A Guide to SPSS/PC+

Related Macmillan titles

Roy Anderson, *Management, Information Systems and Computers*

I.O. Angell, *Advanced Graphics with the IBM Personal Computer*

B.V. Cordingley and D. Chamund, *Advanced BASIC Scientific Subroutines*

N. Frude, *A Guide to SPSS/PC+*

Percy Mett, *Introduction to Computing*

Tony Royce, *COBOL – An Introduction*

Tony Royce, *Structured COBOL – An Introduction*

Barry Thomas, *A PostScript Cookbook*

A Guide to SPSS/PC+

Neil Frude

Second Edition

© Neil Frude 1987, 1993

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No paragraph of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, 90 Tottenham Court Road, London WC1P 9HE.

Any person who does any unauthorised act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

First edition 1987

Reprinted 1989, 1990, 1992

Second edition 1993

Published by

THE MACMILLAN PRESS LTD

Houndmills, Basingstoke, Hampshire RG21 2XS
and London

Companies and representatives
throughout the world

ISBN 978-0-333-58908-3

ISBN 978-1-349-13164-8 (eBook)

DOI 10.1007/978-1-349-13164-8

A catalogue record for this book is available
from the British Library.

Typeset by

Ponting–Green Publishing Services, Sunninghill, Berks

SPSS/PC+, SPSS/PC, SPSS/PC+ Tables, SPSS/PC+ Advanced Statistics, SPSS/Pro, SPSS, and SCSS are the trademarks of SPSS Inc. for its proprietary computer software.

A Guide to SPSS/PC+ is not sponsored or approved by or connected with SPSS Inc. All references in the text of this book to SPSS products are to the trademarks of SPSS Inc.

IBM, IBM PC, IBM/XT and IBM/AT are the registered trademarks of International Business Machines Corporation.

All names of computer programs are hereby acknowledged as trademarks (registered or otherwise), whether or not the symbol appears in the text.

Contents

<i>Preface</i>	ix
<i>A guide to the Appendices</i>	xiv
<i>Acknowledgements</i>	xvii
1 Introducing SPSS/PC+	1
1.1 Collecting and coding data	1
1.2 Entering and saving data	3
1.3 Annotating data	4
1.4 Using the annotated data	7
2 Collecting and Coding Data	13
2.1 Variables, cases and values	13
2.2 Coding strategies	14
2.3 Preparation of the questionnaire and codebook	17
2.4 Conducting the survey	27
2.5 Coding the data	28
3 SPSS/PC+ and the Computer	33
3.1 The computer keyboard – a guided tour	33
3.2 Drives, directories and files	35
3.3 The three program levels	37
3.4 Moving between program levels	38
3.5 The SHOW and SET commands	42
4 Creating a Data File	45
4.1 Using REVIEW	45
4.2 Entering a data matrix and saving it as a data file	50
4.3 To save the file	50
4.4 To copy and print the file	51
4.5 Entering REVIEW directly from DOS	52
4.6 Looking forward	53
5 Creating a Data Definition File	55
5.1 The DATA LIST command	57

5.2. The VARIABLE LABELS command	59
5.3 The VALUE LABELS command	61
5.4 The MISSING VALUES command	62
5.5 Writing the data definition file to disk	63
5.6 Testing data and data definition files for errors	63
5.7 Obtaining information about the file variables and values	67
5.8 The DISPLAY command	67
5.9 The LIST command	69
5.10 Using active files for analysis	70
6 Statistical Analysis I: FREQUENCIES and CROSSTABS	72
6.1 Using procedures	73
6.2 The FREQUENCIES procedure	75
6.3 The CROSSTABS procedure	79
6.4 Inferential statistics: the use of chi-square	85
6.5 Using the listing file to view results	88
7 Statistical Analysis II: MEANS and T-TEST	91
7.1 The MEANS procedure	91
7.2 The T-TEST procedure	94
8 Transforming Variables; Selecting and Sorting Cases	102
8.1 The RECODE command	102
8.2 The COMPUTE command	104
8.3 Conditional transformations: IF	110
8.4 Selecting cases	114
8.5 Sorting cases	120
9 Command Files and System Files	123
9.1 Types of command	123
9.2 Command files	124
9.3 Using the log file	129
9.4 System files	132
10 Further Explorations of REVIEW	137
10.1 The story so far	137
10.2 REVIEW help	138
10.3 Paging and searching	138
10.4 Obtaining a list of variables	141
10.5 Calling another file into REVIEW	142
10.6 Listing file names	142
10.7 Using the 'help' facility	143
10.8 Using the on-line glossary	145

