

# STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES

## SECOND EDITION

STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES													
SECOND EDITION													
COUNT	I	COL	PCT	INONE	ONE	TWO	THREE	FOUR	FIVE	SEVEN	ROW	TOTAL	
					1.1	2.1	3.1	4.1	5.1	7.1			
1	2	1	0	1	0	1	0	1	0	1	0	1	2
	8.3	1	1	0.0	1	0.0	1	0.0	1	0.0	1	0.0	3.0
2	6	1	3	1	3	1	1	0	1	0	1	0	17
RS	0	1	3	1	33.3	1	0.0	1	0.0	1	0.0	1	25.8
3.	1	1	1	1	1	1	1	0	1	0	1	0	9
	5.0	1	1	0.0	1	0.0	1	0.0	1	0.0	1	0.0	13.6
4.	1	2	1	1	2	1	1	0	1	0	1	0	15
	20.	1	4.3	1	0	1	0	1	0.0	1	0.0	1	22.7
	1	3	1	1	3	1	1	0	1	1	1	1	14
IN	4	1	4	1	1	1	1	0	1	100.0	1	1	21.2
6.	1	1	1	1	1	1	1	1	1	0	1	0	6
GRAD	8.3	1	1	1	1	1	1	33.3	1	1	0.0	1	9.1
7.	1	0	1	0	1	0	1	1	1	1	1	1	3
PLUS	1	0.0	1	0.0	1	0.0	1	33.3	1	1	1	1	4.5
COLUMN	24	20	3	1	1	1	1	1	1	1	1	1	6
TOTAL	36.4	30.3	21.	4.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	10
R = 58.19641 WITH 36 DEGREES OF FREEDOM													
V = 0.38335													
CORRELATION COEFFICIENT = 0.68453													
S TAU R = 0.22786													
S TAU C = 0.20355													
0.29096													
0.24020													
RUN													
STUDY (CREATION DATE = 07/25/69) STUDY OF ORGANIZATION AND ACTIVITIES IN PENNSYLVANIA													

# SPSS

**STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES**  
**SECOND EDITION**

**NORMAN H. NIE**

Department of Political Science  
and  
National Opinion Research Center  
University Of Chicago

**C. HADLAI HULL**

Computation Center  
University of Chicago

**JEAN G. JENKINS**

National Opinion Research Center  
University of Chicago

**KARIN STEINBRENNER**

National Opinion Research Center  
University of Chicago

**DALE H. BENT**

Faculty of Business Administration  
and  
Computing Services  
The University of Alberta

**McGRAW-HILL  
BOOK COMPANY**

New York  
St. Louis  
San Francisco  
Auckland  
Düsseldorf  
Johannesburg  
Kuala Lumpur  
London  
Mexico  
Montreal  
New Delhi  
Panama  
Paris  
São Paulo  
Singapore  
Sydney  
Tokyo  
Toronto

No written materials describing the SPSS system may be produced or distributed without the express written permission of the copyright holders of both the SPSS programs and the SPSS manual.

**NOTE TO THE EIGHTH PRINTING:** Since the writing of this manual, which documents Release 6, SPSS Inc. has issued Releases 7 and 8, adding new facilities to the package and implementing significant improvements. New subprograms include MULT RESPONSE for multiple-response variables, NPAR TESTS for nonparametric data, RELIABILITY for computing reliability coefficients for multiple-item scales, SURVIVAL for life tables and associated graphs and statistics, and REPORT for generating publishable documents. Major improvements include faster operation or more precise calculations in most subprograms, enhanced data-management capabilities, and a more flexible facility for sorting cases. In addition, new facilities have been added for interaction between SPSS and SCSS, the SPSS Inc. conversational system. These and other changes are documented in SPSS UPDATE MANUAL, available from McGraw-Hill Book Company.

The statistical algorithms used in SPSS programming are now fully documented in SPSS STATISTICAL ALGORITHMS, available from SPSS Inc., Suite 3300, 444 North Michigan Avenue, Chicago, IL 60611. A POCKET GUIDE TO SPSS syntax, options, and statistics, also available from SPSS Inc., is a valuable companion to the larger manuals for users already familiar with SPSS.

SPSS and SCSS are the trademarks of SPSS, Inc. for its proprietary computer software. No materials describing such software may be produced or distributed without the written permission of the owners of the trademark and license rights in the software and the copyrights in the published materials.

Library of Congress Cataloging in Publication Data  
Nie, Norman H

SPSS: statistical package for the social sciences.

1. SPSS (Electronic computer system) 2. Electronic data processing—Social sciences. I. Title. II. Title: Statistical package for the social sciences.

HA33.N48 1975

029.7

74-16223

ISBN 0-07-046531-2 (pbk)

ISBN 0-07-046532-0 (hbk)

### **STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES**

Copyright © 1975, 1970 by McGraw-Hill, Inc. All rights reserved. Printed in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

11 12 13 14 15 16 17 18 WCWC 7 9 8 7 6 5 4 3 2 1 0

This book was set in Times Roman by Creative Book Services, division of McGregor & Werner, Inc. The editors were Kenneth J. Bowman and Matthew Cahill; the designer was Joseph Gillians; the production supervisor was Thomas J. LoPinto. New drawings were done by J & R Services, Inc.

Webcrafter, Inc. was printer and binder.

## PREFACE

*Statistical Package for the Social Sciences* (SPSS), the system of computer programs described in this volume, represents nearly a decade of systems design and programming and documentation on the part of the authors and others. When the first edition of this manual was published in 1970, SPSS was being used at approximately 60 installations. It is now being used at nearly 600 installations, including conversions to almost 20 different operating systems and computers. The current version of SPSS has almost double the amount of statistical procedures and data-management facilities as were documented in the first edition. Needless to say, we are thankful for the opportunity to continue developing SPSS and to serve an ever-increasing audience.

