

The background of the cover features a 3D grid of cubes. The cubes are arranged in a pattern that recedes into the distance. The color of the cubes transitions from a bright orange at the top to a lighter yellow at the bottom. The surface they sit on is white and highly reflective, creating clear, inverted reflections of each cube. The lighting is soft and even, highlighting the three-dimensional nature of the objects.

Sixth edition

Dennis Howitt &
Duncan Cramer

Introduction to

SPSS in Psychology

Introduction to SPSS in Psychology

Visit the *Introduction to SPSS in Psychology, sixth edition* Companion Website at www.pearsoned.co.uk/howitt to find valuable student learning material, including:

- Learning objectives for each chapter
- Datasets and syntax
- Annotated links to relevant sites on the web

PEARSON

At Pearson, we have a simple mission: to help people make more of their lives through learning.

We combine innovative learning technology with trusted content and educational expertise to provide engaging and effective learning experiences that serve people wherever and whenever they are learning.

From classroom to boardroom, our curriculum materials, digital learning tools and testing programmes help to educate millions of people worldwide – more than any other private enterprise.

Every day our work helps learning flourish, and wherever learning flourishes, so do people.

To learn more, please visit us at www.pearson.com/uk



Introduction to SPSS in Psychology

For Version 22 and earlier

Sixth edition

Dennis Howitt, Loughborough University

Duncan Cramer, Loughborough University

PEARSON

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Auckland • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • São Paulo • Mexico City • Madrid • Amsterdam • Munich • Paris • Milan

Pearson Education Limited

Edinburgh Gate
Harlow CM20 2JE
United Kingdom
Tel: +44 (0)1279 623623
Web: www.pearson.com/uk

First published 1997 (print)
Second edition published 2000 (print)
Second revised edition 2002 (print)
Third edition 2005 (print)
Fourth edition 2008 (print)
Fifth edition 2011 (print)
This edition published 2014 (print and electronic)

© Prentice Hall 1997 (print)
© Pearson Education Limited 2000, 2002, 2005, 2008, 2011 (print)
© Pearson Education Limited 2014 (print and electronic)

The rights of Dennis Howitt and Duncan Cramer to be identified as authors of this work have been asserted by them in accordance with the Copyright, Designs and Patents Act 1988.

The print publication is protected by copyright. Prior to any prohibited reproduction, storage in a retrieval system, distribution or transmission in any form or by any means, electronic, mechanical, recording or otherwise, permission should be obtained from the publisher or, where applicable, a licence permitting restricted copying in the United Kingdom should be obtained from the Copyright Licensing Agency Ltd, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

The ePublication is protected by copyright and must not be copied, reproduced, transferred, distributed, leased, licensed or publicly performed or used in any way except as specifically permitted in writing by the publishers, as allowed under the terms and conditions under which it was purchased, or as strictly permitted by applicable copyright law. Any unauthorised distribution or use of this text may be a direct infringement of the author's and the publishers' rights and those responsible may be liable in law accordingly.

All trademarks used herein are the property of their respective owners. The use of any trademark in this text does not vest in the author or publisher any trademark ownership rights in such trademarks, nor does the use of such trademarks imply any affiliation with or endorsement of this book by such owners.

The screenshots in this book are copyright IBM SPSS Inc. IBM SPSS is a registered trademark and the other product names are registered trademarks of IBM SPSS Inc.

ISBN: 978-1-292-00069-5 (print)
978-1-292-00070-1 (CSeTxt)
978-1-292-00073-2 (PDF)

British Library Cataloguing-in-Publication Data

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

Library of Congress Cataloging-in-Publication Data

A catalog record for the print edition is available from the Library of Congress

10 9 8 7 6 5 4 3 2 1
18 17 16 15 14

Print edition typeset in 9.5/12pt Sabon LT Std by 35
Printed by DZS Grafic, Slovenia

NOTE THAT ANY PAGE CROSS REFERENCES REFER TO THE PRINT EDITION

Summary of contents

<i>Introduction</i>	<i>xxvi</i>
<i>Acknowledgements</i>	<i>xxix</i>
<i>Key differences between SPSS Statistics 22 and earlier versions</i>	<i>xxxii</i>
Part 1 Introduction to SPSS Statistics	1
1 A brief introduction to statistics	3
2 Basics of SPSS Statistics data entry and statistical analysis	17
Part 2 Descriptive statistics	31
3 Describing variables: Tables	33
4 Describing variables: Diagrams	41
5 Describing variables numerically: Averages, variation and spread	57
6 Shapes of distributions of scores	66
7 Standard deviation: The standard unit of measurement in statistics	77
8 Relationships between two or more variables: Tables	85
9 Relationships between two or more variables: Diagrams	94
10 Correlation coefficients: Pearson's correlation and Spearman's rho	106
11 Regression: Prediction with precision	121
Part 3 Significance testing and basic inferential tests	133
12 Standard error	135
13 The <i>t</i> -test: Comparing two samples of correlated/related/paired scores	141
14 The <i>t</i> -test: Comparing two groups of unrelated/unrelated scores	151
15 Confidence intervals	159
16 Chi-square: Differences between unrelated samples of frequency data	163
17 McNemar's test: Differences between related samples of frequency data	178
18 Ranking tests for two groups: Non-parametric statistics	185
19 Ranking tests for three or more groups: Non-parametric statistics	196
Part 4 Analysis of variance	207
20 The variance ratio test: Using the <i>F</i> -ratio to compare two variances	209
21 Analysis of variance (ANOVA): Introduction to the one-way unrelated or uncorrelated ANOVA	215
22 Analysis of variance for one-way correlated scores or repeated measures	223
23 Two-way analysis of variance for unrelated/unrelated scores	232
24 Multiple comparison in ANOVA	245
25 Analysis of variance for two-way correlated scores or repeated measures	253
26 Two-way mixed analysis of variance (ANOVA)	266
27 Analysis of covariance (ANCOVA)	276
28 Multivariate analysis of variance (MANOVA)	287
29 Discriminant (function) analysis (for MANOVA)	297

Part 5	More advanced correlational statistics	309
30	Partial correlation	311
31	Factor analysis	318
32	Item reliability and inter-rater agreement	330
33	Stepwise multiple regression	342
34	Simultaneous or standard multiple regression	353
35	Simple mediational analysis	366
36	Hierarchical multiple regression	377
37	Moderator analysis with continuous predictor variables	387
Part 6	Advanced qualitative or nominal techniques	401
38	Log-linear analysis	403
39	Multinomial logistic regression	413
40	Binomial logistic regression	429
Part 7	Data handling procedures	441
41	Reading ASCII or text files into the Data Editor	443
42	Missing values	450
43	Recoding values	458
44	Computing a scale score with no missing values	466
45	Computing a scale score with some values missing	475
46	Computing a new group variable from existing group variables	482
47	Selecting cases	489
48	Samples and populations: Generating a random sample	495
49	Inputting a correlation matrix	500
50	Checking accuracy of data input	505
Part 8	Other statistical procedures	515
51	Statistical power analysis: Sample size estimation	517
52	Meta-analysis	527
	<i>Appendix: Some other statistics in SPSS Statistics</i>	<i>537</i>
	<i>Glossary</i>	<i>539</i>
	<i>Index</i>	<i>547</i>

Contents

<i>Introduction</i>	<i>xxvi</i>
<i>Acknowledgements</i>	<i>xxix</i>
<i>Key differences between SPSS Statistics 22 and earlier versions</i>	<i>xxx</i>

Part 1 Introduction to SPSS Statistics 1

1	A brief introduction to statistics	3
	<i>Overview</i>	3
1.1	Basic statistical concepts essential in SPSS Statistics analyses	4
1.2	Basic research designs: comparative versus correlational designs	4
1.3	The different types of variables in statistics	7
1.4	Descriptive and inferential statistics compared	9
1.5	Related versus unrelated designs	11
1.6	Quick summaries of statistical analyses	12
1.7	Which procedure or test to use	13
2	Basics of SPSS Statistics data entry and statistical analysis	17
	<i>Overview</i>	17
2.1	What is SPSS Statistics?	18
2.2	Accessing SPSS Statistics	19
2.3	Entering data	20
2.4	Moving within a window with the mouse	21
2.5	Moving within a window using the keyboard keys with the mouse	21
2.6	Saving data to disk	21
2.7	Opening up a data file	23
2.8	Using 'Variable View' to create and label variables	24
2.9	More on 'Data View'	27
2.10	A simple statistical calculation with SPSS	29
2.11	SPSS Statistics output	29
	<i>Summary of SPSS Statistics steps for a statistical analysis</i>	30

Part 2 Descriptive statistics 31

3	Describing variables: Tables	33
	<i>Overview</i>	33
3.1	What are tables?	34
3.2	When to use tables	34
3.3	When not to use tables	35
3.4	Data requirements for tables	35
3.5	Problems in the use of tables	35
3.6	The data to be analysed	35
3.7	Entering summarised categorical or frequency data by weighting	36
3.8	Percentage frequencies	38
3.9	Interpreting the output	39
	<i>Reporting the output</i>	39
	<i>Summary of SPSS Statistics steps for frequency tables</i>	39
4	Describing variables: Diagrams	41
	<i>Overview</i>	41
4.1	What are diagrams?	42
4.2	When to use diagrams	43
4.3	When not to use diagrams	43
4.4	Data requirements for diagrams	43
4.5	Problems in the use of diagrams	44
4.6	The data to be analysed	44
4.7	Entering summarised categorical or frequency data by weighting	45
4.8	Pie-diagram of category data	47
4.9	Adding labels to the pie diagram and removing the legend and label	49
4.10	Changing the colour of a pie diagram slice to a black and white pattern	51
4.11	Bar chart of category data	53
4.12	Histograms	54
	<i>Summary of SPSS Statistics steps for charts</i>	56
5	Describing variables numerically: Averages, variation and spread	57
	<i>Overview</i>	57
5.1	What are averages, variation and spread?	58
5.2	When to use averages, variation and spread	60
5.3	When not to use averages, variation and spread	60
5.4	Data requirements for averages, variation and spread	60
5.5	Problems in the use of averages, variation and spread	61
5.6	The data to be analysed	61
5.7	Entering the data	62