11 Statistical Analysis III: CORRELATION and PLOT	147
11.1 The CORRELATION procedure	147
11.2 The PLOT procedure	151
12 Statistical Analysis IV: ONEWAY and ANOVA	159
12.1 The ONEWAY procedure	160
12.2 The ANOVA procedure	166
13 Statistical Analysis V: Non-parametric Statistics – NPAR TESTS and RANK	169
13.1 Choosing the appropriate test	170
13.2 One-sample tests	172
13.3 One- or two-sample tests	175
13.4 Two-sample tests	177
13.5 Two- or <i>k</i> -sample tests	182
13.6 <i>k</i> -sample tests	184
13.7 NPAR: Statistics and Options	186
13.8 The RANK procedure	187
14 Statistical Analysis VI: RELIABILITY	194
14.1 The concept of reliability	194
14.2 The assessment of reliability	194
14.3 The RELIABILITY procedure	196
14.4 The use of RELIABILITY: an example using ‘alpha’	198
14.5 The use of RELIABILITY: an example using ‘split’	205
14.6 Using the scale	207
15 Joining Files	209
15.1 The JOIN MATCH command	210
15.2 The JOIN ADD command	215
16 Advanced Uses of the System	219
16.1 New data options	219
16.2 Transporting files between computers	227
16.3 Expanding your SPSS/PC+ repertoire	232
16.4 A final word	242
Appendices	
Appendix A: Notes for Trainers	245
Appendix B: Using Word-processor and Other Files with SPSS/PC+	251
Appendix C: The Gender Attitudes Survey: QUESTIONNAIRE	256
Appendix D: The Gender Attitudes Survey: CODEBOOK	261

Appendix E: The Gender Attitudes Survey: DATA	266
Appendix F: The Gender Attitudes Survey: DATA DEFINITION FILE	268
Appendix G: Glossary of Statistical, Computing and SPSS/PC+ Terms	272
Appendix H: Types of File Used or Created by SPSS/PC+	311
Appendix I: Variations between Different Versions of SPSS/PC+	312
Appendix J: Brief guide to SPSS/PC+ Commands and Subcommands	314
Appendix K: Brief Guide to REVIEW Commands	326
<i>Index</i>	328

Preface

SPSS/PC+ (the PC version of the Statistical Package for the Social Sciences) is the best-selling statistical analysis program for microcomputers. It enables data from surveys and experiments to be analyzed fully and flexibly. It has facilities for the extensive manipulation and transformation of data, and includes a wide range of procedures for both simple and highly complex statistical analysis. It also provides the opportunity for the researcher to produce fully labelled tables and graphs which can be easily incorporated into a final project report.

Over the 25 years since it was first devised, the versatile SPSS system has become an indispensable tool for many workers in social science research (including psychology, sociology, politics, human geography, business management, etc.) and in business and government. Many of the largest and most important surveys in the past quarter century have been analyzed using one or other version of the system. SPSS (including the microcomputer version) is regularly used by government agencies, and by many major industrial corporations, market research companies and opinion poll organizations.

Familiarity with SPSS/PC+ will do more than merely enable students to analyze their current projects. Hands-on experience with the SPSS system will help them to develop valuable statistical skills and will provide a working knowledge of a system which will certainly continue to play a central role in social science research, and many commercial fields, for the foreseeable future. When you have worked through this book you can feel confident about having a basic knowledge of a package which is the best of its kind, and the skills you will have gained are likely to remain useful in a range of contexts, and however advanced your project work becomes.

Unfortunately, there is a problem. Although SPSS/PC+ is not difficult to *use*, many beginners find it difficult to *learn to use* the package. Although the documentation for SPSS/PC+ is excellent in many respects, the encyclopaedic nature of the manuals can prove a major obstacle to those wishing to become acquainted with the system for the first time. While the standard SPSS/PC+ manuals are indispensable as reference texts for the experienced user, they are not ideal for beginners who wish to learn to use the system.