SPSS has been a success, much more so than the authors ever expected. We feel that the major reason for this success is that SPSS was developed through the close cooperation of three types of specialists: practicing social science researchers, computer scientists, and statisticians. At each stage in the development of SPSS we have attempted to satisfy these criteria:

- 1 That the statistical procedures be mathematically and statistically correct
- 2 That the program design and code be computationally efficient
- 3 That the logic and syntax of the system parallel the way in which social scientists approach data analysis
- 4 That the system provide statistical procedures and data-management facilities tailored to the particular needs of empirical social researchers

Without the contribution of experts in each field SPSS could not have effectively satisfied these goals.

Vying for importance in the success of SPSS has been our continuing concern for documentation. While many of our users are very sophisticated social scientists, they are not necessarily sophisticated statisticians and computer specialists, nor should they be expected to have the time or training necessary to master these fields. We feel that SPSS documentation must be comprehensive enough to enable students and researchers alike to use SPSS accurately



and efficiently. We hope that even someone who has no experience with computers will be able to run SPSS jobs successfully by using this manual. Furthermore, since the SPSS system is used in many social science methods courses, we feel it is important that the manual not only explain how to invoke the statistical procedures but also provide brief discussions of the various statistical techniques available. Of course, in attempting to satisfy these various needs the manual has grown to be a very large book indeed. However, we hope that the organization of the manual as well as the inclusion of an index will enable users to find what they need conveniently.

Finally, the availability within SPSS of a wide variety of data-management facilities and statistical procedures simplifies the process of data analysis. It avoids many difficult, time-consuming, and generally unrewarding tasks involved in using a variety of single-purpose computer programs, each with its own idiosyncratic control-card syntaxes and input data formats. The researcher thus spends less time as a data-preparation clerk and more time as a social scientist analyzing substantive results.

The development of SPSS began in 1965 at Stanford University as a result of the frustration felt by several of the authors in trying to serve the research and teaching needs of the political science department and the Institute for Political Studies by a library of single-purpose data analysis programs. We and the users suffered the annoyance of having to learn how to operate many different programs and endured the time-consuming task of transferring data and results between essentially noncompatible programs. Furthermore, the documentation for many of these programs ranged from cryptic to nonexistent. That and the fact that the programs were written in many different languages made them extremely difficult to maintain. Finally, while many statistical analysis programs were available, we were forced to write our own programs to perform even the simplest recoding, data transformation, file editing, and other routine house-keeping chores that are essential in social science data analysis. As a result of our experiences we began to design an integrated system that would automate the routine tasks of data processing and around which a series of statistical programs could be built. The basic data-modification, file-handling, and data-description facilities were programmed, and as time permitted statistical analysis procedures were added.

The system developed into SPSS as it was described in the first edition of the manual. Since that time we have followed our plan of incremental development. New facilities have gradually been added and many flaws in the original design have been corrected. SPSS users have submitted suggestions and even complete procedures which have been incorporated into the system. The update manuals that we published each year finally became so unwieldy that we were compelled to provide this second edition of the manual so that SPSS would once again be fully described in a single volume.

The current version of SPSS still has a number of deficiencies, but we feel that it meets a great many of the needs of social science data analysis. SPSS has always been considered an open-ended system, and we have again included a programmer's guide to SPSS (Appendix I) in the hopes that users who wish to add statistical procedures to the system will do so. Because of the design of SPSS, all future programs incorporated into the system can take complete advantage of the capabilities for file maintenance and data handling which exist in the package; this should be an incentive to those contemplating the addition of other statistical programs.

For some time now we have felt that one of the most serious drawbacks of the current SPSS system lies not in the lack of a wider variety of statistical techniques but in that it operates only as a batch program. We and others have had a great deal of success running SPSS through a text-editing and remote batch-entry system, enabling the user to prepare and enter all jobs from a remote terminal and to retrieve and print small jobs directly on the terminal. However, though this adds convenience, it in no way substitutes for a true conversational statistical analysis package. Consequently, the SPSS project staff is now developing a conversational version of SPSS.

The difference between performing analysis in a batch system and a conversational system can be compared to the difference between interacting with another person by letter and by phone. With a conversational system the researcher is brought into close contact with the data being analyzed; results from a statistical procedure are returned on the researcher's terminal in a

few seconds and may be used immediately to guide the next step of the analysis. This facilitates the iterative process of investigative research in which ideas are tested on the data, the results suggest new ideas and modifications of old ideas, the new ideas are then tested, and so forth. A train of thought which could take days (and many trips to a perhaps distant computer center) to explore in a batch system can be explored in a single interactive session. Thus the researcher's concentration and interest are focussed on the relationships being investigated. At the same time that a conversational system facilitates pursuing a single line of thought, the limited printing speed on remote terminals discourages the grand fishing expeditions and the production of excessive amounts of output that plague research in a batch environment. Finally, a good conversational system is somewhat self-teaching; it can guide the user in formulating the correct requests, and the almost instant feedback that facilitates the use in performing statistical analysis is also invaluable in identifying syntax errors and enabling the user to correct them immediately. In short, we feel that a conversational statistical analysis program could help revolutionize empirical social research.

