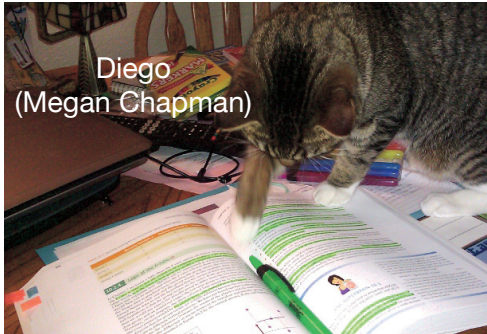
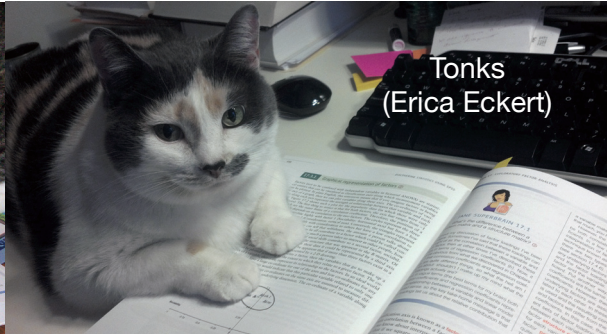


# DISCOVERING STATISTICS USING IBM SPSS STATISTICS

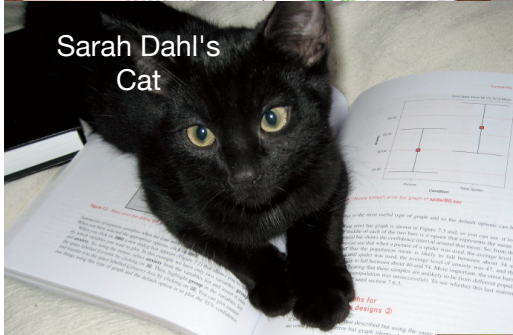
# CATISFIED CUSTOMERS



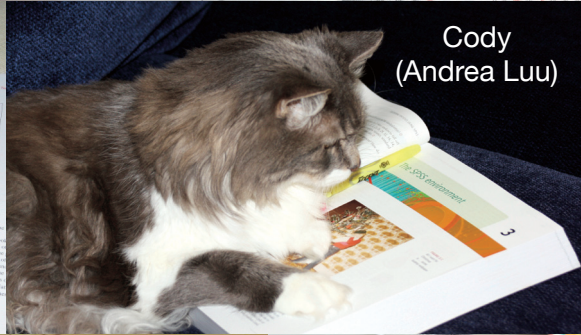
Diego  
(Megan Chapman)



Tonks  
(Erica Eckert)



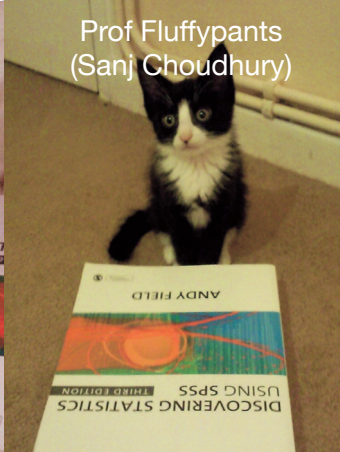
Sarah Dahl's  
Cat



Cody  
(Andrea Luu)



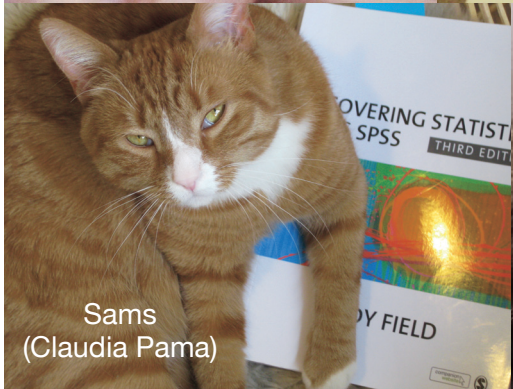
Coltrane  
(Krista Soria)



Prof Fluffypants  
(Sanj Choudhury)



Nightshade  
(Dave Verkuilj)



Sams  
(Claudia Pama)



Kamaji & Choccy  
(Emily Margaret-Gay)

