

Seventh Edition

Introduction to SPSS in Psychology

Dennis Howitt and Duncan Cramer



Introduction to SPSS in Psychology

Visit the *Introduction to SPSS in Psychology, seventh 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 23 and earlier

Seventh 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)
Sixth edition published 2014 (print and electronic)
Seventh edition published 2017 (print and electronic)

© Prentice Hall 1997 (print)
© Pearson Education Limited 2000, 2002, 2005, 2008, 2011 (print)
© Pearson Education Limited 2014, 2017 (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, Barnard's Inn, 86 Fetter Lane, London EC4A 1EN.

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 authors' and the publisher's 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.

Pearson Education is not responsible for the content of third-party internet sites.
ISBN: 978-1-292-18666-5 (print)
978-1-292-18669-6 (PDF)
978-1-292-18670-2 (ePub)

British Library Cataloguing-in-Publication Data

A catalogue record for the print edition is available from the British Library

Library of Congress Cataloguing-in-Publication Data

Names: Howitt, Dennis, author. | Cramer, Duncan, 1948- author.
Title: Introduction to SPSS in psychology : for version 23 and earlier /
Dennis Howitt, Loughborough University, Duncan Cramer, Loughborough University.
Description: Seventh edition. | Harlow, England ; New York : Pearson, 2017.
Identifiers: LCCN 2017007028 | ISBN 9781292186665 (print) | ISBN 9781292186696
(pdf) | ISBN 9781292186702 (epub)
Subjects: LCSH: Psychometrics. | SPSS for Windows. | Psychometrics--Computer programs.
Classification: LCC BF39 .H718 2017 | DDC 150.1/5195--dc23
LC record available at <https://lcn.loc.gov/2017007028>

10 9 8 7 6 5 4 3 2 1
21 20 19 18 17

Print edition typeset in 9.5/12pt Sabon LT Pro by Spi Global
Printed in Slovakia by Neografia

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>
Part 1 Introduction to SPSS in Psychology	1
1 Brief introduction to statistics	3
2 Basics of SPSS data entry and statistical analysis	17
Part 2 Descriptive statistics	31
3 Describing variables tabularly	33
4 Describing variables diagrammatically	40
5 Describing variables numerically: Averages, variation and spread	55
6 Shapes of distributions of scores	65
7 Relationships between two or more variables Tables	76
8 Relationships between two or more variables Diagrams	86
9 Correlation coefficients: Pearson's correlation and Spearman's rho	99
10 Regression: Prediction with precision	113
Part 3 Significance testing and basic inferential tests	125
11 Related <i>t</i> -test: Comparing two samples of correlated/related/paired scores	127
12 Unrelated <i>t</i> -test: Comparing two groups of unrelated/uncorrelated scores	136
13 Confidence intervals	144
14 Chi-square: Differences between unrelated samples of frequency data	148
15 McNemar's test: Differences between related samples of frequency data	163
16 Ranking tests for two groups: Non-parametric statistics	170
17 Ranking tests for three or more groups: Non-parametric statistics	181
Part 4 Analysis of variance	191
18 One-way analysis of variance (ANOVA) for unrelated/uncorrelated scores	193
19 Analysis of variance for correlated scores or repeated measures	201
20 Two-way analysis of variance for unrelated/uncorrelated scores	210
21 Multiple comparisons in ANOVA	223
22 Two-way analysis of variance for correlated scores or repeated measures	231
23 Two-way mixed analysis of variance	244
24 One-way analysis of covariance (ANCOVA)	254
25 One-way multivariate analysis of variance (MANOVA)	265

Part 5	More advanced statistics	275
26	Partial correlation	277
27	Factor analysis	284
28	Item reliability and inter-rater agreement	297
29	Stepwise multiple regression	310
30	Simultaneous or standard multiple regression	321
31	Simple mediational analysis	334
32	Hierarchical multiple regression	344
33	Log-linear analysis	354
34	Meta-analysis	363
Part 6	Data handling procedures	373
35	Missing values	375
36	Recoding values	382
37	Computing a scale score with some values missing	390
38	Computing a new group variable from existing group variables	395
39	Selecting cases	404
40	Reading ASCII or text files into the 'Data Editor'	410
	<i>Glossary</i>	417
	<i>Index</i>	424

Contents

<i>Introduction</i>	xxvi
<i>Acknowledgements</i>	xxx

Part 1 Introduction to SPSS in Psychology 1

1	Brief introduction to statistics	3
	<i>Overview</i>	3
1.1	Basic statistical concepts essential in SPSS analyses	4
1.2	Basic research designs: comparative versus correlational designs	4
1.3	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	12
2	Basics of SPSS data entry and statistical analysis	17
	<i>Overview</i>	17
2.1	What is SPSS?	18
2.2	Accessing SPSS	18
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 memory device	22
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'	26
2.10	Simple statistical calculation with SPSS	28
2.11	SPSS output	29
	<i>Summary of SPSS steps for a statistical analysis</i>	29

Part 2 Descriptive statistics 31

3	Describing variables tabularly	33
	<i>Overview</i>	33
3.1	What are tables?	34
3.2	When to use tables	35
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	Data to be analysed	36
3.7	Entering summarised categorical or frequency data by weighting	36
3.8	Percentage frequencies	38
3.9	Interpreting the output	38
3.10	Reporting the output	39
	<i>Summary of SPSS steps for frequency tables</i>	39
4	Describing variables diagrammatically	40
	<i>Overview</i>	40
4.1	What are diagrams?	41
4.2	When to use diagrams	42
4.3	When not to use diagrams	42
4.4	Data requirements for diagrams	42
4.5	Problems in the use of diagrams	42
4.6	Data to be analysed	43
4.7	Entering summarised categorical or frequency data by weighting	43
4.8	Pie diagram of category data	46
4.9	Adding labels to the pie diagram and removing the legend and label	47
4.10	Changing the colour of a pie-diagram slice to a black-and-white pattern	49
4.11	Bar chart of category data	51
4.12	Histograms	52
	<i>Summary of SPSS steps for charts</i>	54
5	Describing variables numerically: Averages, variation and spread	55
	<i>Overview</i>	55
5.1	What are averages, variation and spread?	56
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	60
5.6	Data to be analysed	61