Conversational SPSS, in addition to including conversational statistical analysis procedures, will feature interactive data- and file-definition capabilities, expanded data-transformation facilities, and ultimately the capacity to process hierarchical and other complex file structures. Although we are very enthusiastic about the prospects for conversational data analysis, we do not expect it to completely supplant batch processing. While a conversational system is ideal for investigative research, hypothesis testing, and rapid retrieval of a small number of specific statistics, it cannot be used efficiently for lengthy routine jobs and/or processing files that contain a very large number of data cases. Thus we expect that researchers will wish to use both the batch and conversational versions of SPSS. This will be reflected in the design of conversational SPSS as well as the evolution of batch SPSS—we are attempting to make the two systems useful as a team. Conversational SPSS will be capable of reading and producing batch-SPSS-format system files, and the control statement syntax for conversational SPSS will be as compatible as possible with batch-SPSS conventions.

We are all very excited about the future of SPSS, and we thank you, the users, for the opportunity to continue its development. Your support, enthusiasm, and constructive criticism have been both impetus and reward to us. We dedicate this book to you in the hope that it will make data analysis an easier task, and that in some small way, through your research, we are contributing to the development of the social sciences.

## ACKNOWLEDGMENTS

For the past five years the SPSS project has been located at the National Opinion Research Center at the University of Chicago. Throughout this critical period of our development NORC has provided us with financial, organizational, and intellectual support. This support has indeed been generous and has been a major factor responsible for our rapid growth and development. We can think of no other organization which would have provided a better home for us. In particular, we are grateful to Norman Bradburn and James Davis, the past and present directors of NORC, and Associate Director Shelby Orrell, without whose backing we might have floundered in the early years when our financial situation could only be described as precarious. We also wish to thank the University of Chicago Computation Center for donating computer time and allowing C. Hadlai Hull the flexibility to continue his work on the SPSS project.

The name of Patrick Bova, the librarian of NORC, is very familiar to many SPSS users. Mr. Bova has organized and directs the distribution of the SPSS system and all associated literature, and while maintaining our financial records, had thus far managed to keep us both solvent and honest. We are all very grateful for his efforts in what has become a very difficult, complicated, and time-consuming effort. The SPSS project is also very fortunate in the recent addition of ViAnn Beadle, who has joined the staff as user consultant. Ms. Beadle gives competent and conscientious attention to whatever questions and problems users may have. She has taken an enormous burden off our programming staff, providing us with the time for, among other things, the completion of this manual.

Special thanks must be given to Jae-On Kim, who has contributed greatly to SPSS over a period of years; Professor Kim helped design the factor analysis and analysis of variance routines, wrote several chapters of this manual, and regularly consults with us about statistical matters. We also wish to thank William R. Klecka, who contributed to the development of the discriminant function analysis procedure and wrote the associated chapter; William C. Mitchell, who designed the multiple-regression processing algorithms; and Bill Reynolds of the University of North Carolina, who designed the SCATTEGRAM procedure. For various improvements and corrections we are thankful to David Muxworthy of the University of Edinburgh, David Specht of Iowa State University, Jonathan B. Fry and Michael Klein of Akron University, Alfred J. Tuchfarber of the University of Cincinnati, Philip Burns of the University of Illinois, Cleve Moler of the University of Michigan, Harold W. Gugel of General Motors, and William L. McKeown of Stanford University. Martin L. Levin of Emory University and Kisun Han of DUALabs have developed SOS, the SPSS-Override System, which enables SPSS to process hierarchical files of the structure found in the Public Use Sample files.<sup>1</sup> Finally, we thank all those installations and users who have contributed suggestions and procedures. Even those which have not been implemented have provided us with ideas and thus have influenced the development of SPSS.

We are very appreciative of the work of Susan Hull, who prepared all of the more than 500 control-card and output examples appearing in this manual, and who implemented the integer version of the BREAKDOWN procedure. We must add that she performed these tasks as a volunteer, yet was as dedicated and painstaking as any other member of the staff. Another large contribution to the completion of this manual was made by Karin Donker, who typed and retyped until she nearly went mad, but who nevertheless cheerfully worked evenings and weekends when we had deadlines to meet.

We are indebted to the Vogelback Computing Center at Northwestern University for continued support of SPSS and in particular to James Tuccy of that center. He not only manages the CDC-6000 version of SPSS but also designed and programmed a number of the facilities described in this manual, including the CANCORR, DISCRIMINANT, ONEWAY, and T-TEST procedures. Space prevents us from individually mentioning each of the other SPSS conversions and the conversion managers, but we are grateful to them, as, we are sure, are their users.

We also express our appreciation to those at Stanford University who supported SPSS during its early development there: the political science department and the Institute of Political Studies, Professor Sidney Verba, Director of the Cross-National Program on Political and Social Change, and Professor Heinz Eulau and Kenneth Prewitt of the City Council Research Project. We thank the Stanford Computer Center for donating computer time. We are also grateful to the many students and researchers both at Stanford and at the University of Chicago who patiently used test versions of SPSS.

A system like SPSS is, of course, never developed in a vacuum, and we would like to acknowledge the contributions made to our general thinking by several previous systems: Data Text, developed at Harvard University, and BMD, developed at the University of California at Los Angeles, were especially significant. The factor analysis subprogram was adapted from a program developed at the University of Alberta. The Guttman Scale subprogram borrowed heavily from a program originally designed by Professor Ronald Anderson. The output format and some of the table statistics used in subprogram CROSSTABS were directly borrowed from the Data Text system. While we wish to acknowledge our great indebtedness to those who developed these programs, we alone are responsible for whatever errors or mistakes are present in their implementation.

Norman H. Nie  
C. Hadlai Hull  
Jean G. Jenkins  
Karin Steinbrenner  
Dale H. Bent

<sup>1</sup>SOS is currently operational for IBM 360/370 and Univac Series 70 computers. Further information may be obtained from Kisun Han of DUALabs, Suite 900, 1601 North Kent Street, Arlington, Virginia 22209.

