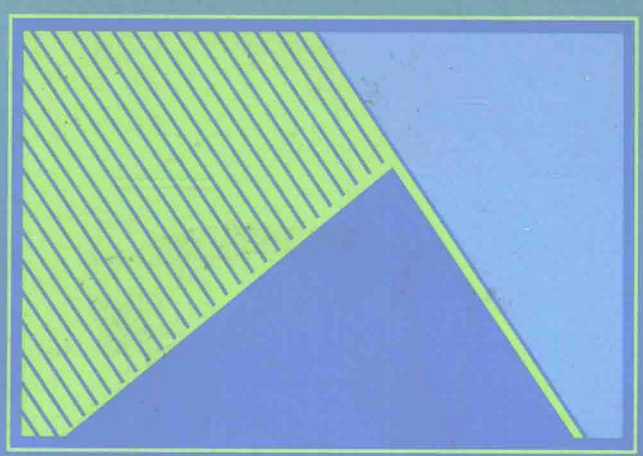
THIRD EDITION

# SPSS FOR WINDOWS

## STEP BY STEP

### A Simple Guide and Reference

### 10.0 UPDATE



Darren George • Paul Mallery

Visit www.abacon.com/george for related data sets

# SPSS® for Windows® Step by Step

## A Simple Guide and Reference

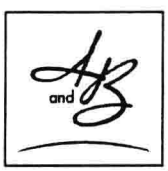
## 10.0 Update

Third Edition

Darren George
*Canadian University College*

Paul Mallery
*La Sierra University*

Allyn and Bacon
Boston London Toronto Sydney Tokyo Singapore





A Pearson Education Company
160 Gould Street
Needham Heights, MA 02494

Internet: www.abacon.com



ISBN 0-205-33127-0

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1 05 04 03 02 01 00

# PREFACE

SPSS is a powerful tool that is capable of conducting just about any type of data analysis used in the social sciences. While mathematics is generally thought to be the language of science, data analysis is the language of research. Research in many fields is critical for human progress, and as long as there is research, there will be the need to analyze data. The present book is designed to make data analysis more comprehendable and less toxic.

In our teaching at UCLA and other places, Paul and I have frequently encountered students so traumatized by the professor who cheerily says "Analyze these data on SPSS; get the manuals if you don't know how," that they dropped the course rather than continue the struggle. It is in response to this anguish that the present book was conceived. My background has been teaching high school mathematics, and my greatest pleasure in teaching was the challenge of making a process that was intrinsically complex as clear as possible. The ultimate goal in all our efforts with the present book has been to make SPSS procedures, above all else, clear.

As the book started to take shape, a second goal began to emerge. In addition to making SPSS procedures clear to the beginner, we wanted to create a tool that was an effective reference for anyone conducting data analysis. This involved the expansion of the original concept to include essentially all major statistical procedures that SPSS covers. The results of almost three years of effort you now hold in your hands.

While the first 16 chapters of the book would be understandable to many with very limited statistical background, the final 12 chapters involve procedures that progressively require a more secure statistical grounding. Those 12 chapters have provided our greatest challenge. At the beginning of each chapter we spend several pages describing the procedure that follows. But, how can one adequately describe factor analysis or discriminant analysis in five or six pages? The answer is simple: The procedures must be described at a common sense, conceptual level that avoids excessive detail and excessive emphasis on computation. However, writing those introductions has not been at all simple. The chapter introductions are the most painstakingly worked sections of the entire book. Although we acknowledge the absence of much detail in our explanation of most procedures, we feel that we have done an adequate job at a project that few would even attempt. How successful have we been at achieving clarity in very limited space? You, the reader, will be our ultimate judge.

SPSS Inc. has produced a collection of manuals to describe everything that their package of programs attempts to accomplish. These volumes, about 3000 pages of fine print documentation, are, if nothing else, comprehensive. Anything that SPSS is able to do is described in the manuals. For the experienced researcher, ownership of the manuals is required. We cannot cover in 350 pages what SPSS does in 3000. However, we are convinced that 95% of analyses that are *ever* conducted in the social sciences could be accomplished with the information presented in our book. For the additional 5% we frequently refer the reader to the SPSS manuals when the level of specificity required extends beyond the scope of the present volume.