5.8	Mean, median, mode, standard deviation, variance and range	62
5.9	Interpreting the output	63
5.10	Other features	63
	<i>Reporting the output</i>	64
	<i>Summary of SPSS Statistics steps for descriptive statistics</i>	65
6	Shapes of distributions of scores	66
	<i>Overview</i>	66
6.1	What are the different shapes of scores?	67
6.2	When to use histograms and frequency tables of scores	70
6.3	When not to use histograms and frequency tables of scores	71
6.4	Data requirements for histograms and frequency tables of scores	71
6.5	Problems in the use of histograms and frequency tables of scores	71
6.6	The data to be analysed	72
6.7	Entering the data	72
6.8	Frequency tables	73
6.9	Interpreting the output	73
	<i>Reporting the output</i>	74
6.10	Histograms	74
6.11	Interpreting the output	75
	<i>Reporting the output</i>	76
	<i>Summary of SPSS Statistics steps for frequency distributions</i>	76
7	Standard deviation: The standard unit of measurement in statistics	77
	<i>Overview</i>	77
7.1	What is standard deviation?	78
7.2	When to use standard deviation	79
7.3	When not to use standard deviation	80
7.4	Data requirements for standard deviation	80
7.5	Problems in the use of standard deviation	80
7.6	The data to be analysed	80
7.7	Entering the data	81
7.8	Standard deviation	81
7.9	Interpreting the output	82
7.10	Z-scores	82
7.11	Other features	83
	<i>Reporting the output</i>	83
	<i>Summary of SPSS Statistics steps for standard deviation</i>	84
8	Relationships between two or more variables: Tables	85
	<i>Overview</i>	85
8.1	What tables are used to show relationships between variables?	86

8.2	When to use tables to show relationships between variables	88
8.3	When not to use tables to show relationships between variables	88
8.4	Data requirements for tables to show relationships between variables	88
8.5	Problems in the use of tables to show relationships between variables	89
8.6	The data to be analysed	89
8.7	Entering the data	89
8.8	Weighting the data	90
8.9	Cross-tabulation with frequencies	91
8.10	Displaying frequencies as a percentage of the total number	92
8.11	Displaying frequencies as a percentage of the column total	93
	<i>Summary of SPSS Statistics steps for contingency tables</i>	93
9	Relationships between two or more variables: Diagrams	94
	<i>Overview</i>	94
9.1	What diagrams are used to show relationships between variables?	95
9.2	When to use diagrams to show relationships between variables	97
9.3	When not to use diagrams to show relationships between variables	98
9.4	Data requirements for diagrams to show relationships between variables	98
9.5	Problems in the use of diagrams to show relationships between variables	98
9.6	The data to be analysed	99
9.7	Entering the data	99
9.8	Weighting the data	100
9.9	Compound (stacked) percentage bar chart	101
9.10	Compound (clustered) bar chart	103
	<i>Summary of SPSS Statistics steps for bar charts</i>	105
10	Correlation coefficients: Pearson's correlation and Spearman's rho	106
	<i>Overview</i>	106
10.1	What is a correlation coefficient?	107
10.2	When to use Pearson and Spearman rho correlation coefficients	110
10.3	When not to use Pearson and Spearman rho correlation coefficients	111
10.4	Data requirements for Pearson and Spearman rho correlation coefficients	111
10.5	Problems in the use of correlation coefficients	111
10.6	The data to be analysed	112
10.7	Entering the data	112
10.8	Pearson's correlation	113
10.9	Interpreting the output	114
	<i>Reporting the output</i>	114
10.10	Spearman's rho	115
10.11	Interpreting the output	115
	<i>Reporting the output</i>	115

10.12	Scatter diagram	116
10.13	Interpreting the output	118
	<i>Reporting the output</i>	118
10.14	Scattergram with more than one case with the same two values	118
	<i>Summary of SPSS Statistics steps for correlation</i>	120
11	Regression: Prediction with precision	121
	<i>Overview</i>	121
11.1	What is simple regression?	122
11.2	When to use simple regression	124
11.3	When not to use simple regression	125
11.4	Data requirements for simple regression	125
11.5	Problems in the use of simple regression	125
11.6	The data to be analysed	126
11.7	Entering the data	126
11.8	Simple regression	127
11.9	Interpreting the output	128
11.10	Regression scatterplot	129
11.11	Interpreting the output	131
	<i>Reporting the output</i>	132
	<i>Summary of SPSS Statistics steps for simple regression</i>	132

Part 3 Significance testing and basic inferential tests 133

12	Standard error	135
	<i>Overview</i>	135
12.1	What is standard error?	136
12.2	When to use standard error	137
12.3	When not to use standard error	137
12.4	Data requirements for standard error	137
12.5	Problems in the use of standard error	138
12.6	The data to be analysed	138
12.7	Entering the data	138
12.8	Estimated standard error of the mean	139
12.9	Interpreting the output	139
	<i>Reporting the output</i>	140
	<i>Summary of SPSS Statistics steps for standard error</i>	140
13	The <i>t</i> -test: Comparing two samples of correlated/related/paired scores	141
	<i>Overview</i>	141
13.1	What is the related <i>t</i> -test?	142

13.2	When to use the related t -test	144
13.3	When not to use the related t -test	145
13.4	Data requirements for the related t -test	145
13.5	Problems in the use of the related t -test	146
13.6	The data to be analysed	146
13.7	Entering the data	147
13.8	Related t -test	147
13.9	Interpreting the output	148
	<i>Reporting the output</i>	149
	<i>Summary of SPSS Statistics steps for related t-test</i>	150
14	The t -test: Comparing two groups of unrelated/uncorrelated scores	151
	<i>Overview</i>	151
14.1	What is the unrelated t -test?	152
14.2	When to use the unrelated t -test	153
14.3	When not to use the unrelated t -test	154
14.4	Data requirements for the unrelated t -test	154
14.5	Problems in the use of the unrelated t -test	154
14.6	The data to be analysed	154
14.7	Entering the data	155
14.8	Unrelated t -test	156
14.9	Interpreting the output	156
	<i>Reporting the results</i>	158
	<i>Summary of SPSS Statistics steps for unrelated t-test</i>	158
15	Confidence intervals	159
	<i>Overview</i>	159
15.1	What are confidence intervals?	160
15.2	The relationship between significance and confidence intervals	161
15.3	Confidence intervals and limits in SPSS Statistics	162
16	Chi-square: Differences between unrelated samples of frequency data	163
	<i>Overview</i>	163
16.1	What is chi-square?	164
16.2	When to use chi-square	166
16.3	When not to use chi-square	167
16.4	Data requirements for chi-square	167
16.5	Problems in the use of chi-square	167
16.6	The data to be analysed	168
16.7	Entering the data using the 'Weighting Cases' procedure	168
16.8	Entering the data case by case	170
16.9	Chi-square	170

16.10	Interpreting the output for chi-square	172
	<i>Reporting the output</i>	173
16.11	Fisher's exact test	173
16.12	Interpreting the output for Fisher's exact test	174
	<i>Reporting the output</i>	174
16.13	One-sample chi-square	175
16.14	Interpreting the output for a one sample chi-square	176
	<i>Reporting the output</i>	177
16.15	Chi-square without ready-made tables	177
	<i>Summary of SPSS Statistics steps for chi-square</i>	177
17	McNemar's test: Differences between related samples of frequency data	178
	<i>Overview</i>	178
17.1	What is McNemar's test?	179
17.2	When to use McNemar's test	180
17.3	When not to use McNemar's test	180
17.4	Data requirements for McNemar's test	180
17.5	Problems in the use of McNemar's test	180
17.6	The data to be analysed	181
17.7	Entering the data using the 'Weighting Cases' procedure	181
17.8	Entering the data case by case	182
17.9	McNemar's test	183
17.10	Interpreting the output for McNemar's test	183
	<i>Reporting the output</i>	184
	<i>Summary of SPSS Statistics steps for McNemar's test</i>	184
18	Ranking tests for two groups: Non-parametric statistics	185
	<i>Overview</i>	185
18.1	What are non-parametric tests?	186
18.2	When to use non-parametric tests	188
18.3	When not to use non-parametric tests	188
18.4	Data requirements for non-parametric tests	188
18.5	Problems in the use of non-parametric tests	188
18.6	The data to be analysed	189
18.7	Entering the data	189
18.8	Related scores: sign test	190
18.9	Interpreting the output for the sign test	190
	<i>Reporting the output for the sign test</i>	191
18.10	Related scores: Wilcoxon test	191
18.11	Interpreting the output for the Wilcoxon test	191
	<i>Reporting the output for the Wilcoxon test</i>	192

18.12	Unrelated scores: Mann–Whitney U test	192
18.13	Entering the data	192
18.14	Mann–Whitney U -test	193
18.15	Interpreting the output for the Mann–Whitney U -test	194
	<i>Reporting the output for the Mann–Whitney U-test</i>	195
	<i>Summary of SPSS Statistics steps for non-parametric tests for two groups</i>	195
19	Ranking tests for three or more groups: Non-parametric statistics	196
	<i>Overview</i>	196
19.1	What are ranking tests?	197
19.2	When to use ranking tests	198
19.3	When not to use ranking tests	198
19.4	Data requirements for ranking tests	198
19.5	Problems in the use of ranking tests	198
19.6	The data to be analysed	199
19.7	Friedman three or more related samples test	199
19.8	Entering the data for the Friedman test	199
19.9	Friedman test	200
19.10	Interpreting the output for the Friedman test	201
	<i>Reporting the output for the Friedman test</i>	201
19.11	Kruskal–Wallis three or more unrelated samples test	202
19.12	Entering the data for the Kruskal–Wallis test	202
19.13	Kruskal–Wallis test	203
19.14	Interpreting the output for the Kruskal–Wallis test	204
	<i>Reporting the output for the Kruskal–Wallis test</i>	205
	<i>Summary of SPSS Statistics steps for non-parametric tests for three or more groups</i>	205