# DISCOVERING STATISTICS USING IBM SPSS STATISTICS

AND SEX AND DRUGS AND ROCK 'N' ROLL

4TH EDITION



## ANDY FIELD

 SAGE

Los Angeles | London | New Delhi  
Singapore | Washington DC



Los Angeles | London | New Delhi  
Singapore | Washington DC

SAGE Publications Ltd  
1 Oliver's Yard  
55 City Road  
London EC1Y 1SP

SAGE Publications Inc.  
2455 Teller Road  
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd  
B 1/1 1 Mohan Cooperative Industrial Area  
Mathura Road  
New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd  
3 Church Street  
#10-04 Samsung Hub  
Singapore 049483

---

Editor: Michael Carmichael  
Development editor: Robin Lupton  
Production editor: Ian Antcliff  
Copyeditor: Richard Leigh  
Proofreaders: Louise Harnby; Kate Harrison;  
Anna Gilding; Jennifer Hinchliffe  
Indexer: David Rudeforth  
Marketing manager: Ben Griffin-Sherwood  
Cover design: Wendy Scott  
Typeset by:  
Printed by:

© Andy Field 2013

First edition published November 2002  
Second edition published 2005  
Third edition published 2009

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

**Library of Congress Control Number: 2012944559**

**British Library Cataloguing in Publication data**

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

ISBN 978-1-4462-4917-8  
ISBN 978-1-4462-4918-5 (pbk)

# CONTENTS

Preface	xix
How to use this book	xxv
Acknowledgements	xxx
Dedication	xxxii
Symbols used in this book	xxxiv
Some maths revision	xxxvi
<b>1 Why is my evil lecturer forcing me to learn statistics?</b>	<b>1</b>
1.1. What will this chapter tell me? ①	1
1.2. What the hell am I doing here? I don't belong here ①	2
1.2.1. The research process ①	3
1.3. Initial observation: finding something that needs explaining ①	4
1.4. Generating theories and testing them ①	4
1.5. Collect data to test your theory ①	7
1.5.1. Variables ①	7
1.5.2. Measurement error ①	11
1.5.3. Validity and reliability ①	12
1.5.4. Correlational research methods ①	13
1.5.5. Experimental research methods ①	14
1.5.6. Randomization ①	18
1.6. Analysing data ①	19
1.6.1. Frequency distributions ①	20
1.6.2. The centre of a distribution ①	21
1.6.3. The dispersion in a distribution ①	24
1.6.4. Using a frequency distribution to go beyond the data ①	28
1.6.5. Fitting statistical models to the data ①	32
1.7. Reporting data ①	34
1.7.1. Dissemination of research ①	34
1.7.2. Knowing how to report data ①	35
1.7.3. Some initial guiding principles ①	35
1.8. Brian's attempt to woo Jane ①	37
1.9. What next? ①	37
1.10. Key terms that I've discovered	38
1.11. Smart Alex's tasks	38
1.12. Further reading	39

<b>2</b>	<b>Everything you never wanted to know about statistics</b>	<b>40</b>
2.1.	What will this chapter tell me? ①	40
2.2.	Building statistical models ①	41
2.3.	Populations and samples ①	42
2.4.	Statistical models ①	44
2.4.1.	The mean as a statistical model ①	46
2.4.2.	Assessing the fit of a model: sums of squares and variance revisited ①	46
2.4.3.	Estimating parameters ①	50
2.5.	Going beyond the data ①	51
2.5.1.	The standard error ①	52
2.5.2.	Confidence intervals ②	54
2.6.	Using statistical models to test research questions ①	60
2.6.1.	Null hypothesis significance testing ①	60
2.6.2.	Problems with NHST ②	74
2.7.	Modern approaches to theory testing ②	78
2.7.1.	Effect sizes ②	79
2.7.2.	Meta-analysis ②	83
2.8.	Reporting statistical models ②	84
2.9.	Brian's attempt to woo Jane ①	85
2.10.	What next? ①	86
2.11.	Key terms that I've discovered	87
2.12.	Smart Alex's tasks	87
2.13.	Further reading	88
<b>3</b>	<b>The IBM SPSS Statistics environment</b>	<b>89</b>
3.1.	What will this chapter tell me? ①	89
3.2.	Versions of IBM SPSS Statistics ①	90
3.3.	Windows versus MacOS ①	90
3.4.	Getting started ①	90
3.5.	The data editor ①	91
3.5.1.	Entering data into the data editor ①	98
3.5.2.	The variable view ①	99
3.5.3.	Missing values ①	107
3.6.	Importing data ①	109
3.7.	The SPSS viewer ①	109
3.8.	Exporting SPSS output ①	113
3.9.	The syntax editor ③	113
3.10.	Saving files ①	115
3.11.	Retrieving a file ①	115
3.12.	Brian's attempt to woo Jane ①	116
3.13.	What next? ①	117
3.14.	Key terms that I've discovered	117
3.15.	Smart Alex's tasks	117
3.16.	Further reading	120
<b>4</b>	<b>Exploring data with graphs</b>	<b>121</b>
4.1.	What will this chapter tell me? ①	121
4.2.	The art of presenting data ①	122
4.2.1.	What makes a good graph? ①	122
4.2.2.	Lies, damned lies, and ... erm ... graphs ①	123