## AUTHORS' BIOGRAPHICAL SKETCHES AND PRESENT ADDRESSES

Darren George is currently an associate professor of Psychology at

Canadian University College
235 College Avenue
College Heights, AB, T4L 2E5,
403-782-3381
dgeorge@cnnet.com

where he teaches personality psychology, social psychology, and research methods. He completed his M.A. in Experimental Psychology (1982) at California State University, Fullerton; taught high school mathematics for nine years (1980-1989) at Mark Keppel High School (Alhambra, CA) and Mountain View High School (El Monte, CA); then completed a Psychology Ph.D. at UCLA (1992) with emphases in personality psychology, social psychology, and measurement & psychometrics.

Paul Mallery is currently an associate professor of Psychology at

La Sierra University
4700 Pierce Street
Riverside, CA, 92515
909-785-2528
pmallery@lasierra.edu

where he teaches social psychology, experimental methodology, and SPSS. He received his Ph.D. in Social Psychology from UCLA (1994), with emphases in statistics and political psychology. Paul formerly worked as a computer specialist, both programming and teaching computer usage.

## ACKNOWLEDGMENTS

As we look over the creative efforts of the past years, Paul and I wish to acknowledge several people we have never met. These are individuals who have reviewed our work and offered invaluable insight and suggestions for improvement. Our gratitude is extended to Richard Froman of John Brown University, Michael A. Britt of Marist College, Marc L. Carter of the University of South Florida, and Randolph A. Smith of Ouachita Baptist University. And then there's the standard (but no less appreciated) acknowledgment of our families and friends who endured us while we wrote this. Particular notice goes to Marcus George (son of the first author) who contributed substantially to the design and format of the present volume, and to Brenda McMann who spent many hours proofing and editing.

# SPSS for Windows Step By Step: A Simple Guide and Reference

*Table of Contents*

# 1

# An Overview of SPSS for Windows Step by Step

THIS BOOK is designed to give you the step-by-step instructions necessary to do most major types of data analysis using SPSS for Windows. This software was originally created by three Stanford graduate students in the late 1960's, and SPSS (once "Statistical Package for the Social Sciences," now "Statistical Product and Service Solutions"), a Chicago-based firm, has grown to be one of the world's largest statistical software companies.

## NECESSARY SKILLS

For this book to be effective when you conduct data analysis with SPSS, you should have certain limited knowledge of statistics and a general acquaintance with the use of a computer. Each issue is addressed in the next two paragraphs.

**Statistics**: You should have had at least a basic course in statistics or be in the process of taking such a course. While it is true that this book devotes the first two or three pages of each chapter to a description of the statistical procedure that follows, these descriptions are designed to refresh the reader's memory, *not* to instruct the novice. While it is certainly possible for the novice to follow the steps in each chapter and get SPSS to produce pages of output, a fundamental grounding in statistics is important for an understanding of which procedures to use and what all the output means. In addition, while the first 16 chapters should be understandable by individuals with limited statistical background, the final 11 chapters deal with much more complex and involved types of analyses. These chapters require substantial grounding in the statistical techniques involved.

**Computer knowledge**: Your knowledge of the computer may be quite limited. The following, however, are necessary. You must:

- Have access to a personal computer that has
  - Windows 95, Windows 98, Windows 2000, or Windows NT installed
  - SPSS for Windows Release 10.0 installed
- Know how to turn the computer on
- Have a working knowledge of the keys on the keyboard and how to use a mouse

This book will take you the rest of the way. If you are using SPSS on a network of computers (rather than your own PC) the steps necessary to access SPSS for Windows may vary slightly from the single step shown in the pages that follow.

## SCOPE OF COVERAGE

SPSS for Windows is a complex and powerful statistical program by any standards. The software occupies over 70MB of your hard drive, and requires at least 32MB of RAM to operate adequately. If you didn't understand the last sentence, don't worry! Despite its size and complexity, SPSS has created a program that is not only powerful but is very user friendly (you're the user, the program tries to be friendly). By creating the windows version, SPSS has done for data analysis what Henry Ford did for the automobile: made it available to the masses. SPSS is able to perform essentially any type of statistical analysis ever used in the social sciences, and recent additions enable this package to operate effectively in the business world and in other scientific disciplines.

