Fourteenth Edition

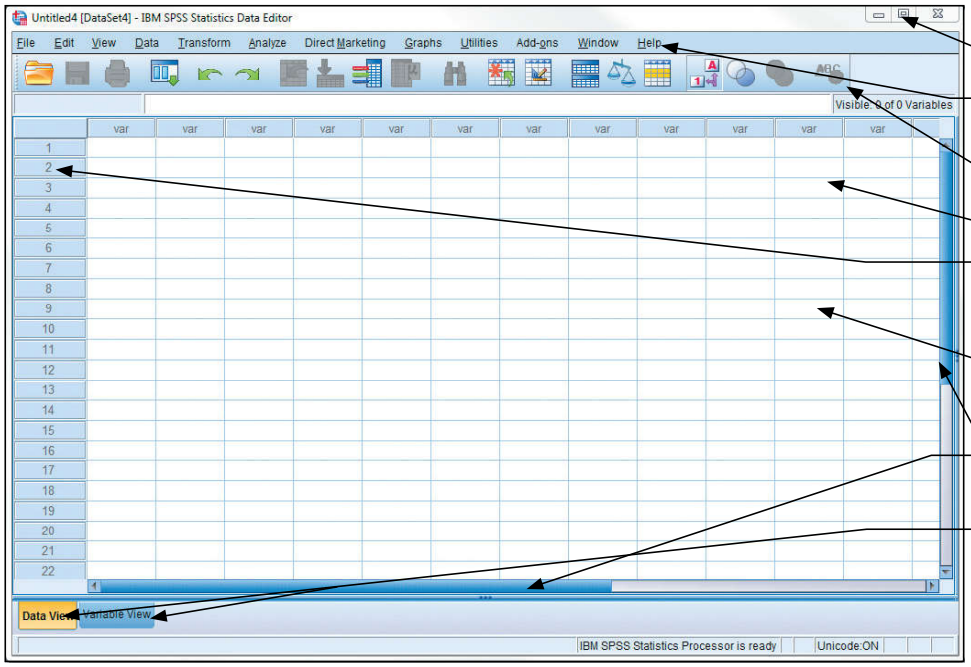
IBM SPSS Statistics 23 Step by Step

A SIMPLE GUIDE AND REFERENCE

Darren George
Paul Mallery

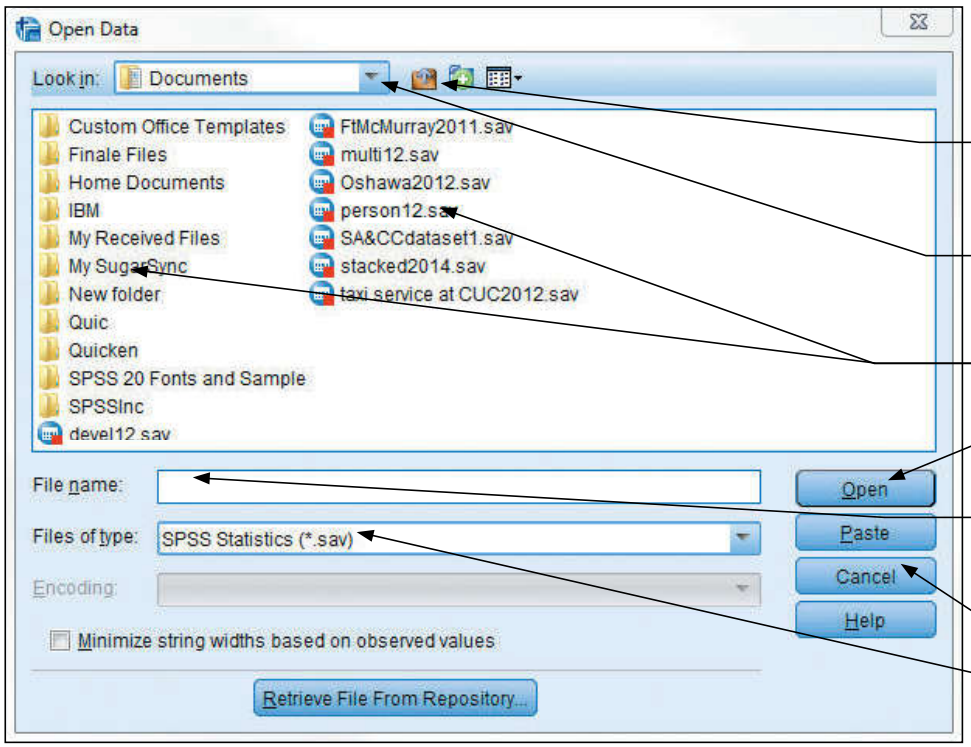
eRESOURCE
ROUTLEDGE

Front 1
Initial data screen



Icon	Function	Icon	Function	Icon	Function
	Click this to open a file		Find data		(upper left corner) the "+" sign indicates that this is the active file
	Save current file		Insert subject or case into the data file		Shifts between numbers and labels for variables with several levels
	Print file		Insert new variable into the data file		Go to a particular variable or case number
	Recall a recently-used command		Split file into subgroups		Use subsets of variables/use all variables
	Undo the last operation		Weight cases		Access information about the current variable
	Redo something you just undid		Select cases		Spell check

Front 2
Open Data Screen



This page intentionally left blank

IBM SPSS Statistics 23 Step by Step

IBM SPSS Statistics 23 Step by Step: A Simple Guide and Reference, 14e, takes a straightforward, step-by-step approach that makes SPSS software clear to beginners and experienced researchers alike. Extensive use of vivid, four-color screen shots, clear writing, and step-by-step boxes guide readers through the program. Exercises at the end of each chapter support students by providing additional opportunities to practice using SPSS.

All datasets used in the book are available for download at:
www.routledge.com/9780134320250

This page intentionally left blank

IBM SPSS Statistics 23 Step by Step

A Simple Guide and Reference

FOURTEENTH EDITION

Darren George
Burman University
(Formerly Canadian University College)

Paul Mallery
La Sierra University

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

Fourteenth edition published 2016
by Routledge
711 Third Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2016 Taylor & Francis

The right of Darren George and Paul Mallery to be identified as authors of this work has been asserted by them in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Thirteenth edition published by Pearson 2014.

Library of Congress Cataloging in Publication Data