Part 4 Analysis of variance

207

20	The variance ratio test: Using the F -ratio to compare two variances	209
	<i>Overview</i>	209
20.1	What is the variance ratio test?	210
20.2	When to use the variance ratio test	211
20.3	When not to use the variance ratio test	211
20.4	Data requirements for the variance ratio test	211
20.5	Problems in the use of the variance ratio test	211
20.6	The data to be analysed	211
20.7	Entering the data	212
20.8	Variance estimate	213

20.9	Calculating the variance ratio from the output	214
	<i>Reporting the variance ratio</i>	214
	<i>Summary of SPSS Statistics steps for the variance ratio test</i>	214
21	Analysis of variance (ANOVA): Introduction to the one-way unrelated or uncorrelated ANOVA	215
	<i>Overview</i>	215
21.1	What is one-way ANOVA?	216
21.2	When to use one-way ANOVA	218
21.3	When not to use one-way ANOVA	218
21.4	Data requirements for one-way ANOVA	218
21.5	Problems in the use of one-way ANOVA	218
21.6	The data to be analysed	219
21.7	Entering the data	219
21.8	One-way unrelated ANOVA	219
21.9	Interpreting the output	220
	<i>Reporting the output</i>	221
	<i>Summary of SPSS Statistics steps for one-way ANOVA</i>	222
22	Analysis of variance for one-way correlated scores or repeated measures	223
	<i>Overview</i>	223
22.1	What is repeated-measures ANOVA?	224
22.2	When to use repeated-measures ANOVA	225
22.3	When not to use repeated-measures ANOVA	225
22.4	Data requirements for repeated-measures ANOVA	226
22.5	Problems in the use of one-way repeated-measures ANOVA	226
22.6	The data to be analysed	226
22.7	Entering the data	226
22.8	One-way repeated-measures ANOVA	227
22.9	Interpreting the output	228
	<i>Reporting the output</i>	231
	<i>Summary of SPSS Statistics steps for repeated-measures ANOVA</i>	231
23	Two-way analysis of variance for unrelated/uncorrelated scores	232
	<i>Overview</i>	232
23.1	What is two-way ANOVA?	233
23.2	When to use two-way ANOVA	236
23.3	When not to use two-way ANOVA	236
23.4	Data requirements for two-way ANOVA	236
23.5	Problems in the use of two-way ANOVA	237
23.6	The data to be analysed	238
23.7	Entering the data	238

23.8	Two-way unrelated ANOVA	239
23.9	Interpreting the output	240
23.10	Editing the graph	242
	<i>Reporting the output</i>	243
	<i>Summary of SPSS Statistics steps for two-way ANOVA</i>	244
24	Multiple comparison in ANOVA	245
	<i>Overview</i>	245
24.1	What is multiple comparison testing?	246
24.2	When to use multiple comparison tests	247
24.3	When not to use multiple comparison tests	247
24.4	Data requirements for multiple comparison tests	248
24.5	Problems in the use of multiple comparison tests	248
24.6	The data to be analysed	249
24.7	Entering the data	249
24.8	Multiple comparison tests	249
24.9	Interpreting the output	250
	<i>Reporting the output</i>	252
	<i>Summary of SPSS Statistics steps for multiple comparison tests</i>	252
25	Analysis of variance for two-way correlated scores or repeated measures	253
	<i>Overview</i>	253
25.1	What is two-way repeated-measures ANOVA?	254
25.2	When to use two-way repeated-measures ANOVA	256
25.3	When not to use two-way repeated-measures ANOVA	257
25.4	Data requirements for two-way repeated-measures ANOVA	257
25.5	Problems in the use of two-way repeated-measures ANOVA	257
25.6	The data to be analysed	257
25.7	Entering the data	258
25.8	Two-way repeated-measures ANOVA	258
25.9	Interpreting the output	260
	<i>Reporting the output</i>	264
	<i>Summary of SPSS Statistics steps for two-way repeated-measures ANOVA</i>	265
26	Two-way mixed analysis of variance (ANOVA)	266
	<i>Overview</i>	266
26.1	What is two-way mixed ANOVA?	267
26.2	When to use two-way mixed ANOVA	268
26.3	When not to use two-way mixed ANOVA	268
26.4	Data requirements for two-way mixed ANOVA	269
26.5	Problems in the use of two-way mixed ANOVA	269
26.6	The data to be analysed	269

26.7	Entering the data	269
26.8	Two-way mixed ANOVA	270
26.9	Interpreting the output	272
	<i>Reporting the output</i>	274
	<i>Summary of SPSS Statistics steps for mixed ANOVA</i>	275
27	Analysis of covariance (ANCOVA)	276
	<i>Overview</i>	276
27.1	What is analysis of covariance (ANCOVA)?	277
27.2	When to use ANCOVA	278
27.3	When not to use ANCOVA	278
27.4	Data requirements for ANCOVA	279
27.5	Problems in the use of ANCOVA	279
27.6	The data to be analysed	279
27.7	Entering the data	279
27.8	One-way ANCOVA	280
27.9	Testing that the slope of the regression line within cells is similar	281
27.10	Interpreting the output	282
27.11	Testing the full model	282
27.12	Interpreting the output	284
	<i>Reporting the output</i>	285
	<i>Summary of SPSS Statistics steps for ANCOVA</i>	286
28	Multivariate analysis of variance (MANOVA)	287
	<i>Overview</i>	287
28.1	What is multivariate analysis of variance (MANOVA)?	288
28.2	When to use MANOVA	289
28.3	When not to use MANOVA	290
28.4	Data requirements for MANOVA	291
28.5	Problems in the use of MANOVA	291
28.6	The data to be analysed	291
28.7	Entering the data	292
28.8	MANOVA	292
28.9	Interpreting the output	293
	<i>Reporting the output</i>	296
	<i>Summary of SPSS Statistics steps for MANOVA</i>	296
29	Discriminant (function) analysis (for MANOVA)	297
	<i>Overview</i>	297
29.1	What is discriminant function analysis?	298
29.2	When to use discriminant function analysis	300

29.3	When not to use discriminant function analysis	300
29.4	Data requirements for discriminant function analysis	300
29.5	Problems in the use of discriminant function analysis	300
29.6	The data to be analysed	301
29.7	Entering the data	301
29.8	Discriminant function analysis	302
29.9	Interpreting the output	304
	<i>Reporting the output</i>	307
	<i>Summary of SPSS Statistics steps for discriminant function analysis</i>	308

Part 5 More advanced correlational statistics 309

30	Partial correlation	311
	<i>Overview</i>	311
30.1	What is partial correlation?	312
30.2	When to use partial correlation	314
30.3	When not to use partial correlation	314
30.4	Data requirements for partial correlation	314
30.5	Problems in the use of partial correlation	314
30.6	The data to be analysed	314
30.7	Entering the data	315
30.8	Partial correlation	315
30.9	Interpreting the output	316
	<i>Reporting the output</i>	317
	<i>Summary of SPSS Statistics steps for partial correlation</i>	317
31	Factor analysis	318
	<i>Overview</i>	318
31.1	What is factor analysis?	319
31.2	When to use factor analysis	322
31.3	When not to use factor analysis	322
31.4	Data requirements for factor analysis	322
31.5	Problems in the use of factor analysis	322
31.6	The data to be analysed	323
31.7	Entering the data	323
31.8	Principal components analysis with orthogonal rotation	324
31.9	Interpreting the output	326
	<i>Reporting the output</i>	329
	<i>Summary of SPSS Statistics steps for factor analysis</i>	329

32	Item reliability and inter-rater agreement	330
	<i>Overview</i>	330
32.1	What are item reliability and inter-rater agreement?	331
32.2	When to use item reliability and inter-rater agreement	334
32.3	When not to use item reliability and inter-rater agreement	334
32.4	Data requirements for item reliability and inter-rater agreement	334
32.5	Problems in the use of item reliability and inter-rater agreement	334
32.6	The data to be analysed for item alpha reliability	335
32.7	Entering the data	335
32.8	Alpha reliability	336
32.9	Interpreting the output	337
	<i>Reporting the output</i>	337
32.10	Split-half reliability	337
32.11	Interpreting the output	338
	<i>Reporting the output</i>	338
32.12	The data to be analysed for inter-rater agreement (kappa)	338
32.13	Entering the data	338
32.14	Kappa	339
32.15	Interpreting the output	340
	<i>Reporting the output</i>	341
	<i>Summary of SPSS Statistics steps for reliability</i>	341
33	Stepwise multiple regression	342
	<i>Overview</i>	342
33.1	What is stepwise multiple regression?	343
33.2	When to use stepwise multiple regression	345
33.3	When not to use stepwise multiple regression	346
33.4	Data requirements for stepwise multiple regression	346
33.5	Problems in the use of stepwise multiple regression	346
33.6	The data to be analysed	347
33.7	Entering the data	347
33.8	Stepwise multiple regression analysis	348
33.9	Interpreting the output	349
	<i>Reporting the output</i>	351
	<i>Summary of SPSS Statistics steps for stepwise multiple regression</i>	352
34	Simultaneous or standard multiple regression	353
	<i>Overview</i>	353
34.1	What is simultaneous or standard multiple regression?	354
34.2	When to use simultaneous or standard multiple regression	358
34.3	When not to use simultaneous or standard multiple regression	358