This book was written for Version 10 of SPSS for Windows. More specifically, the screen shots and output are based on Version 10.0. With only very minor exceptions, everything you see here will be very similar to SPSS Version 7.0 through 7.5, SPSS Version 8.0, and SPSS Version

9.0. Even if you are using Version 6.0 through 6.2, the procedures will be similar in most cases (the GLM procedures are the most noticeable exceptions), but many of the dialog boxes and all of the output will look different (though the content is usually quite similar). The most notable difference of Version 10.0 over Version 9.0 is the change in structure of the modules—something that has little influence on application in actual analyses. Overall, if you are using Version 10.0, this book will be 100% applicable; Version 9.0, 99% applicable; Versions 7.0 through 8.0, 95% applicable; Versions 6.0 through 6.2, 90% applicable.

Our book covers the statistical procedures present in the three *modules* created by SPSS that are most frequently used by social scientists. A module (within the SPSS context) is simply a set of different statistical operations. We include the **Base System Module**, the module covering **Advanced Models**, and the module that addresses **Regression Models**—all described in greater detail later in this chapter. To support their program, SPSS has created a set of comprehensive manuals that cover all procedures these three modules are designed to perform. To a person fluent in statistics and data analysis, the manuals are well written and intelligently organized. To anyone less fluent, however, the organization is often undetectable, and the comprehensiveness (about 3,000 pages of fine-print text) is overwhelming. Our book is about 350 pages long. Clearly we cannot cover in 350 pages as much material as the manuals do in 3,000, but herein lies our major advantage.

The purpose of this book is to make the fundamentals of most types of data analysis clear. To create this clarity requires the omission of much (often unnecessary) detail. Despite brevity, we have been keenly selective in what we have included and believe that the material presented here is sufficient to provide simple instructions that cover 95% of analyses ever conducted by social science researchers. Although we cannot substantiate that exact number, our time in the manuals suggests that at least 2,000 of the 3,000 pages involve detail that few researchers ever consider. How often do you really need 7 different methods of extracting and 6 methods of rotating factors in factor analysis, or 18 different methods for post-hoc comparisons after a one-way ANOVA? (By the way, that last sentence should be understood by statistical geeks only.)

We are in no way critical of the manuals; they do well what they are designed to do and we regard them as important adjuncts to the present book. When our space limitations prevent explanation of certain details, we often refer our readers to the SPSS manuals. Within the context of presenting a statistical procedure, we often show a window that includes several options but describe only one or two of them. This is done without apology except for the occasional "description of these options extends beyond the scope of this book" and cheerfully refer you to the appropriate SPSS manual. The ultimate goal of this format is to create clarity without sacrificing necessary detail.

## OVERVIEW

This chapter introduces the major concepts discussed in this book, giving a brief overview of the book's organization and the basic tools that are needed in order to use it.

If you want to run a particular statistical procedure, have used SPSS for Windows before, and already know which analysis you wish to conduct, you should read the Typographical and Formatting Conventions section in this chapter (pages 6-8) and then go to the appropriate chapter in the last portion of the book (Chapters 6 through 28). Those chapters will tell you exactly what steps you need to perform to produce the output you desire.

If, however, you are new to SPSS for Windows, then this chapter will give you important background information that will be useful whenever you use this book.

## THIS BOOK'S ORGANIZATION, CHAPTER BY CHAPTER

This book was created to describe the crucial concepts of analyzing data. There are three basic tasks associated with data analysis:

A. You must type data into the computer, and organize and format the data so both SPSS and you can identify it easily,
B. You must tell SPSS what type of analysis you wish to conduct, and
C. You must be able to interpret what the SPSS output means.

After this introductory chapter, Chapters 2 deals with basic operations such as types of SPSS windows, the use of the toolbar and menus, saving, viewing and editing the output, printing output, and so forth. While this chapter has been created with the beginner in mind, there is much SPSS-specific information that should be useful to anyone. Chapter 3 addresses the first step mentioned above—creating, editing, and formatting a data file. The SPSS data editor is an instrument that makes the building, organizing, and formatting of data files wonderfully clear and straightforward.

Chapters 4 and 5 deal with two important issues—modification and transformation of data (Chapter 4), and creation of graphs or charts (Chapter 5). Chapter 4 deals specifically with different types of data manipulation, such as creating new variables, reordering, restructuring, merging files, or selecting subsets of data for analysis. Chapter 5 introduces the basic procedures used when making and using graphs; most of the graphs, however, are described in the later chapters.

Chapters 6 through 28 then address Steps B and C—analyzing your data and interpreting the output. It is important to note that each of the analysis chapters is self-contained. If the beginner, for example, were instructed to conduct *t* tests on certain data, Chapter 11 would give complete instructions for accomplishing that procedure. In the Step by Step section, Step 1 is always "create a data file or edit (if necessary) an already existing file," and the reader is then referred to Chapter 3 for instructions if needed. The second step is always "start the SPSS program" and refers the reader to Chapter 2 if there are questions about how to do this. Then the steps that follow explain exactly how to conduct a *t* test.