5.7	Entering the data	61
5.8	Mean, median, mode, standard deviation, variance and range	62
5.9	Interpreting the output	63
5.10	Other features	63
5.11	Reporting the output	64
	<i>Summary of SPSS steps for descriptive statistics</i>	64
6	Shapes of distributions of scores	65
	<i>Overview</i>	65
6.1	What are the different shapes of scores?	66
6.2	When to use histograms and frequency tables of scores	69
6.3	When not to use histograms and frequency tables of scores	70
6.4	Data requirements for using histograms and frequency tables of scores	70
6.5	Problems in using histograms and frequency tables of scores	70
6.6	Data to be analysed	70
6.7	Entering the data	71
6.8	Frequency tables	71
6.9	Interpreting the output	72
6.10	Histograms	73
6.11	Interpreting the output	74
	<i>Summary of SPSS steps for frequency distributions</i>	75
7	Relationships between two or more variables Tables	76
	<i>Overview</i>	76
7.1	What tables are used to show relationships between variables?	77
7.2	When to use tables to show relationships between variables	79
7.3	When not to use tables to show relationships between variables	79
7.4	Data requirements for tables to show relationships between variables	80
7.5	Problems in the use of tables to show relationships between variables	80
7.6	Data to be analysed	80
7.7	Entering the data	81
7.8	Weighting the data	82
7.9	Crosstabulation with frequencies	83
7.10	Displaying frequencies as a percentage of the total number	84
7.11	Displaying frequencies as a percentage of the column total	85
	<i>Summary of SPSS steps for contingency tables</i>	85
8	Relationships between two or more variables Diagrams	86
	<i>Overview</i>	86
8.1	What diagrams are used to show relationships between variables?	87
8.2	When to use diagrams to show relationships between variables	90
8.3	When not to use diagrams to show relationships between variables	90

8.4	Data requirements for diagrams to show relationships between variables	90
8.5	Problems in the use of diagrams to show relationships between variables	91
8.6	Data to be analysed	91
8.7	Entering the data	92
8.8	Weighting the data	93
8.9	Compound (stacked) percentage bar chart	94
8.10	Compound (clustered) bar chart	96
	<i>Summary of SPSS steps for bar charts</i>	98
9	Correlation coefficients: Pearson's correlation and Spearman's rho	99
	<i>Overview</i>	99
9.1	What is a correlation coefficient?	100
9.2	When to use Pearson and Spearman rho correlation coefficients	103
9.3	When not to use Pearson and Spearman rho correlation coefficients	103
9.4	Data requirements for Pearson and Spearman rho correlation coefficients	103
9.5	Problems in the use of correlation coefficients	104
9.6	Data to be analysed	104
9.7	Entering the data	105
9.8	Pearson's correlation	105
9.9	Interpreting the output	106
9.10	Spearman's rho	107
9.11	Interpreting the output	107
9.12	Scatter diagram	108
9.13	Interpreting the output	110
9.14	Scattergram with more than one case with the same two values	110
	<i>Summary of SPSS steps for correlation</i>	112
10	Regression: Prediction with precision	113
	<i>Overview</i>	113
10.1	What is simple regression?	114
10.2	When to use simple regression	116
10.3	When not to use simple regression	116
10.4	Data requirements for simple regression	116
10.5	Problems in the use of simple regression	117
10.6	Data to be analysed	117
10.7	Entering the data	118
10.8	Simple regression	118
10.9	Interpreting the output	119
10.10	Regression scatterplot	120
10.11	Interpreting the output	123
	<i>Summary of SPSS steps for simple regression</i>	124

Part 3	Significance testing and basic inferential tests	125
11	Related <i>t</i> -test: Comparing two samples of correlated/related/paired scores	127
	<i>Overview</i>	127
11.1	What is the related <i>t</i> -test?	128
11.2	When to use the related <i>t</i> -test	130
11.3	When not to use the related <i>t</i> -test	131
11.4	Data requirements for the related <i>t</i> -test	131
11.5	Problems in the use of the related <i>t</i> -test	131
11.6	Data to be analysed	132
11.7	Entering the data	132
11.8	Related <i>t</i> -test	133
11.9	Interpreting the output	133
	<i>Summary of SPSS steps for related <i>t</i>-test</i>	135
12	Unrelated <i>t</i> -test: Comparing two groups of unrelated/uncorrelated scores	136
	<i>Overview</i>	136
12.1	What is the unrelated <i>t</i> -test?	137
12.2	When to use the unrelated <i>t</i> -test	138
12.3	When not to use the unrelated <i>t</i> -test	138
12.4	Data requirements for the unrelated <i>t</i> -test	139
12.5	Problems in the use of the unrelated <i>t</i> -test	139
12.6	Data to be analysed	139
12.7	Entering the data	139
12.8	Unrelated <i>t</i> -test	141
12.9	Interpreting the output	141
	<i>Summary of SPSS steps for unrelated <i>t</i>-test</i>	143
13	Confidence intervals	144
	<i>Overview</i>	144
13.1	What are confidence intervals?	145
13.2	Relationship between significance and confidence intervals	146
13.3	Confidence intervals and limits in SPSS	147
14	Chi-square: Differences between unrelated samples of frequency data	148
	<i>Overview</i>	148
14.1	What is chi-square?	149
14.2	When to use chi-square	151
14.3	When not to use chi-square	151
14.4	Data requirements for chi-square	152
14.5	Problems in the use of chi-square	152