34.4	Data requirements for simultaneous or standard multiple regression	359
34.5	Problems in the use of simultaneous or standard multiple regression	359
34.6	The data to be analysed	360
34.7	Entering the data	360
34.8	Simultaneous or standard multiple regression analysis	361
34.9	Interpreting the output	362
	<i>Reporting the output</i>	364
	<i>Summary of SPSS Statistics steps for simultaneous or standard multiple regression</i>	365
35	Simple mediational analysis	366
	<i>Overview</i>	366
35.1	What is simple mediational analysis?	367
35.2	When to use simple mediational analysis	370
35.3	When not to use simple mediational analysis	371
35.4	Data requirements for a simple mediational analysis	371
35.5	Problems in the use of simple mediational analysis	371
35.6	The data to be analysed	372
35.7	Entering the data	372
35.8	Simultaneous or standard multiple regression analysis	373
35.9	Interpreting the output	374
	<i>Reporting the output</i>	375
	<i>Summary of SPSS Statistics steps for simultaneous or standard multiple regression</i>	376
36	Hierarchical multiple regression	377
	<i>Overview</i>	377
36.1	What is hierarchical multiple regression?	378
36.2	When to use hierarchical multiple regression	380
36.3	When not to use hierarchical multiple regression	380
36.4	Data requirements for hierarchical multiple regression	380
36.5	Problems in the use of hierarchical multiple regression	380
36.6	The data to be analysed	381
36.7	Entering the data	381
36.8	Hierarchical multiple-regression analysis	382
36.9	Interpreting the output	383
	<i>Reporting the output</i>	385
	<i>Summary of SPSS Statistics steps for hierarchical multiple regression</i>	386
37	Moderator analysis with continuous predictor variables	387
	<i>Overview</i>	387
37.1	What is a moderator variable analysis?	388
37.2	When to use moderator variable analysis	390
37.3	When not to use moderator variable analysis	390

37.4	Data requirements for moderator variable analysis	390
37.5	Problems in the use of moderator variable analysis	391
37.6	The data to be analysed	391
37.7	Entering the data	392
37.8	Standardising the variables	393
37.9	Computing the interaction term	393
37.10	Hierarchical multiple-regression analysis	394
37.11	Interpreting the output	395
37.12	Entering the data for predicting the criterion values	396
37.13	Computing the predicted criterion values	397
37.14	Plotting the predicted criterion values	398
	<i>Reporting the output</i>	399
	<i>Summary of SPSS Statistics steps for moderator analysis with hierarchical multiple regression</i>	399

Part 6 Advanced qualitative or nominal techniques 401

38	Log-linear analysis	403
	<i>Overview</i>	403
38.1	What is log-linear analysis?	404
38.2	When to use log-linear analysis	406
38.3	When not to use log-linear analysis	406
38.4	Data requirements for log-linear analysis	407
38.5	Problems in the use of log-linear analysis	407
38.6	The data to be analysed	407
38.7	Entering the data	407
38.8	Log-linear analysis	408
38.9	Interpreting the output	409
	<i>Reporting the output</i>	411
	<i>Summary of SPSS Statistics steps for log-linear analysis</i>	412
39	Multinomial logistic regression	413
	<i>Overview</i>	413
39.1	What is multinomial logistic regression?	414
39.2	When to use multinomial logistic regression	417
39.3	When not to use multinomial logistic regression	418
39.4	Data requirements for multinomial logistic regression	418
39.5	Problems in the use of multinomial logistic regression	418
39.6	The data to be analysed	418

39.7	Entering the data	419
39.8	Stepwise multinomial logistic regression	420
39.9	Interpreting the output	421
	<i>Reporting the output</i>	428
	<i>Summary of SPSS Statistics steps for multinomial logistic regression</i>	428
40	Binomial logistic regression	429
	<i>Overview</i>	429
40.1	What is binomial logistic regression?	430
40.2	When to use binomial logistic regression	433
40.3	When not to use binomial logistic regression	433
40.4	Data requirements for binomial logistic regression	434
40.5	Problems in the use of binomial logistic regression	434
40.6	The data to be analysed	434
40.7	Entering the data	434
40.8	Binomial logistic regression	435
40.9	Interpreting the output	436
	<i>Reporting the output</i>	440
	<i>Summary of SPSS Statistics steps for binomial logistic regression</i>	440

Part 7 Data handling procedures 441

41	Reading ASCII or text files into the Data Editor	443
	<i>Overview</i>	443
41.1	What is an ASCII or text data file?	444
41.2	Entering data into an ASCII or text data file	445
41.3	Reading an ASCII or text data file	446
	<i>Summary of SPSS Statistics steps for inputting an ASCII or text data file</i>	449
42	Missing values	450
	<i>Overview</i>	450
42.1	What are missing values?	451
42.2	Entering the data	452
42.3	Defining missing values	453
42.4	Pairwise and listwise options	454
42.5	Sample output for pairwise deletion	455
42.6	Sample output for listwise deletion	456
42.7	Interpreting the output	456
	<i>Reporting the output</i>	456
	<i>Summary of SPSS Statistics steps for handling missing values</i>	457

43	Recoding values	458
	<i>Overview</i>	458
43.1	What is recoding values?	459
43.2	Entering the data	460
43.3	Recoding values	461
43.4	Recoding missing values	463
43.5	Saving the Recode procedure as a syntax file	463
43.6	Adding some extra cases to Table 43.1	464
43.7	Running the Recode syntax command	464
	<i>Summary of SPSS Statistics steps for recoding values</i>	465
44	Computing a scale score with no missing values	466
	<i>Overview</i>	466
44.1	What is computing a scale score?	467
44.2	Entering the data	468
44.3	Displaying variable labels in dialog boxes	469
44.4	Computing a scale score	471
44.5	Saving the Compute procedure as a syntax file	473
44.6	Adding some extra cases to Table 44.1	473
44.7	Running the Compute syntax command	473
	<i>Summary of SPSS Statistics steps for computing a scale score with no missing values</i>	474
45	Computing a scale score with some values missing	475
	<i>Overview</i>	475
45.1	What is computing a scale score with some values missing?	476
45.2	Entering the data	477
45.3	Computing a scale score with some values missing	478
45.4	Saving the Compute procedure as a syntax file	480
45.5	Adding some extra cases to Table 45.1	480
45.6	Running the Compute syntax command	481
	<i>Summary of SPSS Statistics steps for computing a scale score with some missing values</i>	481
46	Computing a new group variable from existing group variables	482
	<i>Overview</i>	482
46.1	What is computing a new group variable from existing group variables?	483
46.2	Entering the data	485
46.3	Syntax file for computing a new group variable from existing group variables	485
46.4	Running the Compute syntax commands	486
46.5	Computing a new group using menus and dialog boxes	487
	<i>Summary of SPSS Statistics steps for computing a new group variable from existing group variables</i>	488

47	Selecting cases	489
	<i>Overview</i>	489
47.1	What is selecting cases?	490
47.2	Entering the data	491
47.3	Selecting cases	491
	<i>Summary of SPSS Statistics steps for selecting cases</i>	494
48	Samples and populations: Generating a random sample	495
	<i>Overview</i>	495
48.1	What is generating random samples?	496
48.2	Entering the data	496
48.3	Selecting a random sample	497
48.4	Interpreting the results	498
48.5	Statistical analysis on a random sample	499
	<i>Summary of SPSS Statistics steps for generating a random sample</i>	499
49	Inputting a correlation matrix	500
	<i>Overview</i>	500
49.1	What is inputting a correlation matrix?	501
49.2	Syntax file for inputting a correlation matrix and running a stepwise multiple regression	503
49.3	Running the syntax file	503
49.4	Part of the output	503
	<i>Summary of SPSS Statistics steps for inputting a correlation matrix</i>	504
50	Checking the accuracy of data inputting	505
	<i>Overview</i>	505
50.1	What is checking the accuracy of data inputting?	506
50.2	Creating two data files	507
50.3	Combining the two data files	510
50.4	Creating a syntax file for computing the difference between the two entries for the same variables	511
	<i>Summary of SPSS Statistics steps for checking the accuracy of data inputting</i>	514

Part 8 Other statistical procedures 515

51	Statistical power analysis: Sample size estimation	517
	<i>Overview</i>	517
51.1	What is statistical power analysis?	518
51.2	When to use statistical power analysis	521
51.3	When not to use statistical power analysis	522

51.4	Data requirements for statistical power analysis	522
51.5	Problems in the use of statistical power analysis	522
51.6	The data to be analysed	522
51.7	Power analysis	523
51.8	Interpreting the output	525
	<i>Reporting the output</i>	526
52	Meta-analysis	527
	<i>Overview</i>	527
52.1	What is meta-analysis?	528
52.2	When to use meta-analysis	532
52.3	When not to use meta-analysis	533
52.4	Data requirements for meta-analysis	533
52.5	Problems in the use of meta-analysis	533
52.6	The data to be analysed	534
52.7	Meta-analysis	534
52.8	Interpreting the output	535
	<i>Reporting the output</i>	536
	<i>Appendix: Some other statistics in SPSS Statistics</i>	537
	<i>Glossary</i>	539
	<i>Index</i>	547

Companion Website

For open-access **student resources** specifically written to complement this textbook and support your learning, please visit **www.pearsoned.co.uk/howitt**



ON THE WEBSITE

The website contains:

- Datasets relating to exercises from the book
- Additional datasets with questions
- A set of research scenarios with multiple choice questions
- Flashcards

Guided tour

8.1 WHAT TABLES ARE USED TO SHOW RELATIONSHIPS BETWEEN VARIABLES 87

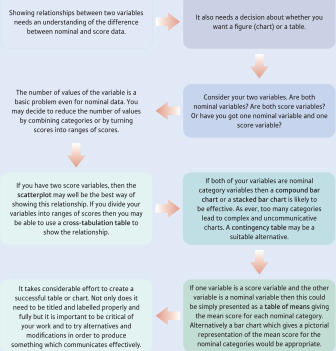


FIGURE 8.1 Steps in showing relationships between two variables

		Knowledge					Total
		1.00	2.00	3.00	5.00	7.00	
Aptitude	1.00	6	0	0	0	0	6
	2.00	0	3	0	0	0	3
	3.00	0	0	3	0	0	3
	4.00	0	0	6	3	0	9
	5.00	0	3	0	0	0	3
	6.00	0	0	3	0	0	3
Total		6	6	12	3	3	30

FIGURE 8.2 A cross-tabulation table

Background details

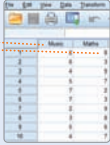
Outlines the background to the techniques discussed in each chapter to encourage a deeper understanding. This includes details of:

- What the technique is
- When you should use it
- When you should not use it
- The data required for the analysis
- Typical problems to be aware of

10.8 PEARSON'S CORRELATION 113

Step 2

In 'Data View' of the 'Data Editor' enter 'Music' scores in the 1st column and 'Maths' scores in the 2nd column. Save this data as a file to use for Chapter 11.