A catalog record for this book has been requested.

ISBN: 978-1-138-68134-7 (hbk)
ISBN: 978-0-13-432025-0 (pbk)
ISBN: 978-1-315-54589-9 (ebk)

Typeset in Palatino LT Pro 9.5/13
by Lumina Datamatics, Inc.

To Elizabeth
—D.G.

To Suzanne
—P.M.

This page intentionally left blank

Contents

Preface xii

1 An Overview of IBM SPSS Statistics 1

Introduction: An Overview of IBM SPSS Statistics 23 1
1.1 Necessary Skills 1
1.2 Scope of Coverage 2
1.3 Overview 3
1.4 This Book's Organization, Chapter by Chapter 3
1.5 An Introduction to the Example 4
1.6 Typographical and Formatting Conventions 5

2a IBM SPSS Statistics Processes for PC 8

2.1 The Mouse 8
2.2 The Taskbar and Start Menu 8
2.3 Common Buttons 10
2.4 The Data and Other Commonly Used Windows 10
2.5 The Open Data File Dialog Window 13
2.6 The Output Window 16
2.7 Modifying or Rearranging Tables 19
2.8 Printing or Exporting Output 22
2.9 The "Options . . . " Option: Changing the Formats 24

2b IBM SPSS Statistics Processes for Mac 26

2.1 Selecting 26
2.2 The Desktop, Dock, and Application Folder 26
2.3 Common Buttons 28
2.4 The Data and Other Commonly used Windows 28
2.5 The Open Data File Dialog Window 30
2.6 The Output Window 34
2.7 Modifying or Rearranging Tables 36
2.8 Printing or Exporting Output 39
2.9 The "Options . . . " Option: Changing the Formats 41

3 Creating and Editing a Data File 43

3.1 Research Concerns and Structure of the Data File 43
3.2 Step by Step 44
3.3 Entering Data 51
3.4 Editing Data 52
3.5 Grades.sav: The Sample Data File 54

Exercises 58

4 Managing Data 59

4.1 Step By Step: Manipulation of Data 60
4.2 The Case Summaries Procedure 60
4.3 The Compute Procedure: Creating Variables 66
4.4 The Recode into Different Variables Procedure Creating New Variables 69
4.5 The Select Cases Option 73
4.6 The Sort Cases Procedure 75
4.7 Merging Files Adding Blocks of Variables or Cases 77
4.8 Printing Results 80

Exercises 82

5 Graphs and Charts: Creating and Editing 83

5.1 Comparison of the Two Graphs Options 83
5.2 Types of Graphs Described 83
5.3 The Sample Graph 84
5.4 Producing Graphs and Charts 85
5.5 Bugs 87
5.6 Specific Graphs Summarized 88
5.7 Printing Results 99