14.6	Data to be analysed	153
14.7	Entering the data using the 'Weighting Cases' procedure	153
14.8	Entering the data case by case	154
14.9	Chi-square	155
14.10	Interpreting the output for chi-square	156
14.11	Fisher's exact test	158
14.12	Interpreting the output for Fisher's exact test	158
14.13	One-sample chi-square	159
14.14	Interpreting the output for a one-sample chi-square	161
14.15	Chi-square without ready-made tables	161
	<i>Summary of SPSS steps for chi-square</i>	162
15	McNemar's test: Differences between related samples of frequency data	163
	<i>Overview</i>	163
15.1	What is McNemar's test?	164
15.2	When to use McNemar's test	164
15.3	When not to use McNemar's test	165
15.4	Data requirements for McNemar's test	165
15.5	Problems in the use of McNemar's test	165
15.6	Data to be analysed	165
15.7	Entering the data using the 'Weighting Cases' procedure	166
15.8	Entering the data case by case	167
15.9	McNemar's test	167
15.10	Interpreting the output for McNemar's test	168
	<i>Summary of SPSS steps for McNemar's test</i>	169
16	Ranking tests for two groups: Non-parametric statistics	170
	<i>Overview</i>	170
16.1	What are non-parametric tests?	171
16.2	When to use non-parametric tests	173
16.3	When not to use non-parametric tests	173
16.4	Data requirements for non-parametric tests	173
16.5	Problems in the use of non-parametric tests	173
16.6	Data to be analysed	174
16.7	Entering the data	174
16.8	Related scores: Sign test	175
16.9	Interpreting the output for the sign test	175
16.10	Related scores: Wilcoxon test	176
16.11	Interpreting the output for the Wilcoxon test	176
16.12	Unrelated scores: Mann-Whitney <i>U</i> -test	177
16.13	Entering the data	177
16.14	Mann-Whitney <i>U</i> -test	178

16.15	Interpreting the output for the Mann–Whitney <i>U</i> -test	179
	<i>Summary of SPSS steps for non-parametric tests for two groups</i>	180
17	Ranking tests for three or more groups: Non-parametric statistics	181
	<i>Overview</i>	181
17.1	What are ranking tests?	182
17.2	When to use ranking tests	183
17.3	When not to use ranking tests	183
17.4	Data requirements for ranking tests	183
17.5	Problems in the use of ranking tests	183
17.6	Data to be analysed	183
17.7	Friedman three or more related samples test	184
17.8	Entering the data for the Friedman test	184
17.9	Friedman test	185
17.10	Interpreting the output for the Friedman test	185
17.11	Kruskal–Wallis three or more unrelated samples test	186
17.12	Entering the data for the Kruskal–Wallis test	187
17.13	Kruskal–Wallis test	188
17.14	Interpreting the output for the Kruskal–Wallis test	189
	<i>Summary of SPSS steps for non-parametric tests for three or more groups</i>	189

Part 4 Analysis of variance

191

18	One-way analysis of variance (ANOVA) for unrelated or uncorrelated scores	193
	<i>Overview</i>	193
18.1	What is one-way unrelated ANOVA?	194
18.2	When to use one-way unrelated ANOVA	195
18.3	When not to use one-way unrelated ANOVA	196
18.4	Data requirements for one-way unrelated ANOVA	196
18.5	Problems in the use of one-way unrelated ANOVA	196
18.6	Data to be analysed	196
18.7	Entering the data	197
18.8	One-way unrelated ANOVA	197
18.9	Interpreting the output	198
	<i>Summary of SPSS steps for one-way unrelated ANOVA</i>	199
19	One-way analysis of variance for correlated scores or repeated measures	201
	<i>Overview</i>	201
19.1	What is one-way repeated-measures ANOVA?	202
19.2	When to use repeated-measures ANOVA	203

19.3	When not to use one-way repeated-measures ANOVA	203
19.4	Data requirements for one-way repeated-measures ANOVA	204
19.5	Problems in the use of one-way repeated-measures ANOVA	204
19.6	Data to be analysed	204
19.7	Entering the data	204
19.8	One-way repeated-measures ANOVA	205
19.9	Interpreting the output	206
	<i>Summary of SPSS steps for one-way repeated-measures ANOVA</i>	209
20	Two-way analysis of variance for unrelated/uncorrelated scores	210
	<i>Overview</i>	210
20.1	What is two-way unrelated ANOVA?	211
20.2	When to use two-way unrelated ANOVA	214
20.3	When not to use two-way unrelated ANOVA	214
20.4	Data requirements for two-way unrelated ANOVA	214
20.5	Problems in the use of two-way unrelated ANOVA	215
20.6	Data to be analysed	216
20.7	Entering the data	216
20.8	Two-way unrelated ANOVA	217
20.9	Interpreting the output	218
20.10	Editing the graph	220
	<i>Summary of SPSS steps for two-way unrelated ANOVA</i>	221
21	Multiple comparisons in ANOVA	223
	<i>Overview</i>	223
21.1	What is multiple-comparisons testing?	224
21.2	When to use multiple-comparisons tests	225
21.3	When not to use multiple-comparisons tests	225
21.4	Data requirements for multiple-comparisons tests	225
21.5	Problems in the use of multiple-comparisons tests	226
21.6	Data to be analysed	226
21.7	Entering the data	227
21.8	Multiple-comparisons tests	227
21.9	Interpreting the output	228
21.10	Reporting the output	229
	<i>Summary of SPSS steps for multiple-comparison tests</i>	230
22	Two-way analysis of variance for correlated scores or repeated measures	231
	<i>Overview</i>	231
22.1	What is two-way repeated-measures ANOVA?	232
22.2	When to use two-way repeated-measures ANOVA	234