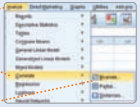


	Music	Maths
1	10	10
2	10	10
3	10	10
4	10	10
5	10	10
6	10	10
7	10	10
8	10	10
9	10	10
10	10	10


10.8 Pearson's correlation

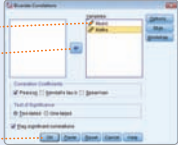
Step 1

Select 'Analyze', 'Correlate' and 'Bivariate'.



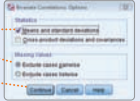
Step 2

Select 'Music' and 'Maths' either singly (or together by holding down the SHIFT, Control and downward keys). Select the  button to put them in the 'Variables:' box as they appear here. If you have more than 2 variables you want to correlate, you can put them all here (see Chapter 49). Select 'Options' if you want means and standard deviations or if you want to select the 'Exclude cases pairwise' option (see Chapter 49) for missing data on more than 2 variables. Otherwise select 'OK'.



Step 3

Select 'Means and standard deviations' to obtain these. If you have some missing data, you may wish to select 'Exclude cases pairwise' which looks at pairs of variables which do not have missing data. Select 'Continue' and then in the previous box 'OK'.



Step by step illustrations and screenshots of SPSS

This presents the stages of data entry and data analysis visually to help you gain confidence in the processes and procedures of SPSS.

10.9 Interpreting the output

There is only one table when only correlations are requested, as in Step 2.

The correlation between 'Maths' and 'Music' is -0.90 . The 2-tailed significance or probability level is .001 or less so the correlation is statistically significant. The number of cases on which this correlation is based is 10. This information is also given in this cell.

Correlations		Music	Maths
Music - Pearson Correlation	1.	1.	.000
		Sig. (2-tailed)	.19
Maths - Pearson Correlation	-.900 ^a		1.
		Sig. (2-tailed)	.001
		N	10

^aCorrelation is significant at the 0.05 level (2-tailed).

When means and standard deviations are also requested, as in Step 3, two tables are presented. The second one is the same as the one above, while the first one contains the means and standard deviations.

This table shows the mean, standard deviation and number of cases (N) for the variables analysed.

Descriptive Statistics			
	Mean	Std. Deviation	N
Music	4.50	1.447	10
Maths	6.20	2.015	10

REPORTING THE OUTPUT

- The correlation between musical ability and mathematical ability is -0.90 . It is usual to round correlations to two decimal places, which would make it -0.90 . This is more than precise enough for most psychological measurements. Note that there is no need to put a 0 before the decimal point (e.g. -0.90) because a correlation cannot be bigger than ± 1.00 .
- The exact significance level to three decimal places is .001. This means that the significance level is less than 0.001. We would suggest that you do not use a string of zeros, as these confuse people. Always change the third zero to a 1. This means that the significance level can be reported as being $p < 0.001$.
- It is customary to present the degrees of freedom (df) rather than the number of cases when presenting correlations. The degrees of freedom are the number of cases minus 2, which makes them 8 for this correlation. There is nothing wrong with reporting the number of cases instead.
- In a report, we would write 'There is a significant negative relationship between musical ability and mathematical ability, $r(8) = -0.90$, $p < 0.001$. Children with more musical ability have lower mathematical ability.'



The significance of the correlation coefficient is discussed in more detail in Chapter 11 of Howitt, D. and Cramer, D. (2014) *Introduction to Statistics in Psychology*, Harlow: Pearson.

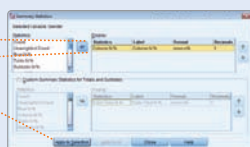
Interpreting the output

Offers a simple explanation of what the important parts of the output mean. SPSS statistical output is presented exactly as it appears on screen to help you become familiar with it.

8.11 DISPLAYING FREQUENCIES AS A PERCENTAGE OF THE COLUMN TOTAL 93

8.11 Displaying frequencies as a percentage of the column total

Select 'Table MN' and the  button to put it back under 'Statistics.' Select 'Column N%' and the  button to put it under 'Display.' Select 'Apply to Selection' to return to original dialog box. Select 'OK'.



If you add the % in each column they total 100.

	Gender	
	Females	Males
Highly satisfied	64.1%	40.0%
Not highly satisfied	35.9%	60.0%

Summary of SPSS Statistics steps for contingency tables**Data**

- In Variable View of the Data Editor, 'name' the 'nominal' variables and select 'nominal' as the measure.
- In Data View of the Data Editor, enter the data under the appropriate variable names.

Analysis

- For a contingency table, select 'Analyze', 'Table' and 'Custom Tables...'.
- Select 'OK', move the appropriate variables to the column and row boxes and select 'OK'.

Output

- Check tables are correct. They can be edited in the Chart Editor if required.

For further resources including data sets and questions, please refer to the website accompanying this book.

Summary of SPSS Statistics steps

Summarises the important steps so that you can easily see how to input, analyse and present data.

Introduction

Statistical Package for the Social Sciences (SPSS) was initially developed in 1965 at Stanford University in California. Since then it has become the leading data analysis package in the field and available all over the world in universities and elsewhere. Modern computing developments have allowed it to be used on home computers. Because of its popularity and universality, using SPSS is one of the most readily transferable of all research skills. Once learnt, SPSS can be used virtually anywhere in the world. SPSS is constantly being updated, both in terms of the range of statistical techniques covered and the detail of the output.

This book is a stand-alone, step-by-step approach to statistical analysis using SPSS for Windows and is applicable to Releases 10 to 21. It is suitable for students and researchers wishing to analyse psychological, sociological, criminological, health and similar data. SPSS does change with time, but this book will also be helpful to those using earlier releases of SPSS as the changes which affect this book are generally hard to notice. Although the last six releases of SPSS have Statistics in the title and releases 17 and 18 were called **PASW Statistics**, we shall generally refer to all versions as SPSS unless we are speaking about particular versions – in which case we will give their release numbers. This is what is generally done by most users. The official name of the latest release at the time of publication is IBM SPSS Statistics Version 22 (Release 22.0.0).

This book updates the fifth edition of *Introduction to SPSS Statistics in Psychology* to cover recent changes in SPSS. The structure provides the fastest possible access to computerised data analysis with SPSS, even when using the most advanced techniques. Each statistical technique is carefully described, step-by-step, using screenshots of SPSS data analysis and output. The user's attention is focused directly on the screenshots, what each of them signifies, and why they are important. In other words, it is as close as is possible in a textbook to face-to-face individual instruction. Users with little or no previous computer skills will be able to quickly analyse quite complex data and appreciate what the output means.

The chapters have a common pattern. The computer steps (which keys to press) are given in exact sequence. However, this is not the end of any data analysis, and so there are also explanations of how to interpret and report the SPSS output. For this new edition, we have added one new SPSS chapter on two-way repeated measures. SPSS is excellent but it does not include certain procedures which a good researcher may need. So we have included instructions in how to carry out a few additional things. The section on power analysis and meta-analysis which uses software freely available on the internet is now included in the book as the eighth section.

The basic structure of the major chapters is:

- A brief bulleted overview of the use of the procedure or statistical technique. This will often be sufficient to get a clear idea of where and when to use the techniques.
- An account of what the technique is for and what needs to be known in preparation for doing the SPSS analysis. It also gives information about when the technique is

used, when it should not be used, the data required for the analysis, and typical problems that we know from long experience cause users difficulties.

- An illustrative example is given of the appropriate sorts of data for each statistical technique. These examples allow the user to work through our computations, and to gain confidence before moving on to their own data.
- Data entry for a particular statistical analysis is presented visually and explained in adjacent text.
- This is followed by a step-by-step, screenshot-by-screenshot, description of how a particular statistical analysis is done using SPSS for Windows.
- The SPSS statistical output is included exactly as it appears on the monitor screen and in printouts of the analysis. This is crucial – SPSS output can be confusing and unclear at first.
- The key features of the statistical output are highlighted on the output itself, together with simple explanations of what the important parts of the output mean – SPSS output is infamous for its over-inclusiveness.
- Suggestions are made on reporting the statistical findings in reports, theses and publications. These include samples of how to describe research findings and present tables clearly. The form followed is that recommended by the American Psychological Association (APA), which is also widely used by other publishers.

This book is based on the latest version of *SPSS Statistics for Windows* (that is, Release 22); but remains suitable for Releases 10 to 21 because of their similarity. Notes after this Introduction describe the main differences between these releases. Although SPSS is updated every year or so, usually there should be little difficulty in adapting knowledge gained on the older versions to the new version.