Exercises 100

6 Frequencies 101

6.1 Frequencies 101
6.2 Bar Charts 101
6.3 Histograms 101
6.4 Percentiles 102
6.5 Step by Step 102
6.6 Printing Results 108
6.7 Output 108

Exercises 111

7 Descriptive Statistics 112

7.1 Statistical Significance 112
7.2 The Normal Distribution 113
7.3 Measures of Central Tendency 114
7.4 Measures of Variability Around the Mean 114
7.5 Measures of Deviation from Normality 114
7.6 Measures for Size of the Distribution 115
7.7 Measures of Stability: Standard Error 115
7.8 Step by Step 115
7.9 Printing Results 118
7.10 Output 119

Exercises 120

8 Crosstabulation and χ^2 Analyses 121

8.1 Crosstabulation 121
8.2 Chi-Square (χ^2) Tests of Independence 121
8.3 Step by Step 123
8.4 Weight Cases Procedure: Simplified Data Setup 127
8.5 Printing Results 129
8.6 Output 129

Exercises 131

9 The Means Procedure 132

9.1 Step by Step 132
9.2 Printing Results 136
9.3 Output 136
Exercises 138

10 Bivariate Correlation 139

10.1 What is a Correlation? 139
10.2 Additional Considerations 141
10.3 Step by Step 142
10.4 Printing Results 146
10.5 Output 147
Exercises 148

11 The *t* Test Procedure 149

11.1 Independent-Samples *t* Tests 149
11.2 Paired-Samples *t* Tests 149
11.3 One-Sample *t* Tests 150
11.4 Significance and Effect Size 150
11.5 Step by Step 151
11.6 Printing Results 155
11.7 Output 155
Exercises 158

12 The One-Way ANOVA Procedure 159

12.1 Introduction to One-Way Analysis of Variance 159
12.2 Step by Step 160
12.3 Printing Results 165
12.4 Output 165
Exercises 168

13 General Linear Models: Two-Way ANOVA 169

13.1 Statistical Power 169
13.2 Two-Way Analysis of Variance 170
13.3 Step by Step 171
13.4 Printing Results 174
13.5 Output 174
Exercises 176

14 General Linear Models: Three-Way ANOVA 177

14.1 Three-Way Analysis of Variance 177
14.2 The Influence of Covariates 178
14.3 Step by Step 179
14.4 Printing Results 181
14.5 Output 181
14.6 A Three-Way Anova that Includes a Covariate 186
Exercises 190

15 Simple Linear Regression 191

15.1 Predicted Values and the Regression Equation 191
15.2 Simple Regression and the Amount of Variance Explained 193
15.3 Testing for a Curvilinear Relationship 193
15.4 Step by Step 196
15.5 Printing Results 200
15.6 Output 201
15.7 A Regression Analysis that Tests for a Curvilinear Trend 202
Exercises 203

16 Multiple Regression Analysis 204

16.1 The Regression Equation 204
16.2 Regression And R^2: The Amount of Variance Explained 206
16.3 Curvilinear Trends, Model Building, and References 206
16.4 Step by Step 208
16.5 Printing Results 213
16.6 Output 213
16.7 Change of Values as Each new Variable is Added 214
Exercises 217

17 Nonparametric Procedures 218

17.1 Step by Step 219
17.2 Are Observed Values Distributed Differently than a Hypothesized Distribution? 221
17.3 Is the Order of Observed Values Non-Random? 223
17.4 Is a Continuous Variable Different in Different Groups? 224
17.5 Are the Medians of a Variable Different for Different Groups? 226
17.6 Are My Within-Subjects (Dependent Samples or Repeated Measures) Measurements Different? 227
17.7 Printing Results 230

18 Reliability Analysis 231

18.1 Coefficient Alpha (α) 232
18.2 Split-Half Reliability 232
18.3 The Example 232
18.4 Step by Step 233
18.5 Printing Results 237
18.6 Output 237
Exercises 242

19 Multidimensional Scaling 243

19.1 Square Asymmetrical Matrixes (The Sociogram Example) 244
19.2 Step by Step 245
19.3 Printing Results 251
19.4 Output 251

20 Factor Analysis 254

20.1 Create a Correlation Matrix 254
20.2 Factor Extraction 254
20.3 Factor Selection and Rotation 255
20.4 Interpretation 257
20.5 Step by Step 258
20.6 Output 264