4.3.	The SPSS chart builder ①	125
4.4.	Histograms ①	127
4.5.	Boxplots (box-whisker diagrams) ①	131
4.6.	Graphing means: bar charts and error bars ①	135
4.6.1.	Simple bar charts for independent means ①	136
4.6.2.	Clustered bar charts for independent means ①	137
4.6.3.	Simple bar charts for related means ①	140
4.6.4.	Clustered bar charts for related means ①	143
4.6.5.	Clustered bar charts for 'mixed' designs ①	145
4.7.	Line charts ①	148
4.8.	Graphing relationships: the scatterplot ①	148
4.8.1.	Simple scatterplot ①	149
4.8.2.	Grouped scatterplot ①	151
4.8.3.	Simple and grouped 3-D scatterplots ①	153
4.8.4.	Matrix scatterplot ①	154
4.8.5.	Simple dot plot or density plot ①	157
4.8.6.	Drop-line graph ①	157
4.9.	Editing graphs ①	158
4.10.	Brian's attempt to woo Jane ①	161
4.11.	What next? ①	161
4.12.	Key terms that I've discovered	161
4.13.	Smart Alex's tasks	162
4.14.	Further reading	162
<b>5</b>	<b>The beast of bias</b>	<b>163</b>
5.1.	What will this chapter tell me? ①	163
5.2.	What is bias? ①	164
5.2.1.	Assumptions ①	165
5.2.2.	Outliers ①	165
5.2.3.	Additivity and linearity ①	167
5.2.4.	Normally distributed something or other ①	168
5.2.5.	Homoscedasticity/homogeneity of variance ②	172
5.2.6.	Independence ②	176
5.3.	Spotting bias ②	176
5.3.1.	Spotting outliers ②	176
5.3.2.	Spotting normality ①	179
5.3.3.	Spotting linearity and heteroscedasticity/heterogeneity of variance ②	192
5.4.	Reducing bias ②	196
5.4.1.	Trimming the data ②	196
5.4.2.	Winsorizing ①	198
5.4.3.	Robust methods ③	198
5.4.4.	Transforming data ②	201
5.5.	Brian's attempt to woo Jane ①	210
5.6.	What next? ①	210
5.7.	Key terms that I've discovered	211
5.8.	Smart Alex's tasks	211
5.9.	Further reading	212
<b>6</b>	<b>Non-parametric models</b>	<b>213</b>
6.1.	What will this chapter tell me? ①	213
6.2.	When to use non-parametric tests ①	214