Introduction to SPSS in Psychology is an excellent single source for data analysis. It is complete in itself and contains many features not available elsewhere. Unlike other SPSS books, it meets the needs of students and researchers at all levels. However, it is also part of a package of methodology books by the same authors designed to be comprehensive, authoritative and exhaustive. The three volumes in this series are closely tied to each other. The other two are:

- *Introduction to Statistics in Psychology* (2014) (6th edition) (Pearson Education: Harlow): This is a thorough introduction to statistics for all students. It consists of a basic introduction to key psychological statistics and also covers many intermediate and advanced techniques in detail, while maintaining its accessibility to students. It contains chapters on topics, such as meta-analysis, which are seldom covered in other statistics texts. Importantly, the structure of the statistics textbook is closely linked to this book. Thus, anyone following a chapter in the statistics book will, where appropriate, find an equivalent chapter in this book with details of how to do the analysis using SPSS. Similarly, anyone using this book will be able to find a detailed account of the technique in the statistics textbook.
- *Introduction to Research Methods in Psychology* (2014) (4th edition) (Pearson Education: Harlow): This is a major textbook on research methods in psychology. It covers both quantitative and qualitative methods. There are major chapters on report writing, ethics in psychology and searching the literature. All aspects of experimental, field study, survey and questionnaire construction are covered, and guidance is given on qualitative data collection and analysis. There are numerous cross-references to this book and *Introduction to Statistics in Psychology*.

In other words, the three books offer a comprehensive introduction to conducting research in psychology. They may be used independently or in any combination.

Introduction to SPSS in Psychology can be used alongside virtually any statistics textbook to support a wide variety of statistics and practical courses. The range of statistical techniques covered is large and includes the simplest as well as the most important advanced statistical techniques. The variety of techniques described and the relative ease of using SPSS Statistics for Windows ensure that this guide can be used at introductory, intermediate and advanced levels of statistics teaching. The structure of the book is such that statistical procedures are described more or less in order of conceptual difficulty. Generally speaking, computing with SPSS is as easy for advanced statistical techniques as it is for simple ones.

Chapter 2 is essential reading, as it explains data entry and basic computer operating. However, the subsequent chapters can be used on a stand-alone basis if desired. Users with insufficient time to work through the guide chapter by chapter should find enough detail in the relevant chapters to complete an SPSS analysis successfully. Table 1.1, at the end of Chapter 1, states which chapter is most appropriate for which purpose, thereby enabling the reader to move directly to that part of the book.

Those who work steadily through the book will profit by doing so. They will have a much better overview of SPSS computing procedures. For most readers, this is possible in a matter of hours, especially if they have prior knowledge of statistics.

SPSS has an extensive catalogue of statistical procedures – far more than could be included. We have selected those suitable for most purposes when the range of possibilities is likely to confuse the average reader. The quickness and ease of SPSS mean that more advanced users can explore the alternatives by using the menus and dialog boxes. Most users will find our coverage more than sufficient.

The data and statistical analyses carried out in this book correspond almost always to those in the authors' accompanying statistics text, *Introduction to Statistics in Psychology* (2014) (6th edition) (Pearson Education: Harlow). This book is referred to as *ISP*, followed by the corresponding chapter or table number.

Dennis Howitt
Duncan Cramer

Acknowledgements

Our names may appear on the cover of this book but it is the talent and hard work of other people that makes it what it is. The team that Pearson provides to do all of the clever, difficult stuff makes all of the difference between a dull manuscript full of errors and this book which, we believe, is a credit to us but undeservedly so. These back-room stars deserve most of the praise and this is our opportunity to thank them. The team for this edition is international and much of the scarily efficient coordination of everything was in the hands of Priyadharshini Dhanagopal. She was the Managing Project Editor working in Chennai. Regrettably, the Internet ensured that we had no place to run and no place to hide. Without her doing all manner of things in the background which we cannot hope to understand, our lives would have been more difficult and more disorganized than ever. We can only wonder what her job specification is and how one person can fulfil it. But she can and we are more than grateful for that. Angela Hawksbee was the Production Controller. She liaised with type-setters, printers and the like and made sure that everything happened when it should happen. Fortunately, this did not involve us directly as she is formidable in her work judging by the smooth way with which everything progressed.

Janey Webb was the Acquisitions Editor though now she has the title of Publisher. We are not sure what this is but roughly translated it seems to mean 'tower of strength'. We have been working with Janey for quite a few editions now and it never fails to amaze us how she always has the energy and determination to get the best new edition out of us that we are capable of. This is all done in ways far too subtle for two psychologists to understand. Somehow, we found ourselves wanting to do what we did not want to do and feeling happy about it. How she gets inside of the topic in question even for arcane things like SPSS and statistics will never cease to amaze us.

If you like the way that this book looks (and we think it looks great) there were two main people to thank. The Cover Designer was Nicola Woowat. How she comes up with fresh ideas which make SPSS look so exciting we do not know but we are glad that she can. Kevin Ancient designed the text. His work is so important to a book like this one. It gives a coherent structure to each chapter, which the manuscript alone cannot. It is far easier to read the book than the manuscript because of this. Kevin's master plan was executed by the Copy-editor Anne Henwood. She has the manuscript in one hand, the design in another hand, and brings them both together in another hand. This, of course, makes her super human too. She is enormously patient and polite even when we are probably being very annoying. The trouble is that she is always right. While doing all of this, she still managed to spot all sorts of errors. When the copy has been turned into the page proofs then the Proof Reader goes to work getting rid of all of those awkward glitches, typos, and mistakes which can make a textbook hard to read and very irritating. Brian Burge is our new best friend for doing this so well and thoroughly. He could spot things that we could not even see even after he told us exactly where they were! Annette Musker was the Indexer. She makes it so much easier to find one's way

around the book. It is a job that anyone can make a mess of but needs a special talent to do well. Annette gets the difficult balance needed for an index exactly right. Of course, there are lots of other people that we ought to thank but can only write how much their contributions are appreciated by us.

Thanks

Dennis Howitt and Duncan Cramer

■ Publisher acknowledgements

We are grateful to the following for permission to reproduce copyright material:

The screenshots in this book are reprinted by permission of IBM SPSS Inc. Screenshots in this book are copyright IBM SPSS Inc. IBM SPSS is a registered trademark and the other product names are registered trademarks of IBM SPSS Inc.

Key differences between IBM SPSS Statistics 22 and earlier versions

SPSS Statistics 21

There seem to be very few differences between 22 and 21 and between 21 and 20 for the procedures described in this book.

SPSS Statistics 20

There seem to be very few differences between 20 and 19 for the procedures described in this book.

SPSS Statistics 19

The name of the software became SPSS Statistics. There seem to be very few differences between 19 and 18 for the procedures described in this book.

PASW Statistics 18

There seem to be very few differences between 18 and 17 for the procedures described in this book.

PASW Statistics 17

This version of SPSS was called PASW, which stands for Predictive Analytic Software. Otherwise, there appear to be very few differences between 17 and 16 for the procedures described in this book. Data Reduction which includes Factor Analysis is now called Dimension Reduction. Right clicking on the keys in the 'Select If: If' box no longer gives a description of what the keys do.

SPSS 16

There seem to be very few differences between 16 and 15 for the procedures described in this book. Basic tables are no longer available in the Tables procedure, so the tables in Chapter 8 are produced with Custom Tables. In the dialog boxes the OK, Paste,

Reset, Cancel and Help options are on the bottom of the box rather than on the right-hand side while the analyses options are on the right-hand side rather than at the bottom of the box.

SPSS 15

There also appear to be very few differences between 15 and 14. The Graph menu in 14 displays all available options. The procedure for Chart Builder is described in Chapter 4 of this book as it was not available in releases before 14. A slightly alternative procedure is shown in Chapters 6, 9 and 11.

SPSS 14

Similarly, there seem to be very few differences between 14 and 13. Release 13 does not have Chart Builder, which was introduced in 14. In 14 the Properties dialog box of the Chart Editor has separate boxes for Text Style and Text Layout.

SPSS 13

The major differences between 13 and 12 are to ‘Compute Variable . . .’, ‘Scatter/Dot . . .’ and the ‘Chart Editor’. Also the plots in the output of 13 are shaded.

In 12 the ‘Compute Variable’ dialog box has a single ‘Functions’ menu from which options can be chosen. ‘Scatter/Dot . . .’ is called ‘Scatter . . .’, the ‘Scatter/Dot’ dialog box is called ‘Scatterplot’ and there is no ‘Dot’ option.

In 12, to label the slices of a pie diagram and add the percentages of cases in each, double click anywhere in the ‘Chart Editor’, double click on the pie diagram (to open the ‘Properties’ dialog box), select ‘Data Value Labels’ (in the ‘Properties’ dialog box), select ‘Count’ in the ‘Contents’ box, select the red ‘X’ (to put ‘Count’ in the ‘Available’ box), select the variable name (e.g. ‘Occupation’), select the curved upward arrow (to put ‘Occupation’ in the ‘Contents’ box), select ‘Percent’ and the curved upward arrow (to put ‘Percent’ in the ‘Contents’ box), select ‘Apply’ and then ‘Close’.

To fit a regression line to a scatterplot, click on a dot in the chart of the ‘Chart Editor’ so that the circles in the plot become highlighted, select ‘Chart’, select ‘Add Chart Element’, select ‘Fit Line at Total’ (which opens the ‘Properties’ dialog box). Assuming that the ‘Fit Line’ tab is active, select ‘Linear’ (this is usually the default) and then ‘Close’.

