendency," and both boxes under "Distribution." The box next to "Values are group midpoints" does not need to be selected, however. When finished, click on **Continue** to return to the previous dialog box (Figure 6.1). Click on **OK** to run the analysis, or follow the instructions in the next section to obtain a graphical frequency distribution in the form of a histogram or bar graph as well.

Histograms and Bar Graphs. To obtain histograms and/or bar charts in addition to frequency tables (and in addition to or instead of descriptive statistics), click on **Charts . . .** at the right in the dialog box illustrated in Figure 6.1 to produce a new dialog box (not illustrated). Initially, SPSS assumes you do not want any graphs ("charts")—the little circle next to "None" is selected by default. Click on the circle next to “Bar chart” or “Histogram” to indicate the kind of frequency distribution graph you prefer. Then click **Continue** to return to the previous dialog box (Figure 6.1), and click on **OK** to run the analysis.

Syntax Method

Open the Syntax Editor as described in Chapter 4. Type the following command (don’t forget the period at the end!) in this window and then click on the *Run* button to execute the analysis:

```
FREQUENCIES /VARIABLES = GENDER SCORE /STATISTICS = ALL.
```

This command tells SPSS to produce frequency distributions. The subcommand **/VARIABLES** specifies the variables for which you would like SPSS to produce a frequency distribution, in this case **GENDER** and **SCORE**. Note that variable names are separated by spaces or commas. Alternatively, you can write **/VARIABLES = ALL**, replacing the list of variable names with the word **ALL**. This tells SPSS to create a frequency distribution for every variable it knows about.

Descriptive Statistics. The subcommand **/STATISTICS = ALL** tells SPSS to produce all available summary statistics for the variables listed on the **/VARIABLES** subcommand. If you want only certain statistics, you need to consult a more detailed manual to determine how to replace **ALL** with names of specific statistics. (The more common ones are **MEAN**, **MEDIAN**, **STDDEV**, and **VARIANCE**.) However, it’s usually easier to just ask for **ALL** statistics and then ignore the ones you don’t care about. If you do not want any descriptive statistics, simply leave off the **/STATISTICS** subcommand entirely.

Histograms and Bar Graphs. To obtain histograms in addition to the tabular frequency distribution, simply add a **/HISTOGRAM** subcommand to the **FREQUENCIES** command, either in addition to or instead of a **/STATISTICS** subcommand. To request a histogram for the variable **SCORE**, for example, use this command:

```
FREQUENCIES /VARIABLES = SCORE /HISTOGRAM.
```

If you wish to produce a bar chart instead of a histogram, replace the subcommand **/HISTOGRAM** with the subcommand **/BARCHART**.

Output

The output produced by SPSS for the sample problem is illustrated in Figure 6.3.

Frequencies

Statistics

		gender	score
N	Valid	30	30
	Missing	0	0
Mean		1.50	78.73
Std. Error of Mean		.093	2.597
Median		1.50	82.50
Mode		1[a]	89
Std. Deviation		.509	14.222
Variance		.259	202.271
Skewness		.000	-1.443
Std. Error of Skewness		.427	.427
Kurtosis		-2.148	2.474
Std. Error of Kurtosis		.833	.833
Range		1	65
Minimum		1	33
Maximum		2	98
Sum		45	2362

a. Multiple modes exist. The smallest value is shown

Frequency Table

gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	15	50.0	50.0	50.0
	2	15	50.0	50.0	100.0
	Total	30	100.0	100.0	

Figure 6.3

For each variable, SPSS produces a frequency distribution and a set of descriptive statistics; we have included only the frequency table for the variable *gender* here due to space considerations. The table of statistics should be self-explanatory. The first column in each Frequency Table lists all values that were found on the variable; the second column ("Frequency") lists