6.3.	General procedure of non-parametric tests in SPSS ①	215
6.4.	Comparing two independent conditions: the Wilcoxon rank-sum test and Mann–Whitney test ①	217
6.4.1.	Theory ②	219
6.4.2.	Inputting data and provisional analysis ①	221
6.4.3.	The Mann–Whitney test using SPSS ①	223
6.4.4.	Output from the Mann–Whitney test ①	224
6.4.5.	Calculating an effect size ②	227
6.4.6.	Writing the results ①	227
6.5.	Comparing two related conditions: the Wilcoxon signed-rank test ①	220
6.5.1.	Theory of the Wilcoxon signed-rank test ②	228
6.5.2.	Running the analysis ①	230
6.5.3.	Output for the ecstasy group ①	231
6.5.4.	Output for the alcohol group ①	233
6.5.5.	Calculating an effect size ②	234
6.5.6.	Writing the results ①	234
6.6.	Differences between several independent groups: the Kruskal–Wallis test ①	236
6.6.1.	Theory of the Kruskal–Wallis test ②	236
6.6.2.	Follow-up analysis ②	238
6.6.3.	Inputting data and provisional analysis ①	239
6.6.4.	Doing the Kruskal–Wallis test in SPSS ①	241
6.6.5.	Output from the Kruskal–Wallis test ①	242
6.6.6.	Testing for trends: the Jonckheere–Terpstra test ②	246
6.6.7.	Calculating an effect size ②	248
6.6.8.	Writing and interpreting the results ①	249
6.7.	Differences between several related groups: Friedman's ANOVA ①	249
6.7.1.	Theory of Friedman's ANOVA ②	251
6.7.2.	Inputting data and provisional analysis ①	252
6.7.3.	Doing Friedman's ANOVA in SPSS ①	253
6.7.4.	Output from Friedman's ANOVA ①	254
6.7.5.	Following-up Friedman's ANOVA ②	256
6.7.6.	Calculating an effect size ②	256
6.7.7.	Writing and interpreting the results ①	257
6.8.	Brian's attempt to woo Jane ①	258
6.9.	What next? ①	259
6.10.	Key terms that I've discovered	259
6.11.	Smart Alex's tasks	259
6.12.	Further reading	261
<b>7</b>	<b>Correlation</b>	<b>262</b>
7.1.	What will this chapter tell me? ①	262
7.2.	Modelling relationships ①	263
7.2.1.	A detour into the murky world of covariance ①	264
7.2.2.	Standardization and the correlation coefficient ①	266
7.2.3.	The significance of the correlation coefficient ③	268
7.2.4.	Confidence intervals for $r$ ③	269
7.2.5.	A word of warning about interpretation: causality ①	270
7.3.	Data entry for correlation analysis using SPSS ①	270
7.4.	Bivariate correlation ①	271
7.4.1.	General procedure for running correlations in SPSS ①	272
7.4.2.	Pearson's correlation coefficient ①	274
7.4.3.	Spearman's correlation coefficient ①	276



7.4.4. Kendall's tau (non-parametric) ①	277
7.4.5. Biserial and point-biserial correlations ③	279
7.5. Partial correlation ②	280
7.5.1. The theory behind part and partial correlation ③	280
7.5.2. Partial correlation in SPSS ③	283
7.5.3. Semi-partial (or part) correlations ②	285
7.6. Comparing correlations ③	285
7.6.1. Comparing independent $r_s$ ③	285
7.6.2. Comparing dependent $r_s$ ③	286
7.7. Calculating the effect size ①	287
7.8. How to report correlation coefficients ①	288
7.9. Brian's attempt to woo Jane ①	290
7.10. What next? ①	290
7.11. Key terms that I've discovered	291
7.12. Smart Alex's tasks	291
7.13. Further reading	292
<b>8 Regression</b>	<b>293</b>
8.1. What will this chapter tell me? ①	293
8.2. An introduction to regression ①	294
8.2.1. The simple linear model ①	294
8.2.2. The linear model with several predictors ②	296
8.2.3. Estimating the model ②	298
8.2.4. Assessing the goodness of fit, sums of squares, $R$ and $R^2$ ①	300
8.2.5. Assessing individual predictors ①	303
8.3. Bias in regression models? ②	304
8.3.1. Is the model biased by unusual cases? ②	304
8.3.2. Generalizing the model ②	309
8.3.3. Sample size in regression ③	313
8.4. Regression using SPSS: One Predictor ①	314
8.4.1. Regression: the general procedure ①	315
8.4.2. Running a simple regression using SPSS ①	316
8.4.3. Interpreting a simple regression ①	318
8.4.4. Using the model ①	320
8.5. Multiple regression ②	321
8.5.1. Methods of regression ②	321
8.5.2. Comparing models ②	324
8.5.3. Multicollinearity ②	324
8.6. Regression with several predictors using SPSS ②	326
8.6.1. Main options ②	327
8.6.2. Statistics ②	328
8.6.3. Regression plots ②	329
8.6.4. Saving regression diagnostics ②	331
8.6.5. Further options ②	332
8.6.6. Robust regression ②	333
8.7. Interpreting multiple regression ②	334
8.7.1. Descriptives ②	334
8.7.2. Summary of model ②	335
8.7.3. Model parameters ②	338
8.7.4. Excluded variables ②	342
8.7.5. Assessing multicollinearity ②	342
8.7.6. Bias in the model: casewise diagnostics ②	345
8.7.7. Bias in the model: assumptions ②	348