## CONTENTS

<b>PREFACE</b>	<b>xxi</b>
<b>1</b>	
<b>AN INTRODUCTION TO COMPUTING WITH SPSS</b>	<b>1</b>
1.1 COMPUTERS AND THE PROCESS OF INDUCTIVE SOCIAL RE- SEARCH	2
1.1.1 Statistical Analysis on Computers: Use and Abuse	3
1.2 USING THE STATISTICAL CAPABILITIES OF SPSS	4
1.2.1 A Note on Levels of Measurement	4
1.2.2 Statistical Procedures in SPSS	6
1.3 AN OVERVIEW OF THE OPERATION OF SPSS	11
1.3.1 Sequencing Calculations	11
1.3.2 Entering and Processing Data	12
1.3.3 Subfiles	16
1.3.4 Missing Data	17
1.3.5 Recoding Data	17
1.3.6 Variable Transformations	17
1.3.7 Sampling, Selecting, and Weighting Data	18
1.3.8 Aggregating Data	18
1.3.9 File Modification and Management	18
1.3.10 Retrieval of Data from the System	19
1.3.11 Output of Results from the System	19
1.4 SUGGESTIONS FOR THE USE OF THIS TEXT	20



<b>2</b>		
<b>ORGANIZATION AND CODING OF DATA FOR INPUT INTO THE SPSS SYSTEM</b>		<b>21</b>
2.1	THE CASE AS THE UNIT OF ANALYSIS	21
2.2	CODING CONVENTIONS FOR DATA	21
2.2.1	Variables and Variable Types	22
2.2.2	Coding Missing Values	22
2.2.3	Organizing Data for Input	23
2.3	SUBFILES: THEIR FUNCTION AND STRUCTURE	25
2.3.1	Using Subfiles	26
2.3.2	Subfile Structure	26
2.4	RECORDING DATA ON MACHINE-READABLE MEDIA	26
2.5	LIMITATIONS ON INPUT DATA	27
<b>3</b>		
<b>SPSS CONTROL CARDS</b>		<b>29</b>
3.1	CONTROL-CARD PREPARATION: GENERAL FORMAT AND CONVENTIONS	29
3.1.1	Control Field	30
3.1.2	Specification Field	30
3.1.3	Summary of General Rules for Control-Card Preparation	32
3.2	NOTATION USED IN PRESENTING CONTROL-CARD FORMATS	33
<b>4</b>		
<b>DEFINING AN SPSS FILE: THE DATA-DEFINITION CARDS</b>		<b>35</b>
4.1	FILE NAME CARD	36
4.2	VARIABLE LIST CARD	36
4.3	SUBFILE LIST CARD	38
4.4	INPUT MEDIUM CARD	39
4.5	N OF CASES CARD	40
4.6	INPUT FORMAT CARD	41
4.6.1	Fixed-Column, Free-Field, or Binary Format	41
4.6.2	Format List for Data in Fixed-Column Format	42
4.6.3	Format Elements	42
4.6.4	VARIABLE LIST and Format List	46
4.6.5	INPUT FORMAT Card for Data in Binary Format	47
4.6.6	INPUT FORMAT Card for Data in FREEFIELD Format	47
4.7	DATA LIST CARD: AN ALTERNATE TO THE VARIABLE LIST AND INPUT FORMAT CARDS	49
4.7.1	DATA LIST for Data in Fixed-Column BCD Format	50
4.7.2	DATA LIST for Data in Binary Format	55
4.8	MISSING VALUES CARD	57
4.8.1	Entering Lists of Variables: General Conventions	58
4.8.2	Specifying the Missing Values	59
4.9	VALUE LABELS CARD	59
4.10	PRINT FORMATS CARD	61
4.11	VAR LABELS CARD	62

<b>4.12</b>	<b>REFERENCING ALL VARIABLES SIMULTANEOUSLY: THE KEYWORD ALL</b>	<b>63</b>
<b>4.13</b>	<b>RULES GOVERNING THE ORDER OF DATA-DEFINITION CARDS</b>	<b>63</b>

**5**

<b>CONTROLLING THE CALCULATIONS: THE TASK-DEFINITION CARDS</b>	<b>65</b>
--	-----------

<b>5.1</b>	<b>PROCEDURE CARDS</b>	<b>65</b>
<b>5.2</b>	<b>OPTIONS CARD</b>	<b>66</b>
<b>5.3</b>	<b>STATISTICS CARD</b>	<b>66</b>
<b>5.4</b>	<b>READ INPUT DATA CARD</b>	<b>67</b>
<b>5.5</b>	<b>RUN SUBFILES CARD</b>	<b>67</b>
<b>5.6</b>	<b>RAW OUTPUT UNIT CARD</b>	<b>69</b>
<b>5.7</b>	<b>SUMMARY OF TASK-DEFINITION CARDS</b>	<b>71</b>

**6**

<b>THE RUN CARDS</b>	<b>72</b>
----------------------	-----------

<b>6.1</b>	<b>RUN NAME CARD</b>	<b>72</b>
<b>6.2</b>	<b>TASK NAME CARD</b>	<b>73</b>
<b>6.3</b>	<b>FINISH CARD</b>	<b>73</b>
<b>6.4</b>	<b>PAGESIZE CARD</b>	<b>74</b>
<b>6.5</b>	<b>PRINT BACK CARD</b>	<b>74</b>
<b>6.6</b>	<b>COMMENT CARD</b>	<b>75</b>
<b>6.7</b>	<b>DOCUMENT CARD</b>	<b>76</b>
<b>6.8</b>	<b>NUMBERED CARD</b>	<b>76</b>