21 Cluster Analysis 267

21.1 Cluster Analysis and Factor Analysis Contrasted 267
21.2 Procedures for Conducting Cluster Analysis 268
21.3 Step by Step 270
21.4 Printing Results 276
21.5 Output 276

22 Discriminant Analysis 281

22.1 The Example: Admission into a Graduate Program 282
22.2 The Steps Used in Discriminant Analysis 282
22.3 Step by Step 284
22.4 Output 289

23 General Linear Models: MANOVA and MANCOVA 296

23.1 Step by Step 297
23.2 Printing Results 304
23.3 Output 305
Exercises 310

24 G.L.M.: Repeated-Measures MANOVA 311

24.1 Step by Step 312
24.2 Printing Results 317
24.3 Output 317
Exercises 321

25 Logistic Regression 322

25.1 Step by Step 323
25.2 Printing Results 327
25.3 Output 328

26 Hierarchical Log-Linear Models 332

26.1 Log-Linear Models 332
26.2 The Model Selection Log-Linear Procedure 333
26.3 Step by Step 334
26.4 Printing Results 338
26.5 Output 338

27 Nonhierarchical Log-Linear Models 344

27.1 Models 344
27.2 A Few Words about Model Selection 345
27.3 Types of Models Beyond the Scope of This Chapter 345
27.4 Step by Step 346
27.5 Printing Results 350
27.6 Output 350

28 Residuals: Analyzing left-over variance 353

28.1 Residuals 353
28.2 Linear Regression: A Case Study 354
28.3 General Log-Linear Models: A Case Study 356
28.4 Accessing Residuals in SPSS 360

Data Files 363
Glossary 367
References 373
Credits 375
Index 377

Preface

SPSS is a powerful tool that is capable of conducting just about any type of data analysis used in the social sciences, the natural sciences, or in the business world. 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 comprehendible and less toxic.

In our teaching, we 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. Darren George's background has been teaching high school mathematics, and Paul Mallery worked his way through college training people to use computers and programming computers. Both of us find great pleasure in the challenge of making a process that is 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 most of the major statistical procedures that SPSS covers in the base module and many of the procedures in the advanced and regression modules as well. The result of years of effort you now hold in your hands.

The 14th edition is improved in both content and style. Most visibly, the new design and the switch from black-and-white to color make the text easier to use, and make the SPSS screens and output in the book more user-friendly. But more importantly than the improvements in style, changes in this edition have incorporated new content, including a much more extensive discussion of effect size throughout the entire book and a complete rewrite of the Nonparametrics chapter (Chapter 17).

Effect size: While effect size was mentioned in several chapters in prior editions, the discussion of this topic now allows the professor to assist students in a comprehensive understanding of this important issue across a number of different procedures. It includes a discussion of Carmer's V for chi-square (χ^2) analysis, r and r^2 for correlations (in which the test statistic and the measure of effect size are the same), Cohen's d for t tests, eta (η) and eta-squared (η^2) for ANOVA, and R and R^2 for regression.

The rewrite of the Nonparametrics chapter reflects both a dramatic shift in the way SPSS handles nonparametric procedures, an introduction that provides a discussion of when nonparametric tests should be used (and when they might be avoided), and a more complete coverage of the material. Rather than simply covering nine different procedures (as in previous editions), five broad categories of nonparametric procedures are addressed: Do observed values differ from a hypothesized distribution? Is the order of observed values non-random? Is a continuous variable dissimilar in different groups? Do the medians differ for different groups? And, do my within-subjects measurements differ?

In the 22nd and 23rd releases, SPSS has cleaned up some bugs in the graphing programs, and there are a few access procedures that are more intuitive than in prior editions. As usual, all screens have been updated, all step-by-step sequences executed, and all outputs scrutinized to make certain everything in the current edition is accurate.

While the first 16 chapters of the book cover basic topics and 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, for instance, factor analysis or discriminant analysis in five or six pages? The answer is simple: We can't, but we can describe the procedures at a common sense, conceptual level that avoids excessive detail and excessive emphasis on computation that is useful as an introduction for beginners or as a useful adjunct to more advanced reading or mentoring for more advanced data analysts. Writing these introductions has not at all been 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? The fact that this book is now in its 14th edition, has been an academic best seller for most of those editions, and is distributed in 85 countries of the world suggests that our efforts have not been in vain.