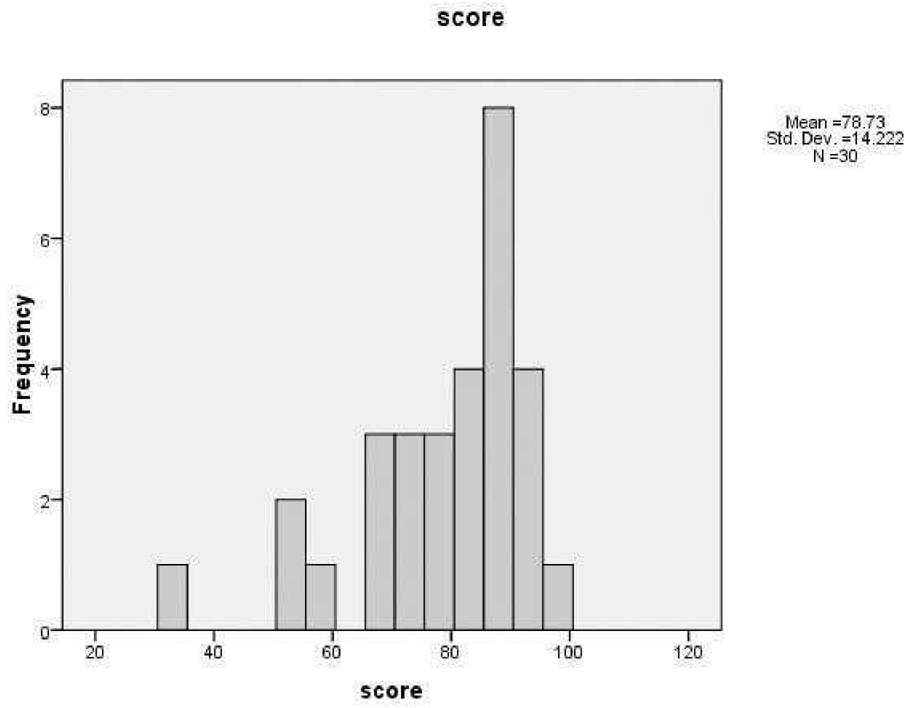


Figure 6.4

the number of cases having that value; the third column ("Percent") is the percentage of all cases having this value (that is, frequency divided by total number of cases); the fourth column ("Valid Percent") is the percentage of all "valid" cases—that is, cases that were not missing data on this variable—having that value; and the fifth column lists the cumulative percentage for that value—that is, the percentage of valid cases having that value or less. Note that "Percent" and "Valid Percent" will be identical if there are no missing values on this variable.

Histograms and Bar Graphs. Figure 6.4 illustrates a histogram for the variable *score*, as produced by the procedure described earlier. Notice that SPSS has divided the range of scores into 5-point intervals (e.g., 30–35, 35–40, and so on), and the heights of the respective bars represent the frequencies of these categories rather than of individual scores. Had you asked SPSS for a *bar chart* instead of a *histogram*, SPSS would not have grouped the scores into intervals like this; instead, each individual value would have its own separate bar representing the frequency of that particular score. For this reason, it is generally preferable to use bar charts for variables that have only a small number of possible values (especially categorical or "nominal-level" variables, such as *gender*), and to use histograms for variables that can take on a large or infinite number of possible values (especially "interval-level" and "ratio-level" variables).

Chapter 7

One-Sample *t*-Test

Sample Problem

The superintendent of the Riverwalk school district claims that the students in his district are more intelligent, on the average, than the general population of students. The mean IQ of the general population of school children is 105. A study was conducted to determine the IQs of a sample of school children in the Riverwalk district. The results are as follows:

Student	*IQ Score*	*Student*	*IQ Score*
1	110	16	110
2	105	17	117
3	102	18	98
4	112	19	124
5	120	20	107
6	107	21	112
7	99	22	122
8	100	23	104
9	109	24	105
10	103	25	110
11	115	26	120
12	125	27	125
13	115	28	120
14	106	29	100
15	110	30	110

In this problem, we are testing the null hypothesis that the mean IQ of all school children in the Riverwalk school district equals 105. This will be assessed by examining a random sample of Riverwalk students. This value was selected because it was already known that 105 is the overall mean IQ of all students in all school districts. We want to know whether the Riverwalk population is any different. Is the mean of this sample significantly different from 105?

Analysis

Following the procedure outlined in Chapter 2, enter the data into the first two columns of the Data Editor and label the variables **student** and **iq**.

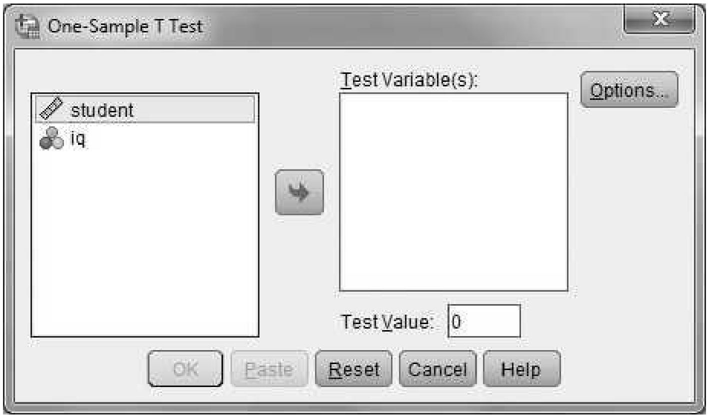


Figure 7.1

Point-and-Click Method

Click on **Analyze** on the menu bar, and then choose **Compare Means**. From the resulting menu (which was illustrated in Figure 3.1), choose **One-Sample T Test . . .** . This produces a dialog box that looks like Figure 7.1.

In the left-hand box, click on **iq,** then click on the right-arrow button between the boxes to move **iq** to the box labeled "Test Variable(s)." Next, click in the box labeled "Test Value" and edit its contents (which appear as **0** initially) so it reads **105**. This is the null hypothesis value against which you are testing the sample mean. (Of course, you would just leave this box as **0** if you wanted to test the hypothesis that the mean IQ equals zero.) Now click on **OK** and the analysis will run.