**7**

<b>GENERATING AND PROCESSING SPSS FILES: CARD ORDER AND DECK SETUP</b>	<b>78</b>
--	-----------

<b>7.1</b>	<b>PROCESSING FROM A RAW-INPUT-DATA FILE ON CARDS, TAPE, OR DISK: CARD ORDER AND DECK SETUP</b>	<b>78</b>
<b>7.2</b>	<b>GENERATING AND RETAINING SPSS SYSTEM FILES: THE SAVE FILE CARD</b>	<b>81</b>
<b>7.3</b>	<b>PROCESSING DATA FROM SPSS SYSTEM FILES</b>	<b>83</b>
<b>7.3.1</b>	<b>GET FILE Card</b>	<b>85</b>
<b>7.3.2</b>	<b>Deck Setup and Card Order for Processing Data from SPSS System Files</b>	<b>85</b>
<b>7.3.3</b>	<b>Modifying or Adding to Data-Definition Information During Processing Runs from SPSS System Files</b>	<b>86</b>

**8**

<b>RECODING AND VARIABLE TRANSFORMATION: THE DATA-MODIFICATION CARDS</b>	<b>89</b>
--	-----------

<b>8.1</b>	<b>RECODING VARIABLES: THE RECODE CARD</b>	<b>90</b>
<b>8.1.1</b>	<b>RECODE Specification List</b>	<b>90</b>
<b>8.1.2</b>	<b>Aids in Recoding: The Keywords LOWEST, HIGHEST, and ELSE</b>	<b>91</b>
<b>8.1.3</b>	<b>Differentiating Between Blanks and Zeros: The Keyword BLANK</b>	<b>92</b>
<b>8.1.4</b>	<b>Converting Variables from Alphanumeric to Numeric</b>	<b>92</b>

8.1.5	Temporary versus Permanent Recoding: the *RECODE Card	94
8.1.6	Example Usages and Deck Setup for RECODE and *RECODE Cards	94
8.1.7	Limitations on the RECODE card	96
8.2	COMPUTING VARIABLES BY MEANS OF ARITHMETIC EXPRESSIONS: THE COMPUTE CARD	96
8.2.1	Constructing Arithmetic Expressions on the COMPUTE Card	97
8.2.2	Example Usages of the COMPUTE Card	99
8.2.3	Temporary Variable Computations: The *COMPUTE Card	101
8.3	VARIABLE TRANSFORMATIONS WITH CONDITIONAL ASSIGNMENTS: THE IF CARD	101
8.3.1	Abbreviating the IF Statement: Implied Operands and Relational Operators	105
8.3.2	Temporary Variable Transformations: The *IF Card	
8.4	LIMITATIONS ON COMPUTE AND IF CARDS	107
8.5	CREATING ADDITIVE INDICES WITH THE COUNT CARD	107
8.5.1	COUNT Value List	108
8.5.2	General Format of the COUNT Card	108
8.5.3	Temporary Counts: The *COUNT Card	109
8.5.4	Limitations on the COUNT Card	110
8.6	STARRED [*] AND NONSTARRED VERSIONS OF RECODE, COMPUTE, IF, AND COUNT CARDS: TEMPORARY VERSUS PERMANENT DATA MODIFICATION	110
8.7	ALLOCATING CORE STORAGE SPACE FOR DATA MODIFICATIONS: THE ALLOCATE CARD	111
8.7.1	Calculating the Amount of TRANSPACE Required	111
8.7.2	Preparation and Use of the ALLOCATE Card	113
8.8	MISSING-DATA SPECIFICATIONS AND OTHER DATA-DEFINITION INFORMATION FOR TRANSFORMED VARIABLES	115
8.8.1	Treatment of Missing Data in Variable Transformations: The ASSIGN MISSING Card	115
8.8.2	An Alternative Treatment of Missing Data in Variable Transformations: The MISSING VALUES Card	119
8.8.3	Initializing Variables Created with IF Statements	120
8.8.4	Inserting Other Data-Definition Information for Transformed Variables	121
8.9	REDUCING CONTROL-CARD PREPARATION WITH THE REPEAT FACILITY: DO REPEAT AND END REPEAT CARDS	121
8.9.1	Valid Control Cards Between DO REPEAT and END REPEAT	123
8.9.2	Limitations on the REPEAT Facility	124
8.10	DECK SETUP AND CARD ORDER FOR DATA-MODIFICATION CARDS	126

<b>9</b>		
<b>DATA-SELECTION CARDS</b>		<b>127</b>
9.1	GENERATING A RANDOM SAMPLE FROM A FILE: SAMPLE AND *SAMPLE CARDS	127
9.2	SELECTING CASES FROM A FILE: SELECT IF AND *SELECT IF CARDS	128
9.3	WEIGHTING CASES IN A FILE: WEIGHT AND *WEIGHT CARDS	128
9.4	DECK SETUP AND CARD ORDER FOR DATA-SELECTION CARDS	131
<b>10</b>		
<b>RETRIEVING DATA AND DESCRIPTIVE INFORMATION FROM SPSS SYSTEM FILES:</b>		
<b>THE FILE RETRIEVAL (LIST AND WRITE) CARDS</b>		<b>133</b>
10.1	LIST FILEINFO CARD	134
10.2	LIST CASES CARD	137
10.3	PRODUCING RAW-OUTPUT-DATA FILES: THE WRITE CASES CARD	139
10.3.1	Variable List for the WRITE CASES Card	140
10.3.2	Format Specification for the WRITE CASES Card	140
10.3.3	A Note on Writing Cases from Weighted Files	143
10.3.4	Listwise Deletion of Missing Data: WRITE CASES Option 1	143
10.3.5	Example Deck Setup and Card Order for the WRITE CASES Procedure	144
10.4	WRITE FILEINFO CARD	145
10.4.1	Variable List and Data-Definition Keywords for the WRITE FILEINFO Card	145
10.4.2	Output Option Keywords for the WRITE FILEINFO Card	146
10.4.3	Using WRITE FILEINFO Cards to Regenerate a System File	148
10.4.4	Producing a VARIABLE LIST with the WRITE FILEINFO Card for Use on EDIT Runs	148
<b>11</b>		
<b>ADDING TO, DELETING FROM, AND REARRANGING DATA IN SPSS SYSTEM FILES:</b>		
<b>THE FILE-MODIFICATION CARDS</b>		<b>149</b>
11.1	DELETING AND RETAINING VARIABLES: DELETE VARS AND KEEP VARS CARDS	149
11.2	ADDING VARIABLES TO AN EXISTING SPSS SYSTEM FILE: ADD VARIABLES AND ADD DATA LIST CARDS	151
11.2.1	Adding Variables with an ADD VARIABLES Card	151
11.2.2	Adding Variables with an ADD DATA LIST Card	153
11.3	MERGING SPSS SYSTEM FILES: THE MERGE FILES CARD	155
11.4	ADDING CASES AND SUBFILES TO EXISTING SPSS SYSTEM FILES: ADD CASES AND ADD SUBFILES CARDS	157
11.4.1	ADD CASES Card	157