As mentioned previously, this book covers three basic modules produced by SPSS: **Base System**, **Advanced Models**, and **Regression Models**. Since some networks at colleges or universities may not have all of these modules (the base module is always present), we organize the book according to the structure SPSS has imposed: In this book we cover ALL procedures included in the Base System module and then selected procedures from the far more complex Advanced Models and Regression Models. Chapters 6-22 deal with processes included in the Base System module. Chapters 23-27 deal with procedures included in the Advanced and Regression Models, and Chapter 28, the analysis of residuals, draws from all three modules.

**Base System Module**: Chapters 6 through 10 describe the most fundamental data analysis methods available, including frequencies, bar charts, histograms, and percentiles (Chapter 6); descriptive statistics such as means, medians, modes, skewness, and ranges (Chapter 7); crosstabulations and chi-square tests of independence (Chapter 8); subpopulation means (Chapter 9); and correlations between variables (Chapter 10).

The next group of chapters (Chapters 11 through 17) explains ways of testing for differences between subgroups within your data or showing the strength of relationships between a dependent variable and one or more independent variables through the use of *t* tests (Chapter 11), ANOVAs (Chapters 12, 13, and 14); linear, curvilinear, and multiple regression analysis (Chapters 15 and 16); and the most common forms of nonparametric tests are discussed in Chapter 17.

Reliability analysis (Chapter 18) is a standard measure used in research that involves multiple response measures, multidimensional scaling (a new topic in this edition) is designed to identify and model the structure and dimensions of a set of stimuli from dissimilarity data (Chapter 19), then factor analysis (Chapter 20), cluster analysis (Chapter 21), and discriminant analysis (Chapter 22) all occupy stable and important niches in research conducted by social scientists.

**Advanced and Regression Models**: The next series of chapters deals with analyses that involve multiple *dependent* variables (SPSS calls these procedures General Linear Models; they are also commonly called MANOVAs). Included under the heading General Linear Model are simple and general factorial models and multivariate models (Chapter 23), and models with repeated measures or within-subjects factors (Chapter 24).

The next three chapters deal with procedures that are only infrequently performed in the social sciences, but they are described here because when these procedures are needed they are indispensable. Chapter 25 describes logistic regression analysis and Chapters 26 and 27 describe hierarchical and nonhierarchical log-linear models, respectively. As mentioned previously, Chapter 28 on residuals closes out the book.

## AN INTRODUCTION TO THE EXAMPLE

A single example is used in 17 of the first 19 chapters of this book. For more complex procedures it has been necessary to select different examples to reflect the particular procedures that are presented. Examples are useful because often, things that appear to be confusing in the SPSS documentation become quite clear when you see an example of how they are done. Although only the most frequently used example is described here, there are a total of nine data sets that are used to demonstrate procedures throughout the book. Data files are available for download at http://www.abacon.com/george. These files can be of substantial benefit to you as you practice some of the processes presented here without the added burden of having to input the data. We suggest that you make generous use of these files by trying different procedures and then comparing your results with those included in the output sections of different chapters.

The example has been designed so that it may be used to demonstrate most of the statistical procedures presented here. It consists of a single data file used by a teacher who teaches three sections of a class with approximately 35 students in each section. For each student, the following information is recorded:

- ID number
- Name
- Gender
- Ethnicity

- Year in school
- Upper- or lower-division classperson
- Previous GPA
- Section
- Whether or not he or she attended review sessions or did the extra credit
- The scores on five 10-point quizzes and one 75-point final exam

In Chapter 4 we describe how to create four new variables. In all presentations that follow (and on the data files that accompanies this book), these four variables are also included:

- The total number of points earned
- The final percent
- The final grade attained
- Whether the student passed or failed the course

The example data file (the entire data set is displayed at the end of Chapter 3) will also be used as the example in the introductory chapters (Chapters 2 through 5). If you enter the data yourself and follow the procedures described in these chapters, you will have a working example data file identical to that used through the first half of this book. Yes, the same material is recorded on the downloadable data files, but it may be useful for you to practice data entry, formatting, and certain data manipulations with this data set. If you have your own set of data to work with, all the better.