Authors' Biographical Sketches and Present Addresses

Darren George is currently a professor of Psychology at (formerly Canadian University College)

Burman University
5415 College Avenue
Lacombe, AB, T4L 2E5
403-782-3381, Ext. 4082
dgeorge@burmanu.ca

where he teaches personality psychology, social psychology, and research methods. He completed his MA 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), and then completed a Psychology PhD at UCLA (1992) with emphases in personality psychology, social psychology, and measurement and psychometrics. Darren has now been a professor at Burman University for 22 years.

Paul Mallery is currently a professor of Psychology at
La Sierra University
4500 Riverwalk Parkway
Riverside, CA, 92515
951-785-2528
pmallery@lasierra.edu

where he teaches social psychology and related courses and experimental methodology (including the application of SPSS). He received his PhD 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, we wish to acknowledge several people 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, Randolph A. Smith of Ouachita Baptist University, Roberto R. Heredia of Texas A&M International University, and several anonymous reviewers. And then there's the standard (but no less appreciated) acknowledgment of our families and friends who endured while we wrote this. Particular notice goes to our wives Elizabeth George and Suzanne Mallery as well as our families for their support and encouragement.

This page intentionally left blank

Chapter 1
An Overview of IBM® SPSS® Statistics

Introduction: An Overview of IBM SPSS Statistics 23

THIS BOOK gives you the step-by-step instructions necessary to do most major types of data analysis using SPSS. The software was originally created by three Stanford graduate students in the late 1960s. The acronym "SPSS" initially stood for "**S**tatistical **P**ackage for the **S**ocial **S**ciences." As SPSS expanded their package to address the hard sciences and business markets, the name changed to "Statistical Product and Service Solutions." In 2009 IBM purchased SPSS and the name morphed to **"IBM SPSS Statistics**." SPSS is now such a standard in the industry that IBM has retained the name due to its recognizability. No one particularly cares what the letters "SPSS" stand for any longer. IBM SPSS Statistics is simply one of the world's largest and most successful statistical software companies. In this book we refer to the program as **SPSS**.

1.1 Necessary Skills

For this book to be effective when you conduct data analysis with SPSS, you should have certain limited knowledge of statistics and have access to a computer that has the necessary resources to run SPSS. 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 12 chapters deal with much more complex and involved types of analyses. These chapters require substantial grounding in the statistical techniques involved.

COMPUTER REQUIREMENTS You must:

- Have access to a personal computer that has
  - Microsoft® Windows Vista® or Windows® 7 or 8.1 or 10; MAC OS® 10.8 (Mountain Lion) or higher installed
  - IBM SPSS Statistics 23.0 installed
- Know how to turn the computer on

- Have a working knowledge of the keys on the keyboard and how to use a mouse—or other selection device such as key board strokes or touch screen monitors.

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 or MAC) the steps necessary to *access* IBM SPSS Statistics may vary slightly from the single step shown in the pages that follow.

1.2 Scope of Coverage

IBM SPSS Statistics is a complex and powerful statistical program by any standards. The software occupies about 800 MB of your hard drive and requires at least 1 GB of RAM to operate adequately. Despite its size and complexity, SPSS has created a program that is not only powerful but is user friendly (you're the user; the program tries to be friendly). By improvements over the years, 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, in the business world, and in other scientific disciplines.

This book was written for Version 23 of IBM SPSS Statistics. More specifically, the screen shots and output are based on Version 23.0. With some exceptions, what you see here will be similar to SPSS Version 7.0 and higher. Because only a few parts of SPSS are changed with each version, most of this book will apply to previous versions. It's 100% up-to-date with Version 23.0, but it will lead you astray only about 2% of the time if you're using Version 21.0 or 22 and is perhaps 60% accurate for Version 7.0 (if you can find a computer and software that old).

Our book covers the statistical procedures present in three of the *modules* created by SPSS that are most frequently used by researchers. A module (within the SPSS context) is simply a set of different statistical operations. We include the **Base Module** (technically called **IBM SPSS Statistics Base**), the module covering advanced statistics (**IBM SPSS Advanced Statistics**), and the module that addresses regression models (**IBM SPSS Regression**)—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 (the equivalent of almost 2,000 pages of fine-print text) is overwhelming. To the best of our knowledge, hard-copy manuals are no longer available but most of this information may now be accessed from SPSS as PDF downloads. The same information is also available in the exhaustive online Help menu. Despite changes in the method of accessing this information, for sake of simplicity we still refer to this body of information as "SPSS manuals" or simply "manuals." Our book is about 400 pages long. Clearly we cannot cover in 400 pages as much material as the manuals do in 2,000, but herein lies our 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 researchers. Although we cannot substantiate that exact number, our time in the manuals suggests that at least 1,600 of the 2,000 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.