11.4.2	ADD SUBFILES Card	157
11.4.3	Setting Up Runs Which Add Cases or Subfiles	158
11.5	DELETING SUBFILES WITH THE DELETE SUBFILES CARD	159
11.6	DELETING CASES: ANOTHER USE OF THE SELECT IF CARD	159
11.7	REORDERING THE SEQUENCE OF VARIABLES IN AN SPSS SYSTEM FILE: THE REORDER VARS CARD	160
11.8	CHANGING THE SEQUENCE OF CASES IN AN SPSS SYSTEM FILE: THE SORT CASES CARD	162
11.8.1	Handling Missing Values While Sorting Cases	164
11.8.2	Sorting Alphanumeric Variables: Don't	164
11.8.3	Determining the Number of Cases When Sorting for New Subfiles	164
11.8.4	Deck Location of the SORT CASES Control Card	164
11.8.5	Special Operating System Control Statements for Use with SORT CASES	165
11.9	CREATING OR REPLACING THE SUBFILE STRUCTURE WITH THE SUBFILE LIST CARD	165
11.10	SUMMARY OF FILE-MODIFICATION CARDS AND THEIR DECK PLACEMENT	165
 <b>12</b>		
<b>CREATING, RETRIEVING, AND MANIPULATING FILES WITH MORE THAN 500 VARIABLES: GET ARCHIVE, SAVE ARCHIVE, AND LIST ARCHINFO CARDS</b>		<b>167</b>
12.1	ACCESSING INPUT FILES WITH THE GET ARCHIVE CARD	168
12.2	PERMANENTLY SAVING AN ARCHIVE FILE WITH THE SAVE ARCHIVE CARD	169
12.3	USING THE ARCHIVE FACILITY: EXAMPLES	171
12.3.1	Creating an Archive File Which Includes New Variables Produced by Data Transformations	171
12.3.2	Creating an Archive File by Merging Several Input Files	171
12.3.3	Creating an Archive File Which Includes New Variables Input Via the ADD VARIABLES Facility	172
12.4	SAVING THE ACTIVE VARIABLES ON A STANDARD SPSS SYSTEM FILE: THE SAVE FILE CARD	172
12.5	USING ARCHIVE FILES FOR ROUTINE STATISTICAL PROCESSING	173
12.6	RETRIEVING DESCRIPTIVE INFORMATION FROM ARCHIVE INPUT FILES: THE LIST ARCHINFO CARD	173
12.7	LIMITATIONS FOR THE ARCHIVE FACILITY	174
 <b>13</b>		
<b>EDITING SPSS CONTROL-CARD DECKS: THE EDIT FACILITY</b>		<b>176</b>
13.1	EDITING DECKS WHICH ACCESS A RAW-INPUT-DATA FILE	177
13.2	EDITING DECKS WHICH ACCESS AN SPSS SYSTEM FILE	177
13.3	CORE STORAGE SPACE REQUIRED FOR EDIT RUNS	177
13.4	PRINTED OUTPUT FROM EDIT RUNS	178
13.5	CONTROL-CARD DECK ERRORS THAT THE EDIT FACILITY WILL NOT DETECT	178