One final note: All of the data in the **grades** file are totally fictional, so any findings exist only because we created them when we made the file.

## TYPOGRAPHICAL and FORMATTING CONVENTIONS

**Chapter organization**: Chapters 2 through 5 describe SPSS for Windows formatting and procedures, and the material covered dictates each chapter's organization. Chapters 6 through 28 (the analysis chapters) are, with only occasional exceptions, organized identically. This format includes:

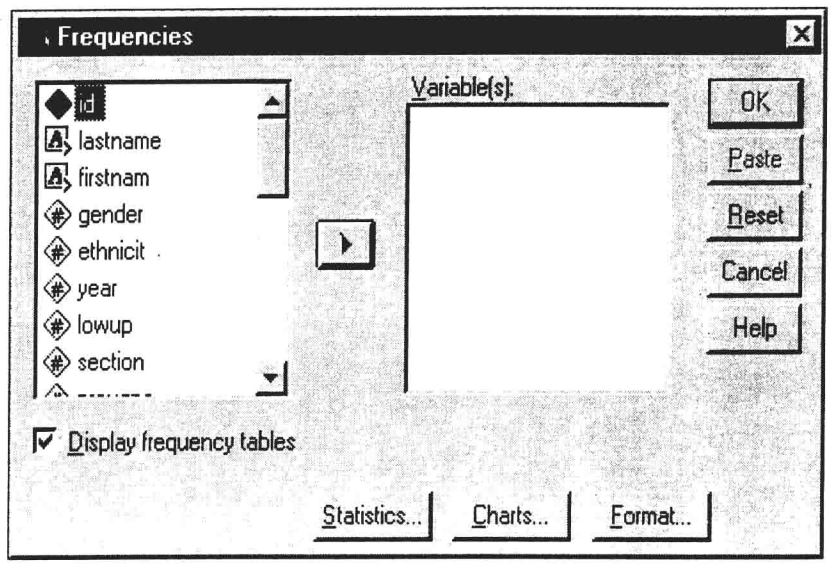
1. The **Introduction** in which the procedure that follows is described briefly but concisely. These introductions vary in length from one to seven pages depending on the complexity of the analysis being described.
2. The **Step by Step** section in which the actual steps necessary to accomplish particular analyses are presented. Most of the typographical and formatting conventions described below refer to the Step by Step sections.
3. The **Output** section, in which the results from analyses described earlier are displayed (often abbreviated), text clarifies the meaning of the output, and all of the critical output terms are defined.

**The screens**: Due to the very visual nature of SPSS, every chapter contains pictures of *screens* or *windows* that appear on the computer monitor as you work. The first picture from Chapter 6 (following page) provides an example. These pictures are labeled "Screens" despite the fact that sometimes what is pictured is a screen (everything that appears on the monitor at a given time) and other times is a *portion* of a screen (a window, a dialog box, or something smaller). This is done for sake of simplicity and ease of identification. If the reader sees reference to

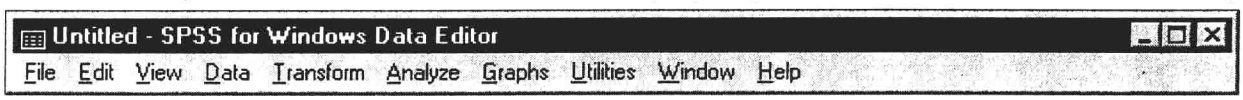
Screen 13.3, she knows that this is simply the third picture in Chapter 13. The screens are typically positioned within breaks in the text (the screen icon and a title are included) and are used for sake of reference as procedures involving that screen are described. Sometimes the screens are separate from the text and labels identify certain characteristics of the screen (see the inside front cover for an example). Because screens take up a lot of space, frequently used screens are included on the inside front and back covers of this book. At other times, within a particular chapter, a screen from a different chapter may be cited to save space.

1.1

*The Frequencies Window*




Sometimes a portion of a screen or window is displayed (such as the menu bar included here)




and is embedded within the text without a label.