There are several reasons for regretting the fact that beginners often

experience difficulties in getting acquainted with the SPSS system, and that teachers often prefer to introduce the student to relevant areas of data analysis by employing less sophisticated packages:

- The power and flexibility of SPSS/PC+ means that the analysis of even modest datasets can be considerably enhanced through its use
- Experience with SPSS/PC+ provides students with an opportunity to explore the use of advanced multivariate techniques employed in many of the major studies about which they read
- SPSS has become a standard package within the social sciences and many related fields, so that students experienced in its use can feel confident that they have added significantly to their range of research- and employment-related skills.

This book aims to teach the use of SPSS/PC+ as clearly and as simply as possible. It is a stand-alone text covering many of the basic SPSS/PC+ procedures and may be used by the student working independently or may form the basis for class teaching. In either case it is assumed that the student has regular access to an SPSS/PC+ facility and will be able to work through parts of the text while interacting with the microcomputer.

Although this book does not assume a prior knowledge of survey work, statistics, computer programming or report writing, it is clear that one introductory book cannot teach all of these skills. Although this book describes aspects of survey methodology (particularly questionnaire design), students who need basic information about statistical tests may need to consult an introductory statistics textbook while working through this book. *A Guide to SPSS/PC+* is not concerned with computer programming (although it is concerned with computer *use*). SPSS/PC+ does not require the user to write programs. It is a computer program which enables the user to enter data, to describe, label and manipulate the data, and to perform simple and complex analyses. The various facilities offered by the program are activated by issuing commands that request procedures to be operated on specified data. The commands consist largely of keywords combined together following simple rules of syntax.

In writing this book I have placed the emphasis on simplicity and clarity rather than on exploring the full complexity of the system. When you have worked through this book (with hands-on experience of SPSS/PC+, using either your own data or that supplied in Appendix E) you will be well placed to cope with and benefit from the information provided in the SPSS/PC+ manuals. This book is not intended to replace these manuals and they remain the ultimate reference. For simple analyses, however, this book can act as a stand-alone text.

Unlike the SPSS/PC+ manuals, this book provides a progressive tutorial. The simplest and most fundamental concepts are introduced first. Later

sections build on the knowledge which has been acquired at earlier stages. After a brief overview of the SPSS system, the reader is taught how to use a microcomputer to explore SPSS/PC+. Experience has shown that students often find the various types of 'file' used by SPSS/PC+ a source of some confusion, and special care has therefore been taken to distinguish carefully between the different types of file. These are introduced progressively through the text. This book focuses on data collected in a questionnaire survey, although SPSS/PC+ can also be used to analyze data from other types of study, including experiments. (Incidentally, although the word 'data' is sometimes thought of as singular, it is in fact plural, as in 'the data were collected' – *not* 'the data was collected'; a single data item is a 'datum'.)

A Guide to SPSS/PC+ provides step-by-step instructions for collecting information from a survey and then using SPSS/PC+ to analyze the data.

Although the book focuses particularly on data obtained from a questionnaire study, SPSS/PC+ can also be used to analyze data from other types of study, including epidemiological research and laboratory work. Whichever of these types of study is involved, the process leading up to the writing of a report can be thought of as comprising four stages. As you work through this book you will be given instructions on how to:

- collect raw data and code them into numerical form
- enter numerical data into the computer and save them
- define these data and annotate them with descriptions and labels
- use the dataset to produce summaries and tables, and to perform statistical tests

The chapters are arranged in a 'tutorial sequence', but after you have worked through the book once, you are likely to refer back to particular chapters when you wish to perform particular statistical analyses or to transform your data in various ways.

Chapter 1 provides a very concise overview of SPSS/PC+.

Chapter 2 shows how the data which will form the raw material for SPSS/PC+ analysis (using the example of a questionnaire survey) can be collected and coded.

Chapter 3 introduces the SPSS/PC+ system as it is implemented on a microcomputer. Details are provided of how the system can be accessed and how the user may switch between levels of program.

Chapter 4 takes the reader, step-by-step, through the procedures involved in using the SPSS/PC+ editor, REVIEW, to enter data into a computer file.