8.8.	What if I violate an assumption? Robust regression ②	349
8.9.	How to report multiple regression ②	351
8.10.	Brian's attempt to woo Jane ①	353
8.11.	What next? ①	354
8.12.	Key terms that I've discovered	354
8.13.	Smart Alex's tasks	354
8.14.	Further reading	355
<b>9</b>	<b>Comparing two means</b>	<b>357</b>
9.1.	What will this chapter tell me? ①	357
9.2.	Looking at differences ①	358
9.2.1.	An example: are invisible people mischievous? ①	359
9.2.2.	Categorical predictors in the linear model ①	360
9.3.	The <i>t</i> -test ①	364
9.3.1.	Rationale for the <i>t</i> -test ①	364
9.3.2.	The independent <i>t</i> -test equation explained ①	365
9.3.3.	The matched-samples <i>t</i> -test equation explained ①	368
9.4.	Assumptions of the <i>t</i> -test ①	371
9.5.	The independent <i>t</i> -test using SPSS ①	371
9.5.1.	The general procedure ①	371
9.5.2.	Exploring data and testing assumptions ①	372
9.5.3.	Compute the independent <i>t</i> -test ①	372
9.5.4.	Output from the independent <i>t</i> -test ①	373
9.5.5.	Calculating the effect size ②	376
9.5.6.	Reporting the independent <i>t</i> -test ①	377
9.6.	Matched-samples <i>t</i> -test using SPSS ①	378
9.6.1.	Entering data ①	378
9.6.2.	Exploring data and testing assumptions ①	378
9.6.3.	Computing the paired-samples <i>t</i> -test ①	383
9.6.4.	Calculating the effect size ①	386
9.6.5.	Reporting the paired-samples <i>t</i> -test ①	388
9.7.	Between groups or repeated measures? ①	388
9.8.	What if I violate the test assumptions? ②	388
9.9.	Brian's attempt to woo Jane ①	389
9.10.	What next? ①	389
9.11.	Key terms that I've discovered	389
9.12.	Smart Alex's tasks	390
9.13.	Further reading	391
<b>10</b>	<b>Moderation, mediation and more regression</b>	<b>392</b>
10.1.	What will this chapter tell me? ①	392
10.2.	Installing custom dialog boxes in SPSS ②	393
10.3.	Moderation: interactions in regression ③	395
10.3.1.	The conceptual model ③	395
10.3.2.	The statistical model ②	397
10.3.3.	Centring variables ②	398
10.3.4.	Creating interaction variables ②	400
10.3.5.	Following up an interaction effect ②	400
10.3.6.	Running the analysis ②	401
10.3.7.	Output from moderation analysis ②	402
10.3.8.	Reporting moderation analysis ②	407
10.4.	Mediation ②	408

10.4.1. The conceptual model ②	408
10.4.2. The statistical model ②	409
10.4.3. Effect sizes of mediation ③	411
10.4.4. Running the analysis ②	413
10.4.5. Output from mediation analysis ②	414
10.4.6. Reporting mediation analysis ②	418
10.5. Categorical predictors in regression ③	419
10.5.1. Dummy coding ③	419
10.5.2. SPSS output for dummy variables ③	422
10.6. Brian's attempt to woo Jane ①	426
10.7. What next? ①	427
10.8. Key terms that I've discovered	427
10.9. Smart Alex's tasks	427
10.10. Further reading	428
<b>11 Comparing several means: ANOVA (GLM 1)</b>	<b>429</b>
11.1. What will this chapter tell me? ①	429
11.2. The theory behind ANOVA ②	430
11.2.1. Using a linear model to compare means ②	430
11.2.2. Logic of the $F$ -ratio ②	434
11.2.3. Total sum of squares ( $SS_T$ ) ②	436
11.2.4. Model sum of squares ( $SS_M$ ) ②	438
11.2.5. Residual sum of squares ( $SS_R$ ) ②	439
11.2.6. Mean squares ②	440
11.2.7. The $F$ -ratio ②	441
11.2.8. Interpreting $F$ ②	442
11.3. Assumptions of ANOVA ③	442
11.3.1. Homogeneity of variance ②	442
11.3.2. Is ANOVA robust? ③	443
11.3.3. What to do when assumptions are violated ②	445
11.4. Planned contrasts ②	445
11.4.1. Choosing which contrasts to do ②	446
11.4.2. Defining contrasts using weights ②	449
11.4.3. Non-orthogonal comparisons ②	454
11.4.4. Standard contrasts ②	456
11.4.5. Polynomial contrasts: trend analysis ②	457
11.5. <i>Post hoc</i> procedures ②	458
11.5.1. Type I and Type II error rates for <i>post hoc</i> tests ②	458
11.5.2. Are <i>post hoc</i> procedures robust? ②	459
11.5.3. Summary of <i>post hoc</i> procedures ②	459
11.6. Running one-way ANOVA in SPSS ②	460
11.6.1. General procedure of one-way ANOVA ②	460
11.6.2. Planned comparisons using SPSS ②	462
11.6.3. <i>Post hoc</i> tests in SPSS ②	463
11.6.4. Options ②	464
11.6.5. Bootstrapping ②	465
11.7. Output from one-way ANOVA ②	466
11.7.1. Output for the main analysis ②	466
11.7.2. Output for planned comparisons ②	469
11.7.3. Output for <i>post hoc</i> tests ②	470
11.8. Calculating the effect size ②	472
11.9. Reporting results from one-way independent ANOVA ②	474
11.10. Key terms that I've discovered	475