Syntax Method

Open the Syntax Editor, type the following command (don't forget the period at the end!), and click on the *Run* button to execute the analysis:

```
T-TEST TESTVAL = 105 /VARIABLES = IQ.
```

The syntax for this command, as you can see, is quite simple. The meaning of the command **T-TEST**, which is also used in the next two chapters, is self-evident. The subcommand **TESTVAL =** specifies the null hypothesis value to be tested; in this case, the value of interest is 105. The **/VARIABLES** subcommand then specifies the variables to be tested. In this case, we want to know only if the mean of the variable **IQ** is significantly different from 105. If we had several variables, and we wanted to test the mean of each (separately) against the null hypothesis value of 105, we would list the other variable names after **IQ**, separated by spaces or commas.

Output

The output produced by SPSS for the sample problem is shown in Figure 7.2.

SPSS first prints some descriptive statistics for the variable *iq*, including number of cases, mean, standard deviation, and standard error. Beneath the descriptive statistics are the results of the significance test. "Mean Difference" is the difference between the observed sample mean (110.73) and the hypothesized mean (105). SPSS also prints out a 95% confidence interval for the difference between means, which in this case goes from 2.73 to 8.74.

Is a mean difference of 5.733 large enough to be significantly different from 105? The results of the *t*-test show that $t = 3.900$, with 29 ($N - 1$) degrees of freedom ("df"). The two-tailed *p*-value for this result is .001 (rounded off to three decimal places). The result is considered statistically significant if the *p*-value is less than the chosen alpha level (usually .05 or .01). In this case, *p* is definitely less than .05 (and .01), so the result is considered statistically significant and the null hypothesis is rejected.

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
iq	30	110.73	8.051	1.470

One-Sample Test

	Test Value = 105					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
iq	3.900	29	.001	5.733	2.73	8.74

Figure 7.2

If you were doing the problem by hand, you would use a table in your statistics textbook to determine the critical *t*-value associated with 29 degrees of freedom—this value is 2.045 (for alpha = .05)—and then compare the observed *t*-value to the critical value. In this case, the observed *t* of 3.90 is greater than the critical value, so again you would have rejected the null hypothesis.

Note that the output from most statistical procedures in SPSS works this same way. That is, SPSS prints the *p*-value associated with the test statistic, and it is up to you to decide whether *p* is sufficiently small to reject the null hypothesis (that is, whether *p* is smaller than your chosen alpha level of .05, .01, or whatever). Both methods—comparing critical versus observed *t*-values and comparing critical (alpha) versus observed (*p*) probabilities—lead to the exact same conclusion for any given problem.

Independent-Samples *t*-Test

Sample Problem

The psychology department at Rockhaven University conducted a study to determine the effectiveness of an integrated statistics/experimental methods course as opposed to the traditional method of taking the two courses separately. It was hypothesized that the students taking the integrated course would conduct better quality (more controlled and statistically sound) research projects than students in the traditional courses as a result of their integrated training. To determine whether there actually was a difference in student performance as a result of integrated versus traditional training, the final research projects of 20 students from an integrated course and 20 students from a traditional course were evaluated. Their scores are listed here. (For condition, 1 = integrated course and 2 = traditional course.)

Student	*Condition*	*Score*	*Student*	*Condition*	*Score*
1	1	87	21	2	82
2	1	95	22	2	72
3	1	89	23	2	95
4	1	74	24	2	60
5	1	73	25	2	90
6	1	92	26	2	87
7	1	63	27	2	89
8	1	90	28	2	86
9	1	94	29	2	76
10	1	84	30	2	74
11	1	91	31	2	85
12	1	90	32	2	75
13	1	75	33	2	90
14	1	93	34	2	91
15	1	87	35	2	88
16	1	85	36	2	63
17	1	90	37	2	70
18	1	89	38	2	72
19	1	87	39	2	84
20	1	85	40	2	60

In this problem, we are testing the null hypothesis that there is no difference in student performance as a result of the integrated versus traditional courses; that is, the mean difference between the conditions in the population from which the sample was drawn is zero. The alternative hypothesis reflects the psychology department's belief that the population means for the two groups of students are not equal (that is, the belief that course format had some kind of effect on quality of research projects).

Analysis

Following the procedure outlined in Chapter 2, enter the data into the first three columns of the Data Editor and label the variables **student**, **cond**, and **score**.

Point-and-Click Method

Click on **Analyze** on the menu bar, and then choose **Compare Means**. From the resulting menu, choose **Independent-Samples T Test . . . .** This produces a dialog box that looks like Figure 8.1.