Chapter 5 shows how a data definition file may be written. This provides definitions and descriptors so that the dataset is fully specified to the system.

Chapter 6 discusses in detail two of the many statistical procedures which form the heart of the SPSS/PC+ system. These simple procedures (FRE-

QUENCIES and CROSSTABS) allow the user to examine some basic features of the dataset.

Chapter 7 introduces two further statistical procedures, MEANS and T-TEST.

Chapter 8 shows the reader how data can be transformed, re-ordered and selected for analysis.

Chapter 9 examines two further types of file which can be used when conducting SPSS/PC+ analysis.

Chapter 10 provides a further exploration of the editing and other facilities offered by the editor program, REVIEW.

Chapter 11 provides details of another statistical procedure, CORRELATION, and introduces PLOT, a procedure which can be used to plot frequency data in graphical form.

Chapter 12 adds analysis of variance to the range of analyses which the user can perform with SPSS/PC+.

Chapter 13 provides a detailed guide showing how SPSS/PC+ can be used to perform a wide range of non-parametric statistical procedures using the commands NPAR TESTS and RANK.

Chapter 14 examines the concept of the reliability of a scale and describes how the RELIABILITY procedure can be used to examine characteristics of a scale.

Chapter 15 describes how two or more related files can be joined together using the JOIN command.

Chapter 16 describes alternative ways of coding and entering data. It also discusses how 'portable files' can be created and transported from one computer to another. Finally, there are brief descriptions of many aspects of the SPSS/PC+ system which are beyond the scope of this introductory book.

A number of appendices are also provided, and a guide to these is presented on pp. xiv–xvi.

One way to use this book most effectively would be to devise a small study of your own – to construct a questionnaire and carry out a pilot survey (following Chapters 1 and 2), to create a **data file** and an appropriate SPSS/PC+ **data definition file** (following Chapters 4 and 5) and then to proceed to a simple analysis of your data (Chapters 6 and 7) before exploring some of the more complex aspects of the system (Chapters 8 to 10). The later chapters introduce a range of further statistical procedures and describe ways in which the system can be used with optimal efficiency and flexibility. When you have worked through this book you will be in a good position to explore more complex forms of SPSS/PC+ analysis with the aid of the SPSS/PC+ manuals. As an alternative to performing your own survey you might choose to make use of the data from the survey which was performed to provide many of the worked examples in this book. This survey concerns gender attitudes and was conducted on British psychology students. The questionnaire was specially

constructed so that the data might allow various aspects of the SPSS/PC+ system to be explored. The original questionnaire is reproduced as *Appendix C*, and other appendices include the complete ‘codebook’ (used to code the data) and the data definition file used to make the data ‘meaningful’ to the system. The *complete* set of data is also reproduced so that students may, if they wish, work through many of the examples given in the book, following the various procedures on their own computer and exploring new ways of analyzing the data provided.

A Guide to the Appendices

Appendix A: Notes for Trainers

This appendix is intended for those who are involved in teaching others to use SPSS/PC+. The particular approach followed in this book is briefly explained, and there is some discussion about how the computer should be set up for the maximum ease of use by students who follow the instructions in the early chapters. The initial confusion which so often causes the beginner bewilderment and anxiety can be avoided if the trainer sets up the system to respond in ways that the student is led to expect, and if the trainer's teaching follows the strategies recommended in the text.

Appendix B: Using Word-processor and Other Files with SPSS/PC+

The use of the SPSS/PC+ editor, REVIEW, is discussed in a number of chapters of this book, and its use is recommended. However, it is sometimes more convenient to use another editing system (for example, a word-processing program) to produce files to be used by SPSS/PC+. This appendix describes how files produced by 'foreign' editor and word-processor systems can be made compatible with SPSS/PC+. Such files can later be re-edited using REVIEW. In addition, files produced by a wide range of database and spreadsheet programs can also be translated into SPSS/PC+ systems files.

The system automatically produces a listing file during each SPSS/PC+ session. This contains a copy of all of the tables, statistics, and other output produced during the session. *Appendix B* describes how this file can be edited with a word-processor.

Appendix C: The Gender Attitudes Survey: Questionnaire

This appendix introduces the 'Gender Attitudes Survey' and includes the original questionnaire used in the study.