22.3	When not to use two-way repeated-measures ANOVA	235
22.4	Data requirements for two-way related-measures ANOVA	235
22.5	Problems in the use of two-way repeated-measures ANOVA	235
22.6	Data to be analysed	235
22.7	Entering the data	236
22.8	Two-way repeated-measures ANOVA	236
22.9	Interpreting the output	238
22.10	Reporting the output	242
	<i>Summary of SPSS steps for two-way repeated-measures ANOVA</i>	242
23	Two-way mixed analysis of variance	244
	<i>Overview</i>	244
23.1	What is two-way mixed ANOVA?	245
23.2	When to use two-way mixed ANOVA	245
23.3	When not to use two-way mixed ANOVA	246
23.4	Data requirements for two-way mixed ANOVA	247
23.5	Problems in the use of two-way mixed ANOVA	247
23.6	Data to be analysed	247
23.7	Entering the data	247
23.8	Two-way mixed ANOVA	248
23.9	Interpreting the output	250
23.10	Reporting the output	251
	<i>Summary of SPSS steps for mixed ANOVA</i>	252
24	One-way analysis of covariance (ANCOVA)	254
	<i>Overview</i>	254
24.1	What is one-way analysis of covariance (ANCOVA)?	255
24.2	When to use one-way ANCOVA	256
24.3	When not to use one-way ANCOVA	256
24.4	Data requirements for one-way ANCOVA	257
24.5	Problems in the use of one-way ANCOVA	257
24.6	Data to be analysed	257
24.7	Entering the data	257
24.8	One-way ANCOVA	258
24.9	Testing that the slope of the regression line within cells is similar	259
24.10	Interpreting the output	259
24.11	Testing the full model	260
24.12	Interpreting the output	262
24.13	Reporting the output	263
	<i>Summary of SPSS steps for one-way ANCOVA</i>	263

25	One-way multivariate analysis of variance (MANOVA)	265
	<i>Overview</i>	265
25.1	What is one-way multivariate analysis of variance (MANOVA)?	266
25.2	When to use one-way MANOVA	267
25.3	When not to use one-way MANOVA	268
25.4	Data requirements for one-way MANOVA	269
25.5	Problems in the use of one-way MANOVA	269
25.6	Data to be analysed	269
25.7	Entering the data	270
25.8	One-way MANOVA	270
25.9	Interpreting the output	271
25.10	Reporting the output	274
	<i>Summary of SPSS steps for one-way MANOVA</i>	274

Part 5 More advanced statistics

275

26	Partial correlation	277
	<i>Overview</i>	277
26.1	What is partial correlation?	278
26.2	When to use partial correlation	280
26.3	When not to use partial correlation	280
26.4	Data requirements for partial correlation	280
26.5	Problems in the use of partial correlation	280
26.6	Data to be analysed	280
26.7	Entering the data	281
26.8	Partial correlation	281
26.9	Interpreting the output	282
26.11	Reporting the output	283
	<i>Summary of SPSS steps for partial correlation</i>	283
27	Factor analysis	284
	<i>Overview</i>	284
27.1	What is factor analysis?	285
27.2	When to use factor analysis	287
27.3	When not to use factor analysis	288
27.4	Data requirements for factor analysis	288
27.5	Problems in the use of factor analysis	288
27.6	Data to be analysed	289
27.7	Entering the data	289

27.8	Principal components analysis with orthogonal rotation	290
27.9	Interpreting the output	293
27.10	Reporting the output	295
	<i>Summary of SPSS steps for factor analysis</i>	296
28	Item reliability and inter-rater agreement	297
	<i>Overview</i>	297
28.1	What are item reliability and inter-rater agreement?	298
28.2	When to use item reliability and inter-rater agreement	300
28.3	When not to use item reliability and inter-rater agreement	301
28.4	Data requirements for item reliability and inter-rater agreement	301
28.5	Problems in the use of item reliability and inter-rater agreement?	302
28.6	Data to be analysed for item alpha reliability	302
28.7	Entering the data	302
28.8	Alpha reliability	303
28.9	Interpreting the output	304
28.10	Split-half reliability	305
28.11	Interpreting the output	305
28.12	Data to be analysed for inter-rater agreement (kappa)	306
28.13	Entering the data	306
28.14	Kappa	307
28.15	Interpreting the output	308
	<i>Summary of SPSS steps for reliability</i>	309
29	Stepwise multiple regression	310
	<i>Overview</i>	310
29.1	What is stepwise multiple regression?	311
29.2	When to use stepwise multiple regression	312
29.3	When not to use stepwise multiple regression	313
29.4	Data requirements for stepwise multiple regression	314
29.5	Problems in the use of stepwise multiple regression	314
29.6	Data to be analysed	314
29.7	Entering the data	315
29.8	Stepwise multiple regression analysis	315
29.9	Interpreting the output	316
29.10	Reporting the output	319
	<i>Summary of SPSS steps for stepwise multiple regression</i>	319
30	Simultaneous or standard multiple regression	321
	<i>Overview</i>	321
30.1	What is simultaneous or standard multiple regression?	322
30.2	When to use simultaneous or standard multiple regression	325