1.3 Overview

This chapter introduces the major concepts discussed in this book and gives 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 IBM SPSS Statistics before, and already know which analysis you wish to conduct, you should read the Typographical and Formatting Conventions section in this chapter (pages 5–7) 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 IBM SPSS Statistics, then this chapter will give you important background information that will be useful whenever you use this book.

1.4 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, Chapter 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 a number of different graphs; some graphs, however, are described more fully 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 "start the SPSS program" and refers the reader to Chapter 2 if there are questions about how to do this. The second step 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. Then the steps that follow explain exactly how to conduct a *t* test.

As mentioned previously, this book covers three modules produced by SPSS: **IBM SPSS Statistics Base, IBM SPSS Advanced Statistics**, and **IBM SPSS Regression**. Since some computers at colleges or universities may not have all of these modules (the **Base** module is always present), the book is organized according to the structure SPSS has imposed: We cover almost all procedures included in the **Base** module and then selected procedures from the more complex **Advanced** and **Regression** Modules. Chapters 6–22 deal with processes included in the Base module. Chapters 23–27 deal with procedures in the Advanced Statistics and Regression Modules, and Chapter 28, the analysis of residuals, draws from all three.

IBM SPSS STATISTICS BASE, 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 is designed to identify and model the structure and dimensions of a set of stimuli from dissimilarity data (Chapter 19); and then factor analysis (Chapter 20), cluster analysis (Chapter 21), and discriminant analysis (Chapter 22) all occupy stable and important niches in research conducted by scientists.

IBM SPSS ADVANCED STATISTICS AND REGRESSION: 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 or MANCOVAs). 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, 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.

1.5 An Introduction to the Example

A single data file is used in 17 of the first 19 chapters of this book. For more complex procedures it has been necessary to select different data files to reflect the particular procedures that are presented. Example data files 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 sample data file is described here, there are a total of 12 data sets that are used to demonstrate procedures throughout the book, in addition to data sets utilized in the exercises. Data files are available for download at *www.spss-step-by-step.net*. 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 it can 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 class person
- 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 file available on the website), 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.

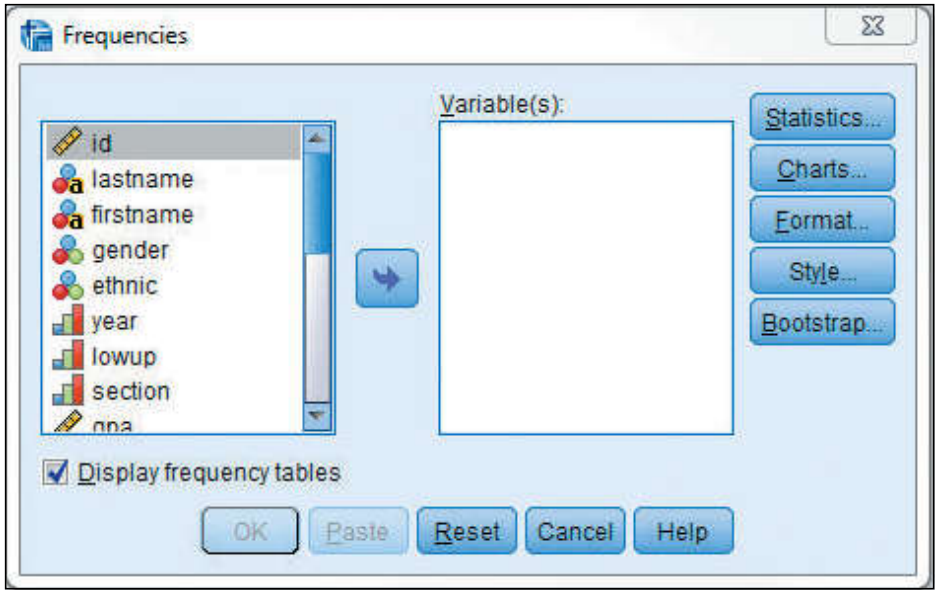
1.6 Typographical and Formatting Conventions

CHAPTER ORGANIZATION Chapters 2 through 5 describe IBM SPSS Statistics 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:

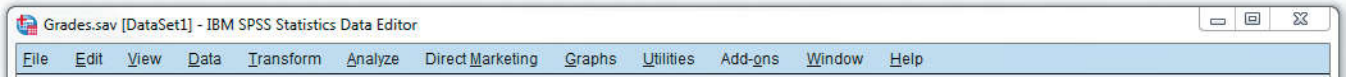
1. The **Introduction** in which the procedure that follows is described briefly and concisely. These introductions vary in length from one to seven pages depending on the complexity of the analysis being described.
2. The **Step by Step** section in which the actual steps necessary to accomplish particular analyses are presented. Most of the typographical and formatting conventions described in the following pages 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 (below) 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). 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.