11.11. Brian's attempt to woo Jane ①	475
11.12. What next? ①	476
11.13. Smart Alex's tasks	476
11.14. Further reading	477
<b>12 Analysis of covariance, ANCOVA (GLM 2)</b>	<b>478</b>
12.1. What will this chapter tell me? ②	478
12.2. What is ANCOVA? ②	479
12.3. Assumptions and issues in ANCOVA ③	484
12.3.1. Independence of the covariate and treatment effect ③	484
12.3.2. Homogeneity of regression slopes ③	485
12.3.3. What to do when assumptions are violated ②	488
12.4. Conducting ANCOVA in SPSS ②	488
12.4.1. General procedure ①	488
12.4.2. Inputting data ①	488
12.4.3. Testing the independence of the treatment variable and covariate ②	488
12.4.4. The main analysis ②	489
12.4.5. Contrasts	490
12.4.6. Other options ②	491
12.4.7. Bootstrapping and plots ②	493
12.5. Interpreting the output from ANCOVA ②	490
12.5.1. What happens when the covariate is excluded? ②	493
12.5.2. The main analysis ②	494
12.5.3. Contrasts ②	497
12.5.4. Interpreting the covariate ②	497
12.6. Testing the assumption of homogeneity of regression slopes ③	499
12.7. Calculating the effect size ②	500
12.8. Reporting results ②	503
12.9. Brian's attempt to woo Jane ①	504
12.10. What next? ②	504
12.11. Key terms that I've discovered	505
12.12. Smart Alex's tasks	505
12.13. Further reading	506
<b>13 Factorial ANOVA (GLM 3)</b>	<b>507</b>
13.1. What will this chapter tell me? ②	507
13.2. Theory of factorial ANOVA (independent designs) ②	508
13.2.1. Factorial designs ②	508
13.2.2. Guess what? Factorial ANOVA is a linear model ③	509
13.2.3. Two-way ANOVA: behind the scenes ②	514
13.2.4. Total sums of squares ( $SS_T$ ) ②	515
13.2.5. Model sum of squares, $SS_M$ ②	516
13.2.6. The residual sum of squares, $SS_R$ ②	519
13.2.7. The $F$ -ratios ②	519
13.3. Assumptions of factorial ANOVA ③	520
13.4. Factorial ANOVA using SPSS ②	520
13.4.1. General procedure for factorial ANOVA ①	520
13.4.2. Entering the data and accessing the main dialog box ②	521
13.4.3. Graphing interactions ②	522
13.4.4. Contrasts ②	523
13.4.5. <i>Post hoc</i> tests ②	524
13.4.6. Bootstrapping and other options ②	524