30.3	When not to use simultaneous or standard multiple regression	326
30.4	Data requirements for simultaneous or standard multiple regression	326
30.5	Problems in the use of simultaneous or standard multiple regression	327
30.6	Data to be analysed	327
30.7	Entering the data	327
30.8	Simultaneous or standard multiple regression analysis	328
30.9	Interpreting the output	329
30.10	Reporting the output	331
	<i>Summary of SPSS steps for simultaneous or standard multiple regression</i>	333
31	Simple mediational analysis	334
	<i>Overview</i>	334
31.1	What is simple mediational analysis?	335
31.2	When to use simple mediational analysis	338
31.3	When not to use simple mediational analysis	338
31.4	Data requirements for a simple mediational analysis	339
31.5	Problems in the use of simple mediational analysis	339
31.6	Data to be analysed	339
31.7	Entering the data	339
31.8	Simultaneous multiple regression analysis	340
31.9	Interpreting the output	341
31.10	Reporting the output	342
	<i>Summary of SPSS steps for simultaneous or standard multiple regression</i>	343
32	Hierarchical multiple regression	344
	<i>Overview</i>	344
32.1	What is hierarchical multiple regression?	345
32.2	When to use hierarchical multiple regression	347
32.3	When not to use hierarchical multiple regression	347
32.4	Data requirements for hierarchical multiple regression	347
32.5	Problems in the use of hierarchical multiple regression	347
32.6	Data to be analysed	348
32.7	Entering the data	348
32.8	Hierarchical multiple regression analysis	349
32.9	Interpreting the output	350
32.10	Reporting the output	352
	<i>Summary of SPSS steps for hierarchical multiple regression</i>	353
33	Log-linear analysis	354
	<i>Overview</i>	354
33.1	What is log-linear analysis?	355
33.2	When to use log-linear analysis	356

33.3	When not to use log-linear analysis	357
33.4	Data requirements for log-linear analysis	358
33.5	Problems in the use of log-linear analysis	358
33.6	Data to be analysed	358
33.7	Entering the data	358
33.8	Log-linear analysis	359
33.9	Interpreting the output	360
33.10	Reporting the output	362
	<i>Summary of SPSS steps for log-linear analysis</i>	362
34	Meta-analysis	363
	<i>Overview</i>	363
34.1	What is meta-analysis?	364
34.2	When to use meta-analysis	367
34.3	When not to use meta-analysis	368
34.4	Data requirements for meta-analysis	368
34.5	Problems in the use of meta-analysis	369
34.6	Data to be analysed	369
34.7	Meta-analysis	369
34.8	Interpreting the output	371
34.9	Reporting the output	371

Part 6 Data handling procedures

373

35	Missing values	375
	<i>Overview</i>	375
35.1	What are missing values?	376
35.2	Entering the data	377
35.3	Defining missing values	378
35.4	Pairwise and listwise options	378
35.5	Sample output for pairwise exclusion	379
35.6	Sample output for listwise exclusion	380
35.7	Interpreting the output	380
35.8	Reporting the output	381
	<i>Summary of SPSS steps for handling missing values</i>	381
36	Recoding values	382
	<i>Overview</i>	382
36.1	What is recoding values?	383
36.2	Entering the data	383

36.3	Recoding values	384
36.4	Recoding missing values	387
36.5	Saving the recode procedure as a syntax file	387
36.6	Adding some extra cases to Table 36.1	388
36.7	Running the Recode syntax command	388
	<i>Summary of SPSS steps for recoding values</i>	388
37	Computing a scale score with some values missing	390
	<i>Overview</i>	390
37.1	What is computing a scale score with some values missing?	391
37.2	Entering the data	392
37.3	Computing a scale score with some values missing	393
37.4	Saving the Compute procedure as a syntax file	395
37.5	Adding some extra cases to Table 37.1	395
37.6	Running the Compute syntax command	396
	<i>Summary of SPSS steps for computing a scale score with some missing values</i>	396
38	Computing a new group variable from existing group variables	397
	<i>Overview</i>	397
38.1	What is computing a new group variable from existing group variables?	398
38.2	Entering the data	400
38.3	Syntax file for computing a new group variable from existing group variables	400
38.4	Running the Compute syntax commands	401
38.5	Computing a new group using menus and dialogue boxes	402
	<i>Summary of SPSS steps for computing a new group variable from existing group variables</i>	403
39	Selecting cases	404
	<i>Overview</i>	404
39.1	What is selecting cases?	405
39.2	Entering the data	406
39.3	Selecting cases	406
	<i>Summary of SPSS steps for selecting cases</i>	409

40	Reading ASCII or text files into the 'Data Editor'	410
	<i>Overview</i>	410
40.1	What is an ASCII or text data file?	411
40.2	Entering data into an ASCII or text data file	412
40.3	Reading an ASCII or text data file	413
	<i>Summary of SPSS steps for inputting an ASCII or text data file</i>	416
	<i>Glossary</i>	417
	<i>Index</i>	424

Companion Website

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



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

Statistical tests are built on assumptions. A common one is that the scores on a variable are normally distributed – bell-shaped.



So departures from the ideal bell shape result in tests of statistical significance not detecting significant differences as well as if the ideal shape was met.



Examining the distribution of scores will help indicate where such problems are likely.



Psychological data often depart from the ideal distribution or shape.

FIGURE 6.6 Influence of distribution shapes on tests of significance

Frequency table

- This is a simple table which indicates how frequent each observation is.
- Where there are many different values, a frequency table can be unwieldy unless you use ranges of values rather than individual values.

Histogram

- This is a pictorial way of representing the frequencies of scores.
- Ideally, each bar of the histogram should be proportionate in width to the range of scores that the bar indicates. Unfortunately, this is not always the case.
- The heights of the bars indicate the relative frequencies of scores.