Appendix D: The Gender Attitudes Survey: CODEBOOK

This reproduces the **codebook** which was produced to provide a guide to the **variables** represented within the Gender Attitudes Survey Questionnaire.

Appendix E: The Gender Attitudes Survey: DATA

This contains the full set of data collected from 50 (real!) respondents.

Appendix F: The Gender Attitudes Survey: DATA DEFINITION FILE

This appendix provides a full listing of the **data definition file** which was used to make the Gender Attitudes Survey data file accessible to SPSS/PC+.

Thus the four appendices, *Appendix C* to *Appendix F* constitute a complete package of materials relating to the Gender Attitudes Survey. Data from this small study are used throughout the book to illustrate the use of many different procedures. You might wish to create your own copy of the Gender Attitudes Survey dataset (by creating a copy of the data file and the data definition file) so that you can reproduce various analyses described throughout the text. You could also conduct your own original analyses of the dataset. However, do not attempt to create the relevant computer files until you have read the relevant chapters (Chapter 1 to Chapter 6) in this book.

If you are conducting your own survey you may wish to consult various aspects of the Gender Attitudes Survey package as a model for your own work.

Appendix G: Glossary of Statistical, Computing and SPSS/PC+ Terms

This provides a comprehensive glossary of the terms used in this book. It includes explanations of statistical terms, relevant computing terms and words which have a special function when used in SPSS/PC+. Whenever you encounter a word with which you are unfamiliar, or are unsure about the meaning of a term, you might find the help you need by consulting this Appendix. (If you are working within REVIEW, an 'on-line' glossary facility is also available – *see* Chapter 10.)

Appendix H: Types of File Used or Created by SPSS/PC+

This is a table that provides a brief description of features of the various types of **file** which you will encounter in using SPSS/PC+.

Appendix I: Variations between Different Versions of SPSS/PC+

If you are using a version of SPSS/PC+ other than the one specifically represented in this book (Version 4.0), you may wish to know how your version differs. This appendix provides an overview of some of the developments in SPSS/PC+ as it has evolved through successive versions. The main operating characteristics of the system have remained much the same through successive versions and are likely to remain similar in future versions. As the system has evolved, however, various enhancements have been added. These affect both the base module and the optional modules.

Appendix J: Brief guide to SPSS/PC+ Commands and Subcommands

This provides a brief synopsis of the SPSS/PC+ commands that are covered in detail elsewhere in this book. It will provide you with a handy reminder of the syntax of SPSS/PC+ commands, and of some of the subcommand options available within particular commands. It should be used only after you have become familiar with the various procedures and commands by reading through the relevant material in the main text.

Appendix K: Brief Guide to REVIEW Commands

This is another ‘handy reminder’ table, showing some of the key commands available when working within the SPSS/PC+ editor, REVIEW.

Acknowledgements

From the first edition

Many people, including students, colleagues, research workers and computer advisors have helped me in the task of creating this book. I approached some of them because of their expertise. Others were selected as 'naïve subjects' and introduced to SPSS for the first time. They worked through early drafts of the book, sitting before an Olivetti SP microcomputer, and made sure that what I said would happen *did* happen. Sometimes it did not, more often because of my mistake than theirs. Students were also asked to comment on difficulties they encountered in understanding the text, and some obviously relished the opportunity for role reversal which this offered. I amended the text in response to each critical note (including 'could be worded more elegantly' and 'are you sure you mean this?').

The 'experts' found bugs of a different species, and suitable pest control measures have been taken to try to eliminate these.

I am most grateful to all those who gave their time so generously. The principal helpers are listed below in alphabetical order. I will leave it to these individuals to judge whether I approached them for expert or 'grass roots' advice:

Annette Davies, Hilary Griggs, Rosemary Jenkins, Pamela Kenealy, Ray Kingdon, Mark McDermott and Martin Read

For the second edition

Preparation of the second edition was made easy by the fact that I have had the opportunity, over the five years since this book was first published, to help many readers (my own students, and the occasional correspondent) as they used the book to guide them through the analysis of their own projects. They, in turn, have helped me to discover the occasional inaccuracy and to recognize which aspects demanded clearer presentation. These students and correspondents are too numerous to mention individually, but I am grateful to them all.