<b>13.6</b>	<b>EXAMPLE OUTPUT FROM EDIT RUNS</b>	<b>178</b>
<b>14</b>		
<b>DESCRIPTIVE STATISTICS AND ONE-WAY FREQUENCY DISTRIBUTIONS</b>		<b>181</b>
<b>14.1</b>	<b>A BRIEF INTRODUCTION TO DESCRIPTIVE STATISTICS</b>	<b>182</b>
<b>14.2</b>	<b>SUBPROGRAM CONDESCRIPTIVE: DESCRIPTIVE STATISTICS FOR CONTINUOUS VARIABLES</b>	<b>185</b>
14.2.1	CONDESCRIPTIVE Procedure Card	186
14.2.2	Processing Options for Subprogram CONDESCRIPTIVE	189
14.2.3	Statistics Available for Subprogram CONDESCRIPTIVE	189
14.2.4	Program Limitations for Subprogram CONDESCRIPTIVE	190
14.2.5	Examples of the Use of Subprogram CONDESCRIPTIVE	191
<b>14.3</b>	<b>SUBPROGRAM FREQUENCIES: ONE-WAY FREQUENCY DIS- TRIBUTIONS WITH DESCRIPTIVE STATISTICS</b>	<b>194</b>
14.3.1	Selecting Between General and Integer Operating Modes	194
14.3.2	FREQUENCIES Procedure Card	195
14.3.3	Printed Output from the FREQUENCIES Procedure	197
14.3.4	Options Available for Subprogram FREQUENCIES	200
14.3.5	Statistics Available for Subprogram FREQUENCIES	201
14.3.6	Limitations for Subprogram FREQUENCIES	201
<b>15</b>		
<b>PRODUCING DESCRIPTIVE STATISTICS FOR AGGREGATED DATA FILES: SUBPROGRAM AGGREGATE</b>		<b>203</b>
<b>15.1</b>	<b>SPECIFYING AGGREGATION GROUPS</b>	<b>205</b>
15.1.1	Required Structure of the Input File	205
<b>15.2</b>	<b>OBTAINING THE OUTPUT FILE OF AGGREGATED DESCRIPTIVE STATISTICS</b>	<b>205</b>
<b>15.3</b>	<b>AGGREGATE PROCEDURE CARD</b>	<b>206</b>
15.3.1	GROUPVARS= Variable List	207
15.3.2	VARIABLES= List	207
15.3.3	AGGSTATS= Keyword List	208
15.3.4	Providing for Missing Values with the RMISS=value/	209
15.3.5	Repetitions of the VARIABLES=, AGGSTATS=, and RMISS= Specifications	210
15.3.6	Outputting the Actual Values of the Grouping Variables with Option 3	211
<b>15.4</b>	<b>CONTENT AND FORMAT OF AGGREGATED OUTPUT DATA FILES</b>	<b>211</b>
15.4.1	Types of Aggregated Output Data Files	212
15.4.2	Contents of Each Case on an Aggregated Output File	212
<b>15.5</b>	<b>PRINTED OUTPUT FROM SUBPROGRAM AGGREGATE</b>	<b>213</b>
<b>15.6</b>	<b>OPTIONS AVAILABLE FOR SUBPROGRAM AGGREGATE</b>	<b>213</b>
<b>15.7</b>	<b>STATISTICS AVAILABLE FOR SUBPROGRAM AGGREGATE</b>	<b>214</b>
<b>15.8</b>	<b>PROGRAM LIMITATIONS FOR SUBPROGRAM AGGREGATE</b>	<b>215</b>
<b>15.9</b>	<b>EXAMPLE DECK SETUP AND OUTPUT FOR SUBPROGRAM AGGREGATE</b>	<b>215</b>



<b>16</b>		
<b>CONTINGENCY TABLES AND RELATED MEASURES OF ASSOCIATION: SUBPROGRAM CROSSTABS</b>		<b>218</b>
<b>16.1</b>	<b>AN INTRODUCTION TO CROSSTABULATION</b>	<b>218</b>
16.1.1	Crosstabulation Tables	219
16.1.2	Summary Statistics for Crosstabulations	222
<b>16.2</b>	<b>SUBPROGRAM CROSSTABS: TWO-WAY TO N-WAY CROSS-TABULATION TABLES AND RELATED STATISTICS</b>	<b>230</b>
16.2.1	Components of the CROSSTABS Procedure Card	231
16.2.2	TABLES= List	231
16.2.3	VARIABLES= List for Integer Mode	234
<b>16.3</b>	<b>FORMAT OF THE CROSSTABS PROCEDURE CARD</b>	<b>236</b>
<b>16.4</b>	<b>PRINTED OUTPUT FROM SUBPROGRAM CROSSTABS</b>	<b>237</b>
16.4.1	A Note on Value Labels for CROSSTABS	239
16.4.2	Including Missing Values Only in Tables	240
<b>16.5</b>	<b>OPTIONS AVAILABLE FOR SUBPROGRAM CROSSTABS: THE OPTIONS CARD</b>	<b>241</b>
<b>16.6</b>	<b>STATISTICS AVAILABLE FOR SUBPROGRAM CROSSTABS: THE STATISTICS CARD</b>	<b>242</b>
<b>16.7</b>	<b>PROGRAM LIMITATIONS FOR SUBPROGRAM CROSSTABS</b>	<b>243</b>
16.7.1	Program Limitations for Subprogram CROSSTABS, General Mode	243
16.7.2	Program Limitations for Subprogram CROSSTABS, Integer Mode	244
<b>16.8</b>	<b>EXAMPLE DECK SETUPS FOR SUBPROGRAM CROSSTABS</b>	<b>245</b>
<b>17</b>		
<b>DESCRIPTION OF SUBPOPULATIONS AND MEAN DIFFERENCE TESTING: SUBPROGRAMS BREAKDOWN AND T-TEST</b>		<b>249</b>
<b>17.1</b>	<b>SUBPROGRAM BREAKDOWN</b>	<b>249</b>
17.1.1	BREAKDOWN Operating Modes: Integer and General	250
17.1.2	TABLES= List	252
17.1.3	VARIABLES= List for Integer Mode	254
17.1.4	Format of the BREAKDOWN Procedure Card	255
17.1.5	Options Available for Subprogram BREAKDOWN: The OPTIONS Card	257
17.1.6	Statistics for Subprogram BREAKDOWN: One-Way Analysis of Variance and Test of Linearity	257
17.1.7	Program Limitations for Subprogram BREAKDOWN	261
17.1.8	Example Deck Setup for Subprogram BREAKDOWN	262
17.1.9	BREAKDOWN Tables Printed in Crosstabular Form: The CROSSBREAK Facility	264
<b>17.2</b>	<b>SUBPROGRAM T-TEST: COMPARISON OF SAMPLE MEANS</b>	<b>267</b>
17.2.1	Introduction to the T-TEST of Significance	267
17.2.2	The T-TEST Procedure Card	271
17.2.3	Options and Statistics for Subprogram T-TEST	273
17.2.4	Program Limitations for Subprogram T-TEST	273