FIGURE 6.8 Ideas that are important in the shapes of distributions

3	5	5	4	4	5	5	3	5	2
1	2	5	3	2	1	2	3	3	3
4	2	5	5	4	2	4	5	1	5
5	3	3	4	1	4	2	5	1	2
3	2	5	4	2	1	2	3	4	1

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

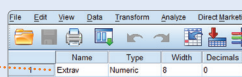
6.2 When to use histograms and frequency tables of scores

It is always highly desirable to obtain histograms and/or tables of frequencies of scores. This is an essential part of data analysis as conducted by a good researcher. Not to do so is basically incompetent, as many important bits of information are contained within these simple statistics. Authorities on statistics have often complained that psychologists fail to use simple statistical techniques to explore their data prior to more complex analyses. Too often this failing is encouraged by SPSS books which hardly mention descriptive statistics.

6.7 Entering the data

Step 1

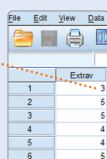
In Variable View of the 'Data Editor', name the variable 'Extrav'.
Remove the two decimal places.



Name	Type	Width	Decimals
Extrav	Numeric	8	0

Step 2

In 'Data View' of the 'Data Editor', enter the extraversion scores in the first column.

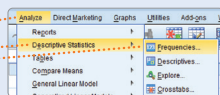


	Extrav
1	3
2	5
3	5
4	4
5	4
6	5

6.8 Frequency tables

Step 1

Select 'Analyze',
'Descriptive Statistics',
and 'Frequencies...'

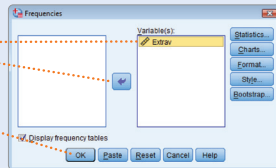


Step-by-step illustrations and screenshots of SPSS

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

Step 2

Select 'Extrav' and the button to put it in the 'Variable(s):' box. Select 'OK'.



6.9 Interpreting the output

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.

The first column shows the five values of extraversion which are labelled 1 to 5.

The second column shows the frequency of these values. There are 7 cases with a value of 1.

Extrav				
		Frequency	Percent	Cumulative Percent
Valid	1	7	14.0	14.0
	2	11	22.0	36.0
	3	10	20.0	56.0
	4	9	18.0	74.0
	5	13	26.0	100.0
	Total	50	100.0	

The fifth column adds these percentages together cumulatively down the table. So 56 percentage of the cases have values of 3 or less.

The third column expresses these frequencies as a percentage of the total number including missing data. Of all cases, 14 percentage have a value of 1.

The fourth column expresses these frequencies as a percentage of the total number excluding missing data. As there are no missing cases, the percentages are the same as in the third column.

3.10 Reporting the output

Only the category labels, the frequency and the percentage frequency need be reported. Consequently, you need to simplify this table if you are going to present it. If the occupation was missing for some of the cases, you would need to decide whether you would present percentages including or excluding them. Missing cases are discussed in detail in Chapter 33. There is no need to present both sets of figures. Also omit the term 'Valid' in the first column, as its meaning may only be familiar to SPSS users.

Summary of SPSS steps for frequency tables

Data

- In 'Variable View' of the 'Data Editor', 'name' the 'nominal' or categorical variable, 'label' each of its values and select 'nominal' as the 'measure'.
- Name the second variable.
- In 'Data View' of the 'Data Editor', enter the value of the category for each case and its frequency.
- Select 'Data', 'Weight Cases . . .', 'Weight cases by' and move the weighting variable to the 'Frequency Variable:' box.

Analysis

- Select 'Analyze', 'Descriptive Statistics' and 'Frequencies . . .'.
- Move the variables to be analyzed to the right-hand box.

Output

- The percentage of cases for each category is given for the whole sample in the 'Percent' column including any missing data, and in the 'Valid Percent' column excluding missing data. If there is no missing data, these percentages will be the same.

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

Summary of SPSS 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, 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, although this has had little effect on the basic SPSS material covered in this text.

This book is a stand-alone, step-by-step approach to statistical analysis using SPSS for Windows and is applicable to Releases 10 to 23 – that is, any version of SPSS that you are likely to use. 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 that 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 23.0. We shall refer to it throughout this book as SPSS because it is shorter, most users refer to it this way and the first letter of the original abbreviation refers to ‘Statistical’, so it seems redundant to repeat it.

This book updates the sixth edition of *Introduction to SPSS 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. We have made the book shorter and less weighty, which should make it even more user-friendly. We have omitted chapters that are less likely to be used. Although the examples are drawn from psychology, the variables can easily be renamed to make them more appropriate to those working in other fields. 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.

The basic structure of the major chapters is as follows:

- 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* (i.e. Release 23), but remains suitable for Releases 10 to 22 because of their similarity. Although SPSS was until recently 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 as follows:

- *Understanding Statistics in Psychology with SPSS* (7th edition) (Harlow: Pearson): 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.

- *Research methods in psychology* (2017) (5th edition) (Harlow: Pearson): 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. It includes numerous cross-references to this book and *Understanding Statistics in Psychology with SPSS*.

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 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. Figure 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 here. 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, *Understanding Statistics in Psychology with SPSS* (2017) (7th edition) (Harlow: Pearson). This book is referred to as *USP*, followed by the corresponding chapter or table number.

*Dennis Howitt
Duncan Cramer*

Acknowledgements

■ Authors' acknowledgements

We would like to thank everyone involved in the production of this book for their hard work, patience, and support. In particular, Janey Webb, who now rejoices in the title Specialist Portfolio Manager, has been a tower of strength throughout several revisions of the book offering sage advice at all times. While she was on maternity leave her place was taken by Laura Aboujieb who cajoled us into some of the changes to this new edition to make it accessible to a wider readership.

Kevin Ancient supplied the striking cover and this edition follows his immaculate text design which makes the book's structure so clear.