In this dialog box, your list of variables appears in the box to the left, and you must (1) move one (or more) of the variables into the box labeled "Test Variable(s)" to select your dependent variable(s) and (2) move *one* of your variables into the box labeled "Grouping Variable" to identify the groups to be compared (that is, to select the independent variable). First, click on **score** (the dependent variable in our example) in the left-hand box to select it; then click on the upper right-arrow button pointing to the "Test Variable(s)" box; **score** disappears from the left-hand box and reappears under "Test Variable(s)." Next, click on **cond** (the independent variable in our example) to select it, and

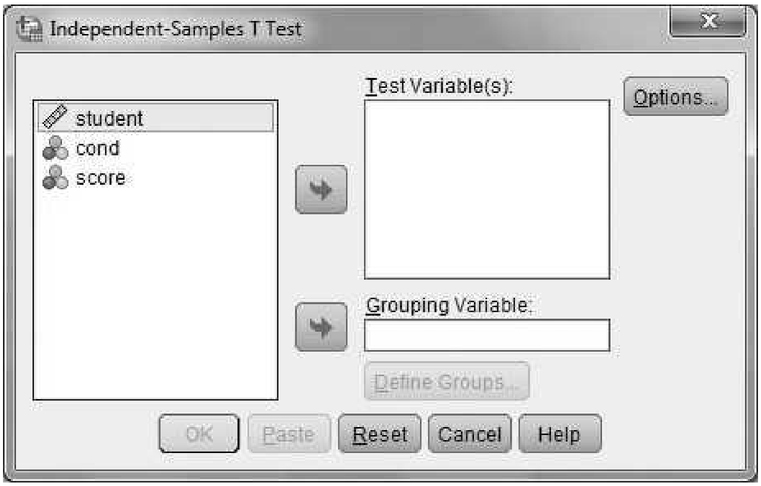


Figure 8.1

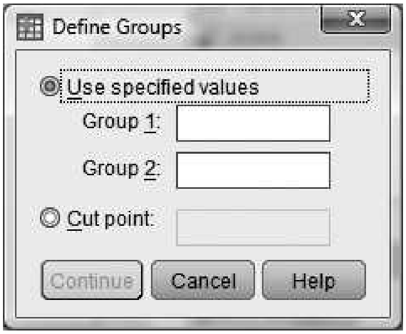


Figure 8.2

then click on the right-arrow button pointing to the "Grouping Variable" box to move it there. The name **cond** now appears under "Grouping Variable," followed by a set of parentheses containing two question marks. This is to call your attention to the fact that one additional specification is required before you can execute this analysis.

When you selected **cond** as your grouping variable, something else happened on your screen as well. The button labeled **Define Groups . . .** suddenly looked different. Whereas it previously appeared fuzzy and with lightly colored lettering, it now appears sharp and distinct. This is because the button was not functional until you selected the grouping variable—but it is now functional and quite important. Click on it and another dialog box appears (see Figure 8.2) in which you must specify the two values of **cond** that represent the two groups you wish to compare. In our case, **cond** was coded simply as **1** (integrated course) and **2** (traditional course). Click in the box next to "Group 1" and, when the cursor appears there, type the number **1**. Then use the mouse to click in the box next to "Group 2" and type the number **2**. Now click on **Continue** to return to the dialog box illustrated in Figure 8.1. In this box, click on **OK** to run the analysis.

Syntax Method

Open the Syntax Editor, type the following command (don't forget the period at the end!), and then click on the *Run* button to execute the analysis:

```
T-TEST /GROUPS = COND (1,2) /VARIABLES = SCORE.
```

To use the **T-TEST** command in SPSS to test a hypothesis about the equality of means of two independent populations, the subcommand **/GROUPS** is used to name the independent variable whose values identify the two groups being compared (in this problem, **COND**). After the independent variable name, you must type, in parentheses, the two values of this variable that identify the two groups you wish to compare (in this problem, **1** and **2**). The two values must be separated by a comma.

The subcommand **/VARIABLES** is used to identify the dependent variable(s) (in this problem, **SCORE**) whose means you wish to compare between groups. You can specify several variable names, separated by commas or spaces, to specify multiple dependent variables on the same command.

Output

The output produced by SPSS for the sample problem is shown in Figure 8.3.

SPSS first prints the number of cases, means, standard deviations, and standard errors on the dependent variable separately for each group. In this case, the two groups are defined by the variable *cond* (1 = integrated; 2 = traditional).

"Levene's Test for Equality of Variances" is provided next by SPSS. This test is probably not of interest to most readers. (It represents a test of the hypothesis that the populations from which the groups were sampled have equal variances.)

Beneath (or to the right of) Levene's Test are the results of the "*t*-test for Equality of Means." The information provided in the row labeled "Equal variances not assumed" reports the results of a *t*-test that is sometimes used when there is reason to believe that the two population variances are not equal. SPSS reports the observed *t*-value, the degrees of freedom ("df"), and the two-tailed *p*-value ("Sig. (2-tailed)"). This test is not often discussed in introductory statistics courses, so if you haven't discussed the test in class, just ignore this part of the output. Also reported on this line are the difference between the means, standard error of the difference, and the 95% confidence interval for the difference between population means.

The most commonly used test is the one listed in the row labeled "Equal variances assumed." Because we are assuming that the two population variances are equal, a pooled variance estimate is used to combine the two sample variances to obtain the most accurate estimate of the variance common to both populations.

The observed *t*-value for this problem is 2.043, with degrees of freedom (total sample size minus 2) equal to 38. The two-tailed probability of .048 is less than .05 and, therefore, the test is considered significant (though barely) at the .05 level.