Screen 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 may designate a procedure, but it is the Step by Step boxes that identify exactly what must be done to execute a procedure. The following box illustrates:

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

Sequence Step 3 means: "Beginning with Screen 1 (displayed on the inside front cover), click on the word **File,** move the cursor to **Open,** and then click the word **Data**. At this point a new window will open (Screen 2 on the inside front cover); type 'grades.sav' and then click the **Open** button, at which point a screen with your data file opens." Notice that within brackets shortcuts are sometimes suggested: Rather than the **File** → **Open** → **Data** sequence, it is quicker to click the icon. Instead of typing grades.sav and then clicking **Open**, it is quicker to double click on the **grades.sav** (with or without the **".sav"** suffix; this depends on your settings) 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 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
	21.4	The fourth screen in Chapter 21

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

Image	What it Means
	A single click of the left mouse button (or select by touch screen or key strokes)
	A double-click of the left mouse button (or select by touch screen or key strokes)
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 journey through the exciting and challenging world of data analysis!

Chapter 2A
IBM SPSS Statistics Processes for PC

WE MENTIONED in the introductory chapter that it is 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 required 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 someone less familiar. But everyone should read at least portions of this chapter carefully; it contains a great deal of information unique to SPSS for Windows.

2.1 The Mouse

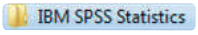
Over the years SPSS has modified their product so that mouse operations parallel those of many major programs. The left-mouse button **point and click**, **double click**, and **dragging** operate in ways similar to major word processing programs, although sometimes SPSS has unique responses to these common applications. Important differences will be noted in the chapters where they apply. A **right-click** and running the cursor over a word or object also produces similar results. In early editions of the book we provided a thorough description of mouse operations. Now, the computer world is moving in a direction where one day the mouse may be obsolete—touch screens, key-stroke operations, and other selection devices may one day predominate. Because of this we have shifted our former "mouse-click" icon ( ) to an icon designed to mean "select" ( ). If you are operating with a mouse this icon still means "left mouse click."

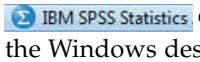
2.2 The Taskbar and Start Menu

Once you have arrived at the Windows desktop, your screen should look something like that shown on the following page. It will certainly not look exactly like this, but it will be similar. There will typically be a number of icons along the left side of the screen, and a bar across the bottom (or top) of the screen (with the word "Start" on the left, and the time on the right).

There are two main types of icons on the Windows desktop: **Program** icons represent a particular program, while **folder** icons actually contain other icons (usually several programs that are related in some way).

The most important thing you need to know about the Windows desktop (at least as long as you are reading this book) is how to start the SPSS program. To do this on most computers, you need to click the button, move the cursor over the