The production of the book was managed by SPi Global in Chennai, Tamil Nadu, India. Greetal Carolyn Jayanandan was the project manager for SPi Global and ensured a speedy and efficient production. Her colleagues, regrettably unnamed, did the essential tasks of copy editing and proofreading. We are also most grateful for the hard work and flair of Jennifer Sargunar, Managing Content Producer, for Pearson Education.

Finally, the book is better for the contribution of the academic reviewers who generously made comments and recommendations for improvements.

Dennis Howitt and Duncan Cramer

■ Publisher's acknowledgements

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

Screenshots

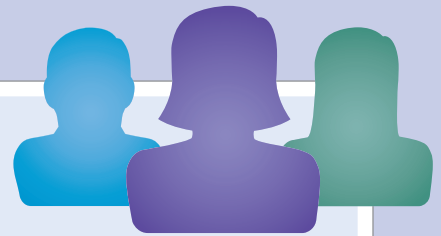
Screenshot 2.2 from PC Home Page Microsoft, Microsoft Corporation, Microsoft screenshot reprinted with permission from Microsoft Corporation; Screenshots 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 2.10, 2.11, 3.7, 3.8, 3.9, 4.7, 4.8, 4.9, 4.10, 4.11, 4.12, 5.7, 5.8, 5.8, 5.9, 6.7, 6.8, 6.9, 6.10, 7.7, 7.8, 7.9, 7.10, 7.11, 8.7, 8.8, 8.9, 8.10, 9.7, 9.8, 9.10, 9.12, 10.7, 10.8, 10.10, 11.7, 11.8, 12.7, 12.8, 14.3, 14.7, 14.8, 14.9, 15.7, 15.8, 15.9, 16.7, 16.8, 16.13, 16.14, 17.8, 17.9, 17.13, 17.14, 18.7, 18.8, 19.8, 19.10, 20.7, 20.8, 20.10, 21.7, 21.8, 22.7, 22.8, 23.7, 23.8, 24.7, 24.8, 24.9, 24.11, 25.7, 25.8, 26.7, 26.8, 27.7, 27.8, 28.7, 28.8, 28.10, 28.13, 28.14, 29.7, 29.8, 30.7, 30.8, 31.7, 31.8, 32.7, 32.8, 33.7, 33.8, 35.2, 35.3, 35.4, 36.2, 36.3, 36.5, 36.7, 37.2, 37.3, 37.4, 37.6, 38.2, 38.3, 38.4, 38.5, 39.2, 39.3, 40.2, 40.3 from IBM SPSS Statistics 23.0, IBM SPSS Inc., screenshots reprinted courtesy of International Business Machines Corporation,

© International Business Machines Corporation. SPSS was acquired by IBM in October, 2009. IBM, the IBM logo, ibm.com, and SPSS are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at 'IBM Copyright and trademark information' at www.ibm.com/legal/copytrade.shtml; Screenshot 34.7 from <http://www.lyonsmorris.com/contact.cfm> and screenshot 34.8 from <http://www.lyonsmorris.com/ma1/index.cfm>, Larry Lyons and Wendy Morris.

PART 1

Introduction to SPSS in Psychology





CHAPTER 1

Brief introduction to statistics

Overview

- Unfortunately, in statistics, there is no gain without some effort. 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.
- 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 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 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:

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). 'Case' is very much SPSS 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

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

by 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. See Figure 1.1 for more information.

■ 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. The 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.