To verify this, you can use your statistics textbook to determine the critical *t*-value associated with 38 degrees of freedom: critical $t_{.05}(38)$ = approximately 2.03. The observed *t*-value (2.04) is greater (barely) than the critical *t*-value; therefore, the null hypothesis is rejected at the .05 level of significance.

T-Test

Group Statistics

	cond	N	Mean	Std. Deviation	Std. Error Mean
score	1	20	85.65	8.242	1.843
	2	20	79.45	10.782	2.411

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
score	Equal variances assumed	3.880	.056	2.043	38	.048	6.200	3.035	.057	12.343
	Equal variances not assumed			2.043	35.551	.049	6.200	3.035	.043	12.357

Figure 8.3

Chapter 9

Dependent-Samples (Matched-Pairs, Paired Samples) *t*-Test

Sample Problem

An investigator believes (based on past research) that parents who use positive verbal statements (polite requests and suggestions) have children who are more socially accepted and more positive in interactions with peers. Although children acquire behavioral information from sources other than parents (TV, peers, and so on), more induction (coaching children by introducing consequences for behaviors and supplying rationales that support them) on the part of parents, as opposed to more power-assertive and permissive types of discipline, facilitates a prosocial behavioral orientation in children that leads to greater social competence and greater acceptance by peers. Twenty first-grade children (who were rated by teachers and peers as aggressive) and their parents were selected for a study to determine whether a seminar instructing on inductive parenting techniques improves social competency in children. The parents attended the seminar for one month. The children were tested for social competency before the course began and were retested six months after the completion of the course. The results of the social competency test are shown here (with higher scores indicating greater social competency):

Child	*Pre*	*Post*	*Child*	*Pre*	*Post*
1	31	34	11	31	28
2	26	25	12	27	32
3	32	38	13	25	25
4	38	36	14	28	30
5	29	29	15	32	41
6	34	41	16	27	37
7	24	26	17	37	39
8	35	42	18	29	33
9	30	36	19	31	40
10	36	44	20	27	28

In this problem, we are testing the null hypothesis that there is no difference between the means of pre- and postseminar social competency scores. In other words, the parenting seminar has no effect on child social competency scores; or, stated yet another way, the population mean of difference scores (preseminar minus postseminar scores, or vice versa) equals zero. The alternative hypothesis reflects the investigator's belief that there is a difference between pairs of pre- and postseminar scores; the population mean of difference scores does not equal zero (that is, the seminar has some effect on social competency scores).

Analysis

Following the procedure outlined in Chapter 2, enter the data into the first three columns of the Data Editor and label the variables **child**, **pre**, and **post**.

Point-and-Click Method

Click on **Analyze** on the menu bar, and then choose **Compare Means**. From the resulting menu (which was illustrated in Figure 3.1), choose **Paired-Samples T Test . . . .** This produces a dialog box that looks like Figure 9.1.

In this dialog box, your list of variables appears in the box in the upper left, and you identify as follows the ones whose means you want to compare. First, click on **pre**, then on the arrow key between the boxes; this will move the first variable over to the box on the right. Then click on **post,** then the arrow, to move the second variable. (Alternatively, if you hold down the *Shift* key while clicking on **pre** and then **post,** they will both be highlighted; clicking on the arrow next will move them both together.) If you make a mistake, just click on the variable name in the right-hand box, and the arrow will switch to point in the other direction. Clicking on it then will move the variable back to the list on the left.

When SPSS computes difference scores, it computes the first variable ("variable 1") minus the second variable ("variable 2"). In this example, the difference scores will be computed as **pre** minus **post.** This can be confusing, as positive differences will mean that scores decreased from pretest to posttest, and negative differences will mean they increased. To avoid this confusion, you can move the variables to the right-hand box in the opposite order, so that differences are computed as **post** minus **pre.** In either case, the significance test tests the hypothesis that in the population, the mean of these difference scores equals zero. Now click on **OK** to execute the analysis.

Figure 9.1

Syntax Method

Open the Syntax Editor, type the following command (don't forget the period at the end!), and then click on the *Run* button to execute the analysis:

```
T-TEST /PAIRS = PRE POST.
```

As usual, the **T-TEST** command in SPSS is used to test the null hypothesis that two population means are equal.

The **/PAIRS** subcommand tells SPSS that you wish to conduct a paired *t*-test to test whether the means of two variables are equal. After the **/PAIRS** subcommand, you must specify the names of the two variables whose means you wish to compare, separated by a space or comma. Difference scores will be computed by subtracting the first variable minus the second—in this case, **pre** minus **post**—so choose the order you prefer.

Output

The output produced by SPSS for the sample problem is shown in Figure 9.2.

SPSS first produces some descriptive statistics (number of pairs, mean, standard deviation, and standard error) for each variable (in this case, *pre* and *post*). In addition, SPSS also calculates the correlation coefficient between the two variables *pre* and *post*, and the two-tailed probability for a test of the null hypothesis that the population correlation coefficient equals zero. These results appear under "Correlation" and "Sig" in the middle section of the output. This test is probably not of interest to you at this point. A more direct way of calculating correlation coefficients is presented in Chapter 14.