menu folder, and then over the 

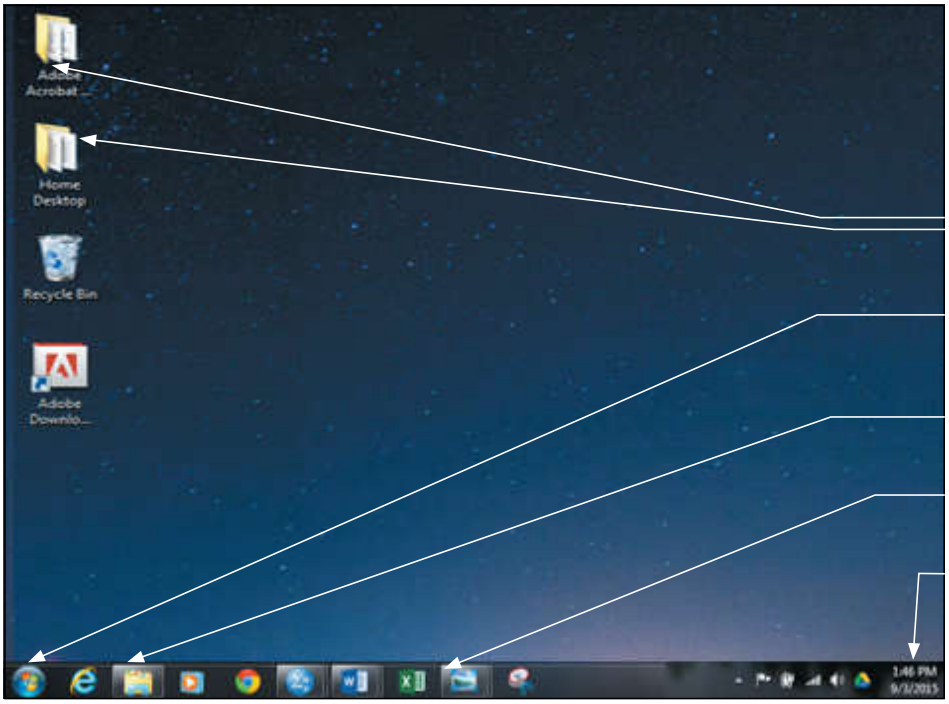
program icon. Once 

emerges, click on the icon, and SPSS will begin. On most computers, the Windows desktop will look similar to that shown in Screen 2.2 (following page) immediately before you click on

.

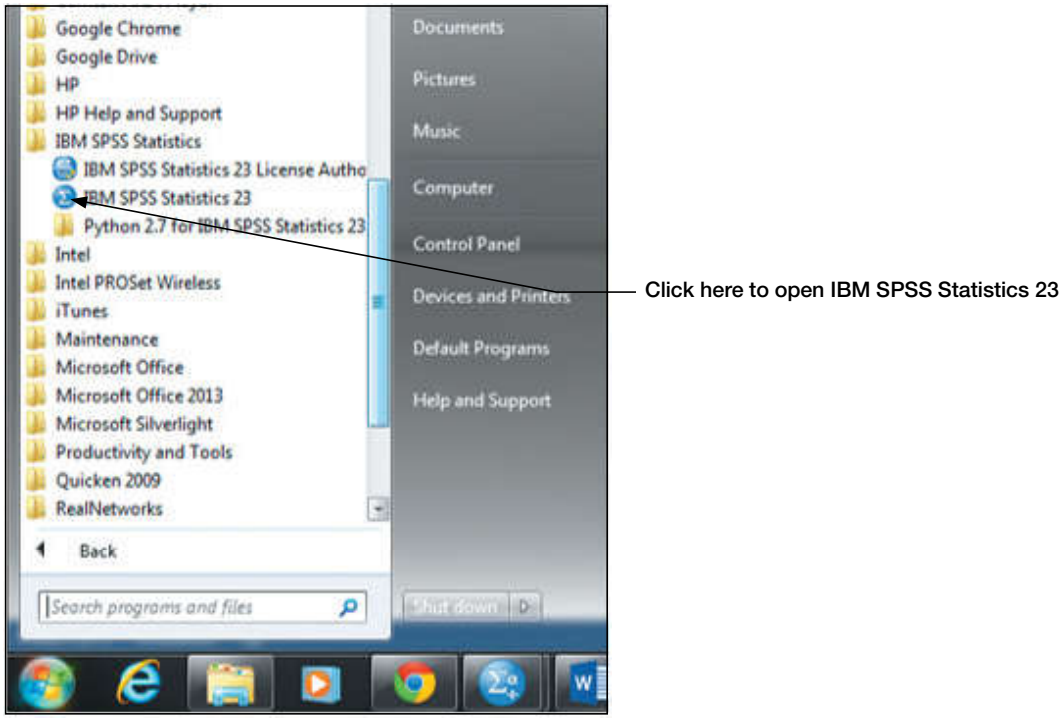
Reprinted with permission of Apple Inc. and Reprinted with permission of Microsoft, Inc.

Screen 2.1 Windows desktop



One word of warning: On some computers, the SPSS program icon may be in a different location within the Start menu. You may have to move the cursor around the Start menu (look especially for any folders labeled "IBM" or "SPSS"). Occasionally, the icon is on the Windows desktop (along the left side), and you don't have to use the Start button at all.

Screen 2.2 View of the Windows desktop Start menu immediately before clicking to start the SPSS program



In addition to starting the SPSS program, the other important required skill when using the Taskbar is changing between programs. This is especially important because SPSS is actually a collection of several programs. When you first start the SPSS system,