17.2.5	Example Deck Setups and Output for Subprogram T-TEST	274
<b>18</b>	<b>BIVARIATE CORRELATION ANALYSIS: PEARSON CORRELATION, RANK-ORDER CORRELATION, AND SCATTER DIAGRAMS</b>	<b>276</b>
<b>18.1</b>	INTRODUCTION TO CORRELATION ANALYSIS	276
<b>18.2</b>	SUBPROGRAM PEARSON CORR: PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENTS	280
18.2.1	PEARSON CORR Procedure Card	281
18.2.2	Options Available for Subprogram PEARSON CORR	283
18.2.3	Statistics Available for Subprogram PEARSON CORR	285
18.2.4	Program Limitations for Subprogram PEARSON CORR	285
18.2.5	Sample Deck Setup and Output for Subprogram PEARSON CORR	286
<b>18.3</b>	SUBPROGRAM NONPAR CORR: SPEARMAN AND/OR KENDALL RANK-ORDER CORRELATION COEFFICIENTS	288
18.3.1	NONPAR CORR Procedure Card	290
18.3.2	Options Available for Subprogram NONPAR CORR	291
18.3.3	Statistics Available for Subprogram NONPAR CORR	291
18.3.4	Program Limitations for Subprogram NONPAR CORR	291
18.3.5	Sample Deck Setup and Output for Subprogram NONPAR CORR	292
<b>18.4</b>	SUBPROGRAM SCATTERGRAM: SCATTER DIAGRAM OF DATA POINTS AND SIMPLE REGRESSION	293
18.4.1	SCATTERGRAM Procedure Card	294
18.4.2	Options Available for Subprogram SCATTERGRAM	296
18.4.3	Statistics Available for Subprogram SCATTERGRAM	297
18.4.4	Program Limitations for Subprogram SCATTERGRAM	297
18.4.5	Sample Deck Setup and Output for Subprogram SCATTERGRAM	298
<b>19</b>	<b>PARTIAL CORRELATION: SUBPROGRAM PARTIAL CORR</b>	<b>301</b>
<b>19.1</b>	INTRODUCTION TO PARTIAL CORRELATION ANALYSIS	302
<b>19.2</b>	PARTIAL CORR PROCEDURE CARD	305
19.2.1	Correlation List	306
19.2.2	Control List	306
19.2.3	Order Value(s)	307
<b>19.3</b>	SPECIAL CONVENTIONS FOR MATRIX INPUT FOR SUBPROGRAM PARTIAL CORR	308
19.3.1	Requirements on the Form and Format for Correlation Matrices	308
19.3.2	Methods for Specifying the Order of Variables on Input Matrices	309
19.3.3	Specifying the Order of Variables on Input Matrices by the Partial List: The Default Method	309
19.3.4	Specifying the Order of Variables on Input Matrices by the Variable List Card: Option 6	310

19.3.5	SPSS Control Cards Required for Matrix Input	311
19.4	OPTIONS AVAILABLE FOR SUBPROGRAM PARTIAL CORR	312
19.5	STATISTICS AVAILABLE FOR SUBPROGRAM PARTIAL CORR	315
19.6	PROGRAM LIMITATIONS FOR SUBPROGRAM PARTIAL CORR	315
19.7	EXAMPLE DECK SETUPS FOR SUBPROGRAM PARTIAL CORR	316
 <b>20</b>		
<b>MULTIPLE REGRESSION ANALYSIS: SUBPROGRAM REGRESSION</b>		<b>320</b>
20.1	INTRODUCTION TO MULTIPLE REGRESSION	321
20.1.1	Simple Bivariate Regression	323
20.1.2	Extension to Multiple Regression	328
20.2	REGRESSION PROCEDURE CARD	342
20.2.1	VARIABLES List	343
20.2.2	REGRESSION Design Statement	343
20.3	SUMMARY OF PROCEDURE CARDS	348
20.4	SPECIAL CONVENTIONS FOR MATRIX INPUT WITH SUBPROGRAM REGRESSION	349
20.4.1	Format of Input Correlation Matrices	349
20.4.2	Format and Position of Input Means and Standard Devia- tions	350
20.4.3	Methods for Specifying the Number and Order of Vari- ables on Input Correlation Matrices	350
20.4.4	Control Cards Required to Enter Matrices	351
20.5	SPECIAL CONVENTIONS FOR HANDLING RESIDUAL OUTPUT	351
20.6	OPTIONS AVAILABLE FOR SUBPROGRAM REGRESSION	352
20.7	STATISTICS AVAILABLE FOR SUBPROGRAM REGRESSION	355
20.8	PROGRAM LIMITATIONS FOR SUBPROGRAM REGRESSION	356
20.9	PRINTED OUTPUT FROM SUBPROGRAM REGRESSION	358
20.10	EXAMPLE DECK SETUPS AND OUTPUT FOR SUBPROGRAM REGRESSION	360
 <b>21</b>		
<b>SPECIAL TOPICS IN GENERAL LINEAR MODELS</b>		<b>368</b>
21.1	NONLINEAR RELATIONSHIPS	368
21.1.1	Data Transformation	369
21.1.2	Examining Polynomial Trends	371
21.1.3	Interaction Terms	372
21.2	REGRESSION WITH DUMMY VARIABLES	373
21.2.1	Dummy Variables: Coding and Interpretation	374
21.2.2	Dummy Variable Regression with Two or More Categori- cal Variables	377
21.3	PATH ANALYSIS AND CAUSAL INTERPRETATION	383
21.3.1	Principals of Path Analysis	384
21.3.2	Interpretation of Path Analysis Results	387
21.3.3	Statistical Inference	392
21.3.4	Standardized versus Unstandardized Coefficients	394