A paired *t*-test is calculated by first computing a set of difference scores in which one variable is subtracted from the other. The mean of these difference scores is given under "Paired Differences" and is equivalent to the difference between the mean of *pre* and the mean of *post*. The standard deviation and standard error of these difference scores are listed next, followed by the 95% confidence interval for the population mean of difference (*pre* – *post*) scores.

Finally, the results of the *t*-test itself are presented. The observed *t*-value, calculated as the mean difference (–3.750) divided by its standard error (.876), is –4.280. The degrees of freedom (number of pairs of observations minus 1) and the two-tailed *p*-value are also printed. Note that although the computed *p*-value for this problem is reported as ".000," this does not mean that it is actually equal to zero. SPSS rounds off *p*-values (in this case, to three decimal places). Therefore, any *p*-value less than .0005 is printed as .000.

To check this, you can use your statistics textbook to determine the critical *t*-value associated with 19 degrees of freedom and an alpha level of .001 (this is the smallest level of significance listed in most textbook tables) for a two-tailed test: critical $t_{.001}(19) = 3.883$. The observed *t*-value is greater than the critical *t*-value; therefore, the null hypothesis is rejected at the .001 level of significance.

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pre	30.45	20	4.019	.899
	post	34.20	20	6.066	1.356

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	pre & post	20	.771	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	pre - post	-3.750	3.919	.876	-5.584	-1.916	-4.280	19	.000

Figure 9.2

Chapter 10

One-Way Between-Groups (Independent Groups) ANOVA

Sample Problem

Sleep researchers have postulated that, in a broad evolutionary sense, sleeping more poorly during periods of perceived environmental threat may have survival value. Previous research has indicated that individuals experiencing anxiety or stress exhibit reduced periods of deep sleep and increased periods of light sleep (because one can most easily be aroused by a sound in the environment while in light sleep). An attachment researcher conducted a study to examine the effects of anxious, avoidant, and secure attachment styles on the physiology of sleep. The investigator hypothesized that children with anxious (and perhaps avoidant) attachment styles experience more sleep disturbances than children with secure attachment styles because they feel responsible for monitoring the external environment and regulating the distance between themselves and their caregivers. These children may find it difficult to sleep when their caregivers are physically absent and may develop patterns of light sleep because of the need to be aware of the caregiver's presence at all times. Deep sleep can be experienced as threatening to the attachment bond and thus dangerous to the child. The sleep patterns of 10 secure, 10 anxious, and 10 avoidant 5-year-old children were monitored. Of primary importance to the attachment researcher was the overall percentage of time that each child spent in deep (delta) sleep. It was hypothesized that children who are insecurely attached to their primary caregivers will spend a lower percentage of time in deep (delta) sleep as compared to their secure counterparts. Following is the average amount of time that each child spent in delta sleep, expressed as a percentage of total sleep time. (For the attachment styles, 1 = secure, 2 = anxious, and 3 = avoidant.)

Participant	*Attachment Style*	*Delta Sleep*	*Participant*	*Attachment Style*	*Delta Sleep*
1	1	21	11	2	17
2	1	21	12	2	17
3	1	25	13	2	15
4	1	23	14	2	15
5	1	24	15	2	15
6	1	23	16	2	14
7	1	23	17	2	20
8	1	22	18	2	13
9	1	22	19	2	14
10	1	22	20	2	19

Participant	*Attachment Style*	*Delta Sleep*	*Participant*	*Attachment Style*	*Delta Sleep*
21	3	18	26	3	17
22	3	20	27	3	15
23	3	18	28	3	16
24	3	19	29	3	17
25	3	17	30	3	18

In this problem, we are testing the null hypothesis that, on average, the three populations of children (secure, anxious, and avoidant attachment styles) spend an equal percentage of time in deep (delta) sleep:

$$H_0: \begin{array}{c}\text{Mean delta for the}\\ \text{population of}\\ \text{secure children}\end{array} = \begin{array}{c}\text{Mean delta for the}\\ \text{population of}\\ \text{anxious children}\end{array} = \begin{array}{c}\text{Mean delta for the}\\ \text{population of}\\ \text{avoidant children}\end{array}$$

The test of this hypothesis, sometimes referred to as the "omnibus" null hypothesis, is conducted as an F-test.

Post-Hoc Test

Rejection of the "omnibus" null hypothesis in ANOVA tells us only that not all population means are equal. It does not indicate which attachment groups are significantly different from which others. One way to further examine group differences is to use *post-hoc* or *a posteriori* multiple comparison tests, such as the Student-Newman-Keuls or Tukey HSD test. In our example, we use Tukey HSD tests to test for differences between all possible pairs of means.

Planned Comparisons/Contrasts

Another approach to multiple comparisons in ANOVA is to test more specific hypotheses, which have been specified in advance, about which means differ from which other means or which combinations of means differ from other combinations of means. These tests are known as *planned comparisons*, *contrasts*, or *a priori* multiple comparison tests.