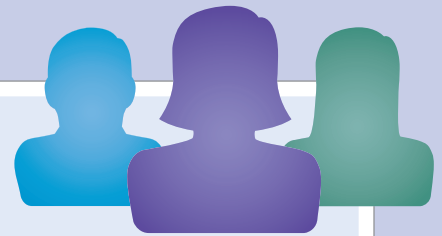
SPSS 12

The major differences between 12 and 11 also apply to 10. They are relatively few. In 11 and 10 variable names cannot begin with a capital letter and are restricted to eight characters. The ‘Data’ and ‘Transform’ options are not available in the ‘Viewer’ or ‘Output’ window. Some output, such as partial correlation and reliability, is not organised into tables. The ‘Chart Editor’ works differently. To fit a regression line to a scatterplot, double click anywhere in the scatterplot to open the ‘Chart Editor’, select ‘Chart’, select ‘Options . . .’ (which opens the ‘Scatterplot Options’ dialog box), select ‘Total’ under ‘Fit Line’ and then ‘OK’.

PART 1

Introduction to SPSS Statistics





CHAPTER 1

A brief introduction to statistics

Overview

- Unfortunately, there is no gain without some effort in statistics. There are a small number of statistical concepts which need to be understood in order to speed up learning how to use statistics in research. Each of these is discussed and explained in this chapter.
- Key ideas covered in this chapter include score variables versus nominal (category) variables; unrelated versus **related designs**; descriptive versus inferential statistics; and significance testing. With a knowledge of each of these it is possible to quickly develop a working knowledge of statistical analysis using SPSS Statistics.
- The appropriate statistical analysis for data depends very much on the particular type of research design employed. You need a basic understanding of what statistics are appropriate for your research design.
- The chapter provides detailed advice on how to select a statistical technique for the analysis of psychological data.

1.1 Basic statistical concepts essential in SPSS Statistics analyses

The basics of statistics are quite simple. The problem is in putting these elements together. Nobody can become expert in statistical analysis overnight but a very small amount of knowledge can lead to quite sophisticated analyses being carried out by even the most inexperienced of researchers. Mathematical ability has very little role to play in data analysis. Much more important is that the researcher understands some basic principles of research design. There are close links between different research designs and what the appropriate statistical analysis methods for one's data are. At the most basic level, there are two broad classes of research design – the comparative and the correlational designs. Of course, each of these has any number of variants. The type of research design involved in the study lays down broadly the sort of statistical tests, etc. which are needed for the analysis of the data from that study. Sometimes the personal preferences of the researcher play a part since, quite often, there are several ways of achieving much the same ends.

But before we can discuss research designs, there are two basic concepts we need to understand as they are part of the jargon of statistics and SPSS Statistics:

Variable: A variable is any concept that can be measured and which varies. Variables are largely inventions of the researcher and they can be very different from study to study. There are a few fairly standard variables, such as age and gender that are very commonly measured. Typically, however, the variables used tend to be specific to particular topics of study. Variables are the means by which psychologists attempt to measure the concepts that they use in their research – a variable, generally, cannot perfectly measure a concept and so is an approximation to the concept. For this reason, it is important to understand that data which involves variables and theory which involves concepts do not always map perfectly with each other.

Cases: A case is simply a member of the sample. In psychology a case is usually a person (i.e. an individual participant in the research). Cases are very much SPSS Statistics jargon. It is a wider and more embracing term than the participants which psychologists talk about.

Variables normally appear in SPSS analyses as the columns of the data spreadsheet. Cases (normally) appear in SPSS analyses as the rows of the data spreadsheet though it is possible to reverse these. In other words, variables and cases can be set out in the form of a **matrix** or a two-dimensional array. The size of the matrix depends on the number of variables and cases involved.

1.2 Basic research designs: comparative versus correlational designs

■ Comparative designs

The basic comparative design compares the typical or average score of a group of participants with that of another group. This might involve comparing a group of men with a group of women or comparing an experimental group with a control group in an experimental study. This design is illustrated in Table 1.1. Usually, in this sort of design, the comparison is between the average score for one group and the average score in the other group. Usually, what most people refer to as the *average* is called by

Table 1.1 Simple comparative design such as an experiment			
Participant (case)	GROUP A (e.g. experimental group)	Participant (case)	GROUP B (e.g. control group)
1	13	11	5
2	12	12	8
3	10	13	6
4	7	14	9
5	5	15	3
6	9	16	6
7	5	17	5
8	14	18	4
9	12		
10	16		
Mean =	10.30	Mean =	5.75

statisticians the *mean*. So the design can be used to assess whether, say, the average time taken by males getting ready for a first date is different from the average time taken by females.

This is the basic version of a whole range of statistical procedures which compare the average scores in different groups in more complex research designs. The **analysis of variance** (ANOVA) involves a whole family of different research designs based on this basic principle. Look at Figure 1.1 for more information.

■ The correlational design

The basic correlational design is one in which the researcher measures several different things at the same time, using a single group of participants. These things which are measured might be gender, age, IQ, extraversion and dogmatism. This basic correlational design is illustrated in Table 1.2.

The statistical analysis of this sort of design is usually based on the **correlation coefficient** or some other closely related statistical procedure based on the correlation coefficient. A correlation coefficient is a numerical index of the size of the relationship between two measures. The data from a correlational design may be analysed using a variety of statistics, as can be seen in Figure 1.1.

Correlational designs are sometimes called cross-sectional studies. They can be more complex, for example, when the researcher adds a time (temporal) dimension to the research design. There are special statistics to deal with these more complex designs (e.g. causal modelling such as LISREL) but these are essentially correlational in nature.

It would be misleading to pretend that the above covers every available statistical technique but a surprising range of statistics can be better understood if the underlying research design is clear to the researcher. Also remember that statistics is a mature discipline in its own right so it is unrealistic to assume that there are instant shortcuts to mastery of statistics in psychology. Getting basic concepts clear goes a long way towards this mastery, as does some experience.

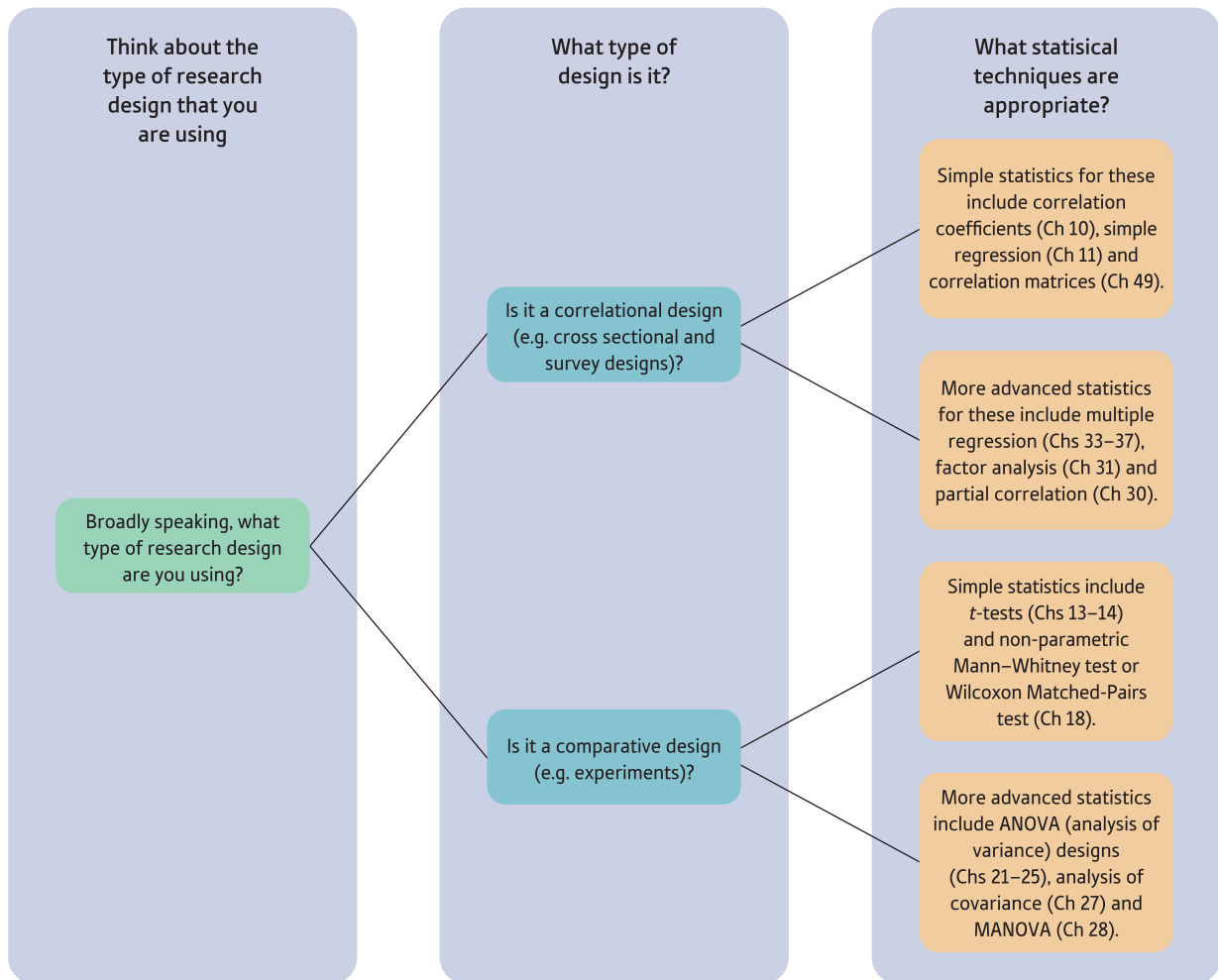


FIGURE 1.1

Fundamentals of design and statistical analysis

Table 1.2

The basic correlational design

PARTICIPANT	GENDER	AGE	IQ	EXTRAVERSION	DOGMATISM
1	Female	26	110	15	9
2	Male	31	130	19	6
3	Female	25	160	22	4
4	Female	22	110	34	8
5	Male	33	170	12	3
6	Female	28	140	17	7
7	Male	29	90	16	6
8	Male	34	130	22	5
9	Female	23	80	26	4
10	Male	27	70	11	2

1.3 The different types of variables in statistics

One's ability to use statistics in a practical context will be made much easier if some basic stuff is learnt about the fundamental different types of variables in statistics. Different types of variables require different kinds of statistical techniques for their analysis. So there are two basic questions that need to be asked:

- What types of variables do I have?
- What statistical tests analyse the data from these variables in the way that I want?