13.5. Output from factorial ANOVA ②	526
13.5.1. Levene's test ②	526
13.5.2. The main ANOVA table ②	526
13.5.3. Contrasts ②	529
13.5.4. Simple effects analysis ③	530
13.5.5. <i>Post hoc</i> analysis ②	532
13.6. Interpreting interaction graphs ②	533
13.7. Calculating effect sizes ③	537
13.8. Reporting the results of two-way ANOVA ②	539
13.9. Brian's attempt to woo Jane ①	540
13.10. What next? ②	541
13.11. Key terms that I've discovered	541
13.12. Smart Alex's tasks	541
13.13. Further reading	542
<b>14 Repeated-measures designs (GLM 4)</b>	<b>543</b>
14.1. What will this chapter tell me? ②	543
14.2. Introduction to repeated-measures designs ②	544
14.2.1. The assumption of sphericity ②	545
14.2.2. How is sphericity measured? ②	545
14.2.3. Assessing the severity of departures from sphericity ②	546
14.2.4. What is the effect of violating the assumption of sphericity? ③	546
14.2.5. What do you do if you violate sphericity? ②	548
14.3. Theory of one-way repeated-measures ANOVA ②	548
14.3.1. The total sum of squares, $SS_T$ ②	551
14.3.2. The within-participant sum of squares, $SS_W$ ②	551
14.3.3. The model sum of squares, $SS_M$ ②	552
14.3.4. The residual sum of squares, $SS_R$ ②	552
14.3.5. The mean squares ②	552
14.3.6. The <i>F</i> -ratio ②	554
14.3.7. The between-participants sum of squares ②	554
14.4. Assumptions in repeated-measures ANOVA ③	555
14.5. One-way repeated-measures ANOVA using SPSS ②	555
14.5.1. Repeated-measures ANOVA: the general procedure ②	555
14.5.2. The main analysis ②	555
14.5.3. Defining contrasts for repeated measures ②	557
14.5.4. <i>Post hoc</i> tests and additional options ③	558
14.6. Output for one-way repeated-measures ANOVA ②	559
14.6.1. Descriptives and other diagnostics ①	559
14.6.2. Assessing and correcting for sphericity: Mauchly's test ②	560
14.6.3. The main ANOVA ②	560
14.6.4. Contrasts ②	563
14.6.5. <i>Post hoc</i> tests ②	565
14.7. Effect sizes for repeated-measures ANOVA ③	566
14.8. Reporting one-way repeated-measures ANOVA ②	568
14.9. Factorial repeated-measures designs ②	568
14.9.1. The main analysis ②	570
14.9.2. Contrasts ②	573
14.9.3. Simple effects analysis ③	573
14.9.4. Graphing interactions ②	574
14.9.5. Other options ②	574
14.10. Output for factorial repeated-measures ANOVA ②	576

14.10.1. Descriptives and main analysis ②	576
14.10.2. Contrasts for repeated-measures variables ②	581
14.11. Effect sizes for factorial repeated-measures ANOVA ③	586
14.12. Reporting the results from factorial repeated-measures ANOVA ②	587
14.13. Brian's attempt to woo Jane ①	588
14.14. What next? ②	589
14.15. Key terms that I've discovered	589
14.16. Smart Alex's tasks	589
14.17. Further reading	590
<b>15 Mixed design ANOVA (GLM 5)</b>	<b>591</b>
15.1. What will this chapter tell me? ①	591
15.2. Mixed designs ②	592
15.3. Assumptions in mixed designs ②	593
15.4. What do men and women look for in a partner? ②	593
15.5. Mixed ANOVA in SPSS ②	594
15.5.1. Mixed ANOVA: the general procedure ②	594
15.5.2. Entering data ②	594
15.5.3. The main analysis ②	595
15.5.4. Other options ②	598
15.6. Output for mixed factorial ANOVA ③	600
15.6.1. The main effect of gender ②	602
15.6.2. The main effect of looks ②	603
15.6.3. The main effect of charisma ②	605
15.6.4. The interaction between gender and looks ②	606
15.6.5. The interaction between gender and charisma ②	607
15.6.6. The interaction between attractiveness and charisma ②	608
15.6.7. The interaction between looks, charisma and gender ③	611
15.6.8. Conclusions ③	614
15.7. Calculating effect sizes ③	615
15.8. Reporting the results of mixed ANOVA ②	617
15.9. Brian's attempt to woo Jane ①	620
15.10. What next? ②	621
15.11. Key terms that I've discovered	621
15.12. Smart Alex's tasks	621
15.13. Further reading	622
<b>16 Multivariate analysis of variance (MANOVA)</b>	<b>623</b>
16.1. What will this chapter tell me? ②	623
16.2. When to use MANOVA ②	624
16.3. Introduction	624
16.3.1. Similarities to and differences from ANOVA ②	624
16.3.2. Choosing outcomes ②	625
16.3.3. The example for this chapter ②	626
16.4. Theory of MANOVA ③	626
16.4.1. Introduction to matrices ③	626
16.4.2. Some important matrices and their functions ③	628
16.4.3. Calculating MANOVA by hand: a worked example ③	629
16.4.4. Principle of the MANOVA test statistic ④	637
16.5. Practical issues when conducting MANOVA ③	642
16.5.1. Assumptions and how to check them ③	642