In our example, we test two contrasts. For each contrast, we are testing a different null hypothesis. First, we want to test the null hypothesis that the mean percentage of time spent in delta sleep for the population of secure children is equal to the mean of the means of the populations of anxious and avoidant children. Mathematically,

$$H_0: \begin{array}{c}\text{Mean delta for the}\\ \text{population of}\\ \text{secure children}\end{array} = \frac{\begin{array}{c}\text{Mean delta for the}\\ \text{population of}\\ \text{anxious children}\end{array} + \begin{array}{c}\text{Mean delta for the}\\ \text{population of}\\ \text{avoidant children}\end{array}}{2}$$

The appropriate contrast coefficients for this problem are **–2** for the secure group (attstyle = 1), **1** for the anxious group (attstyle = 2), and **1** for the avoidant group (attstyle = 3). Note that we could just as easily have reversed all the signs and used the coefficients 2, –1, and –1, respectively.

For the second contrast, we want to test the null hypothesis that the mean percentage of time spent in delta sleep for the population of anxious children is equal to the mean percentage of time spent in delta sleep for the population of avoidant children:

$$H_0: \begin{matrix}\text{Mean delta for the}\\ \text{population of}\\ \text{anxious children}\end{matrix} = \begin{matrix}\text{Mean delta for the}\\ \text{population of}\\ \text{avoidant children}\end{matrix}$$

The appropriate coefficients for attstyles 1 through 3, respectively, are 0, 1, and –1.

In this example, we are testing only two contrasts, and they happen to be orthogonal. However, it is not necessary for contrasts specified in SPSS to be orthogonal, and there is no limit on the number you may specify in a given analysis.

Analysis

Following the procedure outlined in Chapter 2, enter the data into the first three columns of the Data Editor and label the variables **participant**, **attstyle**, and **delta**.

Point-and-Click Method

Click on **Analyze** on the menu bar, and then choose **Compare Means**. From the resulting menu, choose **One-Way ANOVA . . . .** This produces a dialog box that looks like Figure 10.1.

This dialog box is very similar to that for the independent-samples *t*-test, as discussed in Chapter 8. In this dialog box, your list of variables appears in the box to the left, and you must (1) move one (or more) of your variables into the box labeled "Dependent List" to select your dependent variable(s) and (2) move *one* of your variables into the box labeled "Factor" to identify the groups to be compared (that is, to select the independent variable). First, click on **delta** (the dependent variable in our example) in the left-hand box to select it; then click on the upper right-arrow pointing to the "Dependent List" box; **delta** disappears from the left-hand box and reappears under "Dependent List." Next, click on **attstyle** (the independent variable in our example) to select it, and then click on the right-arrow button pointing to the "Factor" box to move it there. The name **attstyle** now appears under "Factor."

If you wish to have SPSS print out means and other descriptive statistics along with the results of the *F*-test—and, of course, you do!—there is still one more step. In the dialog box illustrated in Figure 10.1, click on **Options . . .** to bring up another dialog box. In this box, which is illustrated in Figure 10.2,

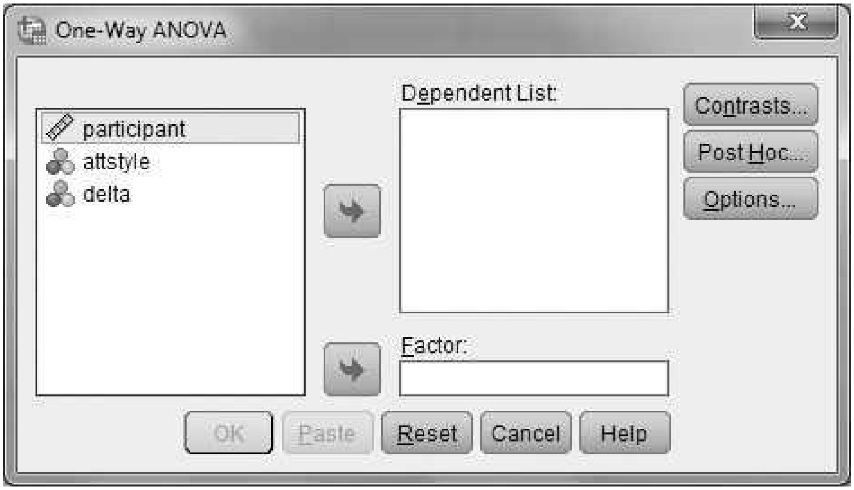


Figure 10.1

click on the box under "Statistics," next to "Descriptive." Then click on **Continue** to return to the previous dialog box.

If you wish to conduct only the "omnibus" analysis of variance to test the hypothesis that all three population means are equal, you are done. Click

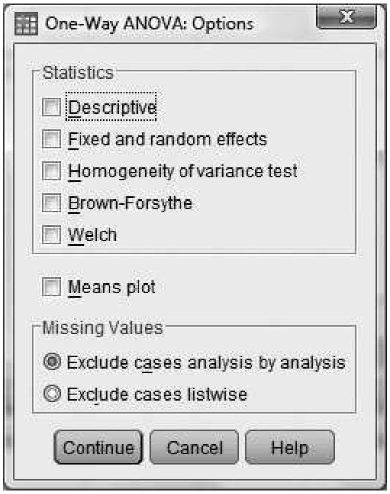


Figure 10.2

on **OK** to execute the analysis. If, in addition, you wish to conduct tests of multiple comparisons, using either post-hoc tests or planned comparisons/contrasts, follow the instructions in the corresponding sections below before clicking on **OK**.