**The Step by Step boxes**: Text that surrounds the screens *describes* a procedure, but it is the step by step boxes that identify exactly what must be done to *do* a procedure. The following box illustrates:

| In Screen | Do This | Step 3 (sample) |
|---|---|---|
| Front1 | File → Open | [ or  ] |
| Front2 | type grades.sav → OK | [ or grades.sav ] Data |

Sequence Step 3 means: "Beginning with Screen 1 (displayed on the inside front cover), click on the word **File,** then click on the word **Open**. At this point a new window will open (Screen 2 displayed on the inside front cover); type the words `grades.sav` then click the **OK** button, at which point a screen with your data file will open." Notice that within brackets shortcuts are sometimes suggested. For instance, rather than clicking **File** than clicking **Open**, it is quicker to click the  icon; instead of typing `grades.sav` then clicking **OK**, it is quicker to double click on the **grades.sav** file name. Items within Step-by-Step boxes include:

**Screens**: A small screen icon will be placed to the left of each group of instructions that are based on that screen. There are three different types of screen icons:

| Type of Screen Icon | Example Icon | Description of Example |
|---|---|---|
| Inside Cover Screens | Front1 | Screen #1 on the inside front cover |
| General Type Screens | Menu | Any screen with the menu bar across the top |
| | Graph | Any screen that displays a graph or chart |
| Chapter Screens | 4.3 | The third screen in Chapter 4 |
| | 13.3 | The third screen in Chapter 13 |

Other Images with special meaning inside of Step by Step boxes include:

| Image | What it Means |
|---|---|
|  | A single click of the left mouse button |
|  | A double-click of the left mouse button |
| type | A type icon appears before words that need to be typed |
| press | A press icon appears when a button such as the **TAB** key needs to be pressed |
| → | Proceed to the next step. |

Sometimes fonts can convey information, as well:

| Font | What it Means |
|---|---|
| `Monospaced font (Courier)` | Any text within the boxes that is rendered in the `Courier` font represents text (numbers, letters, words) to be *typed* into the computer (rather than being clicked or selected). |
| *Italicized text* | *Italicized text* is used for information or clarifications within the Step-by-Step boxes. |
| **Bold font** | The **bold font** is used for words that appear on the computer screen. |

The groundwork is now laid. We wish you a pleasant trip through the exciting and challenging world of data analysis!

# 2

# SPSS Windows Processes

WE MENTIONED in the introductory chapter that it was necessary for the user to understand how to turn the computer on and get as far as the Windows desktop. This chapter will give you the remaining skills that you will need to use SPSS for Windows: How to use the mouse, how to navigate using the taskbar, what the various buttons (on the toolbar and elsewhere) do, and how to navigate the primary windows used in SPSS.

If you are fluent with computers, you may not need to read this chapter as carefully as if you are new to computers. But every one should read at least portions of this chapter carefully; it contains a great deal of information unique to SPSS for Windows.

## THE MOUSE

Most computers today have a mouse attached to them. Not since the days of the Pied Piper of Hamelin (A.D. 1376) have mice proliferated so rapidly. When the mouse is moved on a hard surface (hopefully a mouse pad) an arrow or other symbol (called the *mouse cursor*) makes a corresponding movement on the screen. Three different types of mouse operations are used frequently in this book:

- **Point and click** (or single click) means to position the point of the arrow on the word, symbol, or icon that you desire, then press and release the left mouse button.
- **Double click** refers to holding the mouse so the arrow is positioned correctly and rapidly pressing the left mouse button twice.
- **Dragging** refers to selecting several items with one mouse operation. If, for instance, there is a list of words and you wish to select several of them, you may drag the mouse over all of the words. To do this, position the arrow on the first word and press the left mouse button, continue to hold the button down while you move the arrow down to the last desired word, and then release the mouse button. This will highlight all the words or text you wish for further operations.

Mice usually have at least two buttons (left and right). Many have a third (center) button, and some have dials as well. Within the context of this book, the left mouse button is used so extensively that a click or a double click always refers to a left-mouse-button operation. On the occasion when the right button is clicked, we will specify that in words. As mentioned in the first chapter a small picture of a mouse (  ) always represents the mouse click within Step by Step boxes.

Feel free to experiment with the right mouse button. In SPSS, when you position the cursor on an object or word, a right click may do nothing but often one of two different things occur:

- It provides a brief summary or definition of whatever you right clicked. This is particularly useful when you do not know the definition of a term or when you need to quickly access variable and value-label information. If you can't remember how you coded gender or ethnicity or can't recall the meaning of the variable named "cntbusb", a right click can provide quick relief.
- It provides a short pop-up menu of commands you can select, such as cut, copy, paste, delete, or others.

There are a few occasions in which moving the mouse cursor over something on the screen will provide either more information (in the case of toolbar icons), or will produce a submenu (in the case of the taskbar) that will cause additional information to pop up.