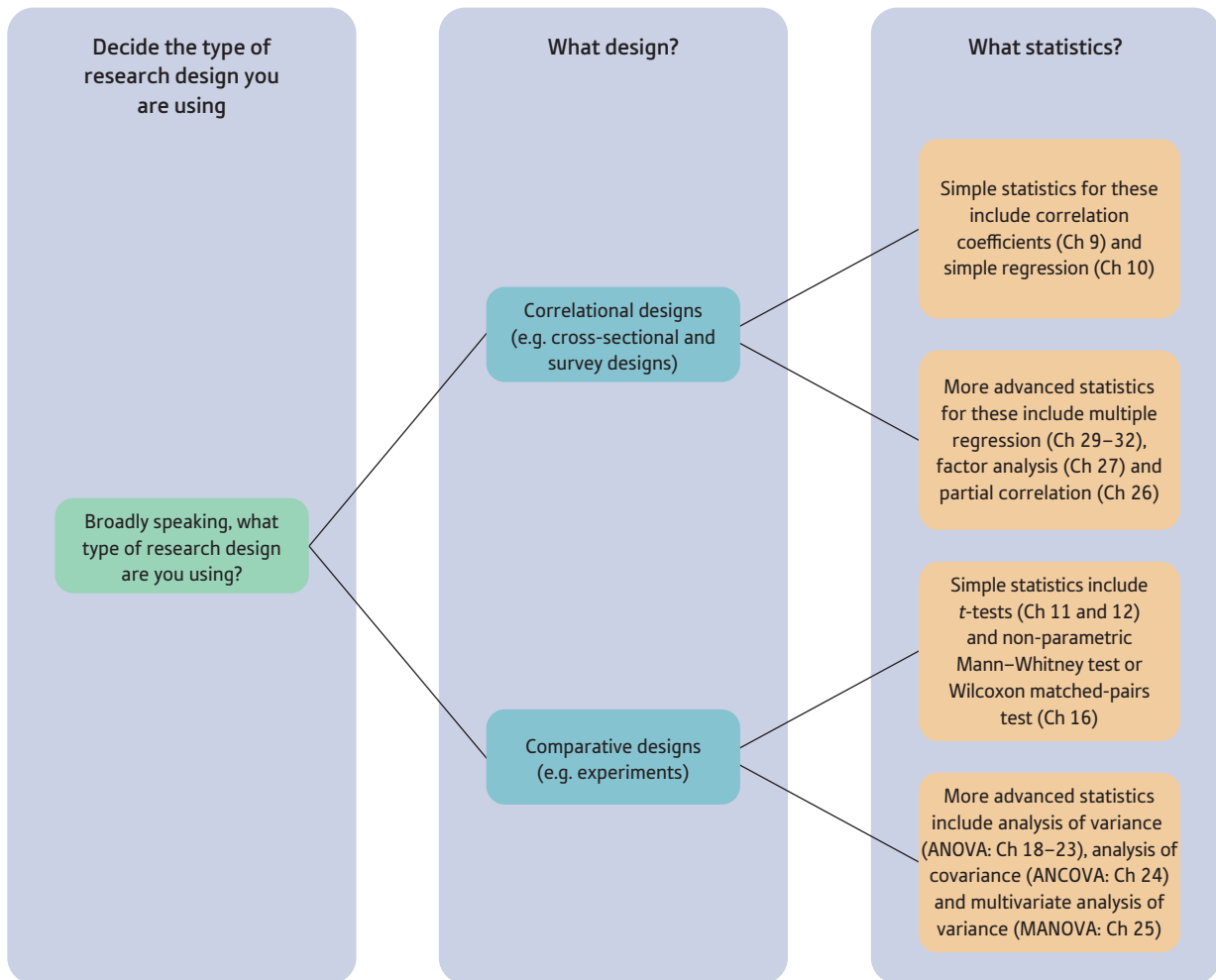


FIGURE 1.1

Fundamentals of design and statistical analysis

Table 1.2

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 Different types of variables in statistics

One's ability to use statistics in a practical context will be made much easier if some basic information 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. 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 variables

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 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 – they 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 characterises 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.

■ 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 selecting

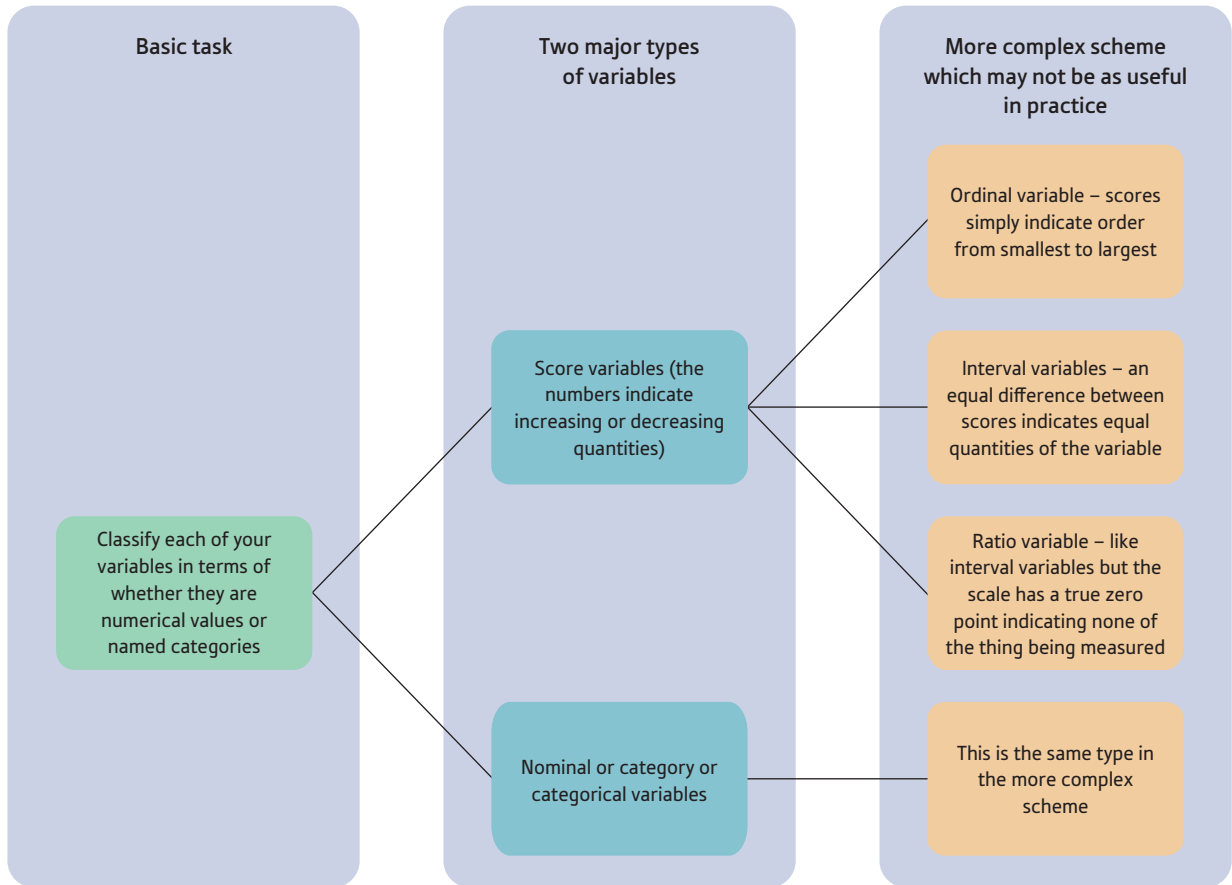


FIGURE 1.2

Two schemes for different types of variables

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 these psychologists. 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 just about 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 their rank order. Since the size of the intervals do not matter for ordinal data, 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. (2017) *research methods in psychology*, Harlow: Pearson.

■ Importance of deciding the types of variables 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 named 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

■ 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 from which they are taken. The phrases ‘statistically significant’ and ‘not statistically significant’ simply indicate that any trends in the data can be accepted as likely (i.e. statistically significant) or not likely enough to rely on (i.e. not statistically significant). A statistically