Post-Hoc Tests. To conduct post-hoc tests in SPSS, specify your variables, and so on, as described in the preceding section; then click on the button labeled **Post Hoc . . .** at the right in the dialog box illustrated in Figure 10.1. This produces a new dialog box (see Figure 10.3) containing a list of available tests. For simple pairwise comparisons, the most commonly used tests are probably Student-Newman-Keuls (labeled "S-N-K") and the Tukey HSD test (labeled simply "Tukey"). Simply click on the box(es) next to the test(s) you wish SPSS to calculate and print. Then click on **Continue** to return to the dialog box illustrated in Figure 10.1, and click on **OK** to run the analysis.

To illustrate, the output produced by specifying Tukey's HSD test is reproduced later in this chapter.

Planned Comparisons/Contrasts. To conduct tests of planned comparisons or contrasts, specify your variables, and so on, as described previously, and then click on the button labeled **Contrasts . . .** at the right in the dialog box illustrated in Figure 10.1. This produces a new dialog box as illustrated in Figure 10.4.

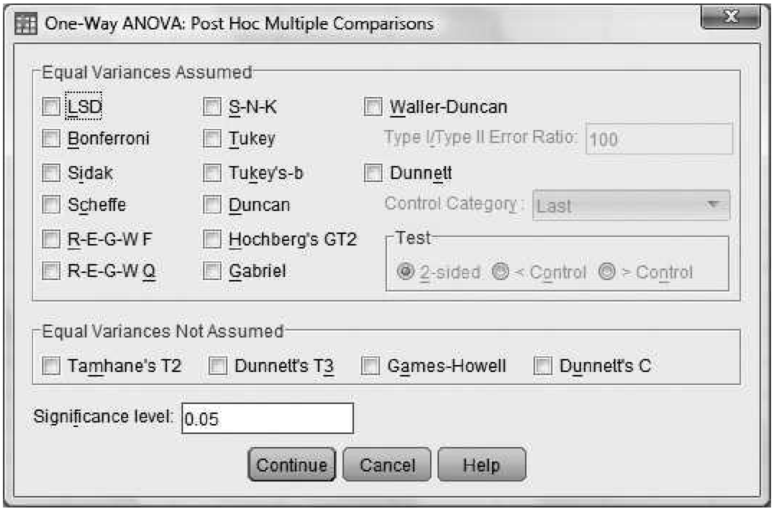


Figure 10.3

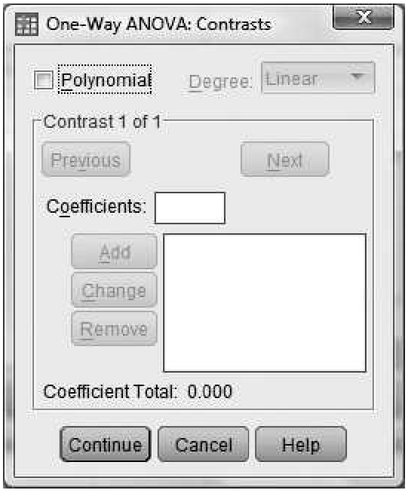


Figure 10.4

Specifying your contrast coefficients in this window can be a little tricky, but experiment a little and you'll get the hang of it. For the first contrast in our example, we wish to use the coefficients –2, 1, and 1, in that order, for *attstyle* groups 1, 2, and 3. To specify these coefficients, click in the box to the right of "Coefficients" and type the number of the coefficient for the first (lowest-numbered) group, in this case **–2**. Now click on the button labeled **Add** to add this coefficient to your list of coefficients (which is displayed in the box to the right of the **Add** button). Now click in the "Coefficients" box again, and type the coefficient for the second group (in this example, type **1**), and then click again on **Add**. Finally, click on the "Coefficients" box again, type **1** (the last coefficient), and click on **Add**. Your list of three coefficients now appears in the box to the right of the **Add** button: –2, 1, and 1, from top to bottom. If this is the only contrast you wish to test, click on **Continue** to return to the main One-Way ANOVA dialog box (Figure 10.1), and then click on **OK** to run the analysis.

In this example, however, we wish to test a second contrast as well. After specifying the coefficients for the first contrast as just described, click on **Next** to the right of the phrase "Contrast 1 of 1." This phrase now changes to "Contrast 2 of 2," which is what you specify now. The box containing your previously entered coefficients is again blank. Repeat the procedure outlined above to enter the coefficients **0**, **1**, and **–1** in order. Once these have been entered, click on **Continue** to return to the main One-Way ANOVA dialog box, and then click on **OK** to run the analysis.