Fortunately, there are just two main types of data, so this is relatively straightforward. On the other hand, there are many different statistical tests and techniques. Of course, the way to learn about each of these is to gain some experience trying each of them out by working through the chapters which follow in this book. Most of the chapters in this book cover just one statistical technique or test in each chapter. The important thing is that each chapter tells you exactly what sorts of data (variables) are appropriate for that test or technique – and then how to do the analysis using a computer.

■ Types of variable

For all practical purposes, variables can be classified as being of *two* types (see Figure 1.2):

- *Score variables* Some variables are scores. A score is when a numerical value is given to a variable for each case in the sample. This numerical value indicates the quantity or amount of the characteristic (variable) in question. So age is a score variable since the numerical value indicates an increasing amount of the variable age. One could also describe this variable as quantitative.
- *Nominal or category or categorical variables* Some variables are measured by classifying cases into one of several named categories. These are also known as nominal, categorical or category variables. A better name for them might be qualitative variables because they measure the qualities of things rather than their quantities. For example, gender has two named categories – male and female. Nationality is another example: English, Welsh, Irish and Scottish are the nationalities of people of Britain. They have *no* numerical implications as such. To say that a person is Scottish is simply to put them into that named category. There is one risk of confusion – categories such as gender are usually entered into SPSS Statistics using different numbers to represent the different categories. For example, the variable gender has two categories – males could be represented by the number 1 and females by the number 2 (or vice versa). The numbers used are arbitrary – it could be 1002 and 2005 if the researcher desired. It is vital not to confuse these numbers, which merely represent different coding categories or qualities, with scores, which indicate the quantity of a variable which characterizes each participant in the research. For this reason, it is important to label the different values of nominal variables in full in the SPSS data spreadsheet since the number codes, in themselves, mean nothing. This is easily done, as is shown on pages 25–26.

■ The alternative traditional classification system

Sometimes variables are classified as nominal, ordinal, interval and ratio. Most textbooks go into arcane explanations about the difference between ordinal, interval and **ratio data**. This is mainly of conceptual interest and of little practical significance in

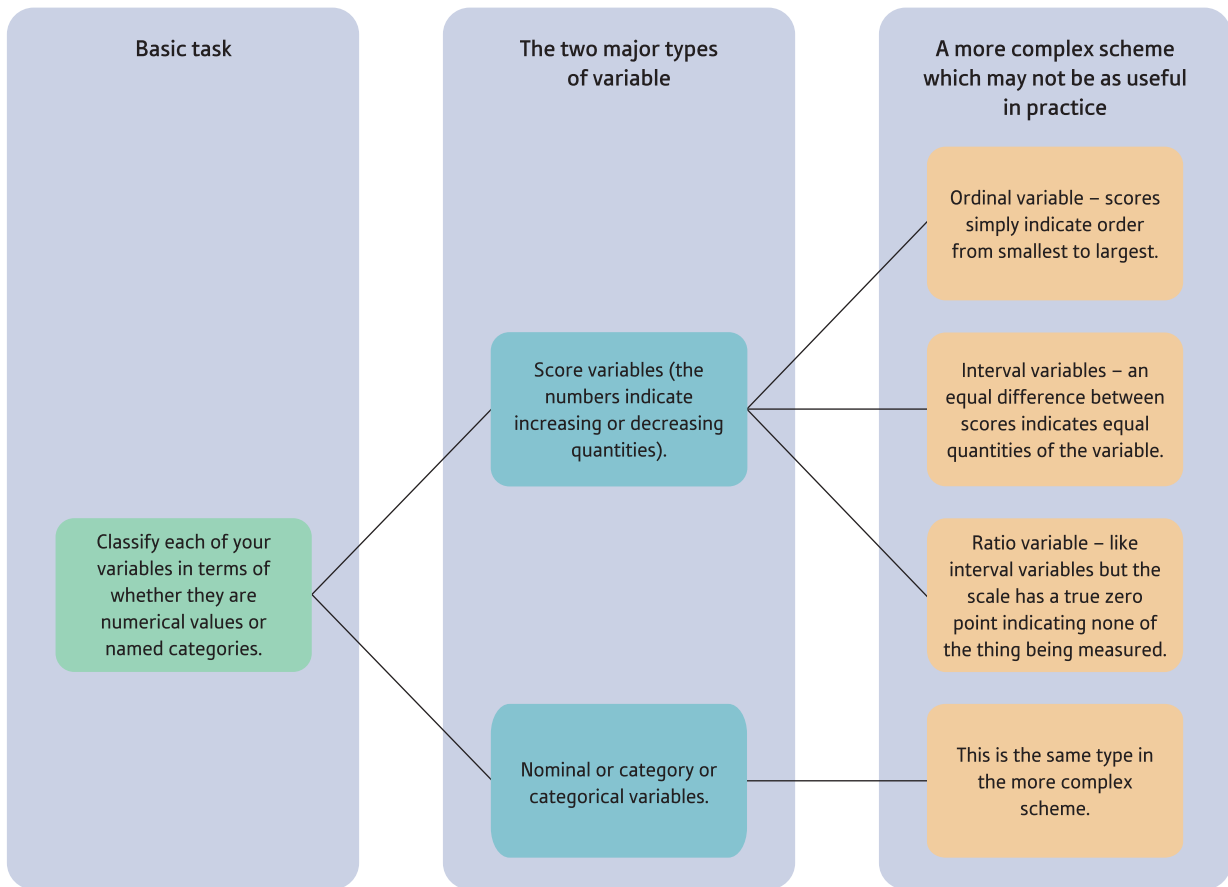


FIGURE 1.2

The two schemes for different types of variables

selecting appropriate statistics. Generally speaking, we would advise that this system should be ignored because it does not correspond with modern practice and it causes great confusion. *Nominal* is exactly the same as our classification of nominal (category) data and is important since a particular set of statistical techniques is called for to analyse category data. Of the other three, *interval* measurement is the most important. Interval measurement is where the steps on the scale of measurement are equal (just as centimetre steps on a rule are equal). Some psychologists are inclined to the view that this scale of measurement should reflect the underlying psychological variable being measured. Unfortunately, it is very difficult (i.e. impossible) to identify whether a psychological measure has equal intervals but, nevertheless, it is a sort of holy grail to them. Others, ourselves included, take the view that, so long as the numerical scale on which variables are measured has equal intervals (which is always the case except for nominal or category data, of course, from this perspective), then there is no problem as it is these numbers on which the statistical calculation is based and not some mystical underlying measurement scale. However, as a concession, we have mentioned equality of intervals as being desirable from time to time in the text. *Ratio* measures have equal intervals and a zero point, which means one can calculate ratios and make statements such as: 'One score is twice as big as another score.' Unfortunately, yet again, it is impossible to identify

any psychological variables which definitely are measured on a ratio measurement scale. Finally, **ordinal data** is data which does not have equal intervals, so scores only give the rank order of scores. Since the size of the intervals do not matter for ordinal data, then it is assumed that any psychological score data corresponds to the ordinal measurement scale at a minimum. For this reason, some psychologists have advocated the use of non-parametric (distribution-free) statistics for the analysis of much psychological data. The problem is that these techniques are not so powerful or flexible as most statistics employed in psychological research. You will find an extended discussion of ordinal, interval and ratio scales of measurement in Howitt, D. and Cramer, D. (2014) *Introduction to Research Methods in Psychology*, Harlow: Pearson.

■ Importance of deciding the types of variable involved

It is essential to decide for each of your variables whether it is a nominal (category) variable or a score variable. Write a list of your variables and classify each of them if you are a beginner. Eventually, you will do it somewhat automatically and usually without much thought. The statistical techniques which are appropriate for score variables are generally inappropriate for nominal or category variables because they measure qualities. So, for example, it is appropriate to calculate the mean (numerical average) of any variable which is a score (e.g. average age). On the other hand, it is totally inappropriate to calculate the mean (average) for variables which consist of categories. It would be nonsense to say that the average nationality is 1.7 since nationality is not a score. The problem is that SPSS works with the numbers in the data spreadsheet and does not know whether they are scores or numerical codes for different categories. (Though SPSS does allow you to classify your variables as ordinal or nominal.)

Do not be surprised to find that all or nearly all of your variables are scores. Psychologists have a predilection for measuring their variables as scores and it is likely that this rubs off on psychology students. What this means is that statistics for nominal or category data are often not needed, thus simplifying the task of data analysis.

1.4 Descriptive and inferential statistics compared

■ The difference between descriptive and inferential statistics

There are two main types of statistical techniques – *descriptive* and *inferential* statistics:

- **Descriptive statistics** chiefly describe the main features of individual variables. So calculating the average age of a sample of people is an example of descriptive statistics. Counting the number of English people would be another example of descriptive statistics. If one variable is considered at a time this is known as **univariate** statistics. **Bivariate** statistics are used when the relationship between two (or more) variables is being described.
- Inferential statistics is a totally distinct aspect of statistics. It only addresses the question of whether one can rely on the findings based on a *sample* of cases rather than *all* cases. The use of samples is characteristic of nearly all modern research. The problem with samples is that some of them are not similar to the populations