16.5.2. What to do when assumptions are violated ③	643
16.5.3. Choosing a test statistic ③	643
16.5.4. Follow-up analysis ③	644
16.6. MANOVA using SPSS ②	644
16.6.1. General procedure of one-way ANOVA ②	644
16.6.2. The main analysis ②	645
16.6.3. Multiple comparisons in MANOVA ②	646
16.6.4. Additional options ③	646
16.7. Output from MANOVA ③	647
16.7.1. Preliminary analysis and testing assumptions ③	647
16.7.2. MANOVA test statistics ③	648
16.7.3. Univariate test statistics ②	649
16.7.4. SSCP matrices ③	650
16.7.5. Contrasts ③	652
16.8. Reporting results from MANOVA ②	652
16.9. Following up MANOVA with discriminant analysis ③	654
16.10. Output from the discriminant analysis ④	656
16.11. Reporting results from discriminant analysis ②	660
16.12. The final interpretation ④	660
16.13. Brian's attempt to woo Jane ①	662
16.14. What next? ②	663
16.15. Key terms that I've discovered	663
16.16. Smart Alex's tasks	664
16.17. Further reading	664
<b>17 Exploratory factor analysis</b>	<b>665</b>
17.1. What will this chapter tell me? ①	665
17.2. When to use factor analysis ②	666
17.3. Factors and components ②	667
17.3.1. Graphical representation ②	668
17.3.2. Mathematical representation ②	669
17.3.3. Factor scores ②	671
17.4. Discovering factors ②	674
17.4.1. Choosing a method ②	674
17.4.2. Community ②	675
17.4.3. Factor analysis or PCA? ②	675
17.4.4. Theory behind PCA ③	676
17.4.5. Factor extraction: eigenvalues and the scree plot ②	677
17.4.6. Improving interpretation: factor rotation ③	678
17.5. Research example ②	682
17.5.1. General procedure ①	682
17.5.2. Before you begin ②	683
17.6. Running the analysis ②	686
17.6.1. Factor extraction in SPSS ②	688
17.6.2. Rotation ②	689
17.6.3. Scores ②	691
17.6.4. Options ②	691
17.7. Interpreting output from SPSS ②	692
17.7.1. Preliminary analysis ②	693
17.7.2. Factor extraction ②	696
17.7.3. Factor rotation ②	701
17.7.4. Factor scores ②	703
17.7.5. Summary ②	705

17.8. How to report factor analysis ①	706
17.9. Reliability analysis ②	706
17.9.1. Measures of reliability ③	706
17.9.2. Interpreting Cronbach's $\alpha$ (some cautionary tales) ②	709
17.9.3. Reliability analysis in SPSS ②	710
17.9.4. Reliability analysis output ②	712
17.10. How to report reliability analysis ②	716
17.11. Brian's attempt to woo Jane ①	716
17.12. What next? ②	717
17.13. Key terms that I've discovered	717
17.14. Smart Alex's tasks	717
17.15. Further reading	719
<b>18 Categorical data</b>	<b>720</b>
18.1. What will this chapter tell me? ①	720
18.2. Analysing categorical data ①	721
18.3. Theory of analysing categorical data ①	721
18.3.1. Pearson's chi-square test ①	721
18.3.2. Fisher's exact test ①	723
18.3.3. The likelihood ratio ②	724
18.3.4. Yates's correction ②	724
18.3.5. Other measures of association ①	725
18.3.6. Several categorical variables: loglinear analysis ③	725
18.4. Assumptions when analysing categorical data ①	735
18.4.1. Independence ①	735
18.4.2. Expected frequencies ①	735
18.4.3. More doom and gloom ①	736
18.5. Doing chi-square in SPSS ①	736
18.5.1. General procedure for analysing categorical outcomes ①	736
18.5.2. Entering data ①	736
18.5.3. Running the analysis ①	738
18.5.4. Output for the chi-square test ①	740
18.5.5. Breaking down a significant chi-square test with standardized residuals ②	743
18.5.6. Calculating an effect size ②	744
18.5.7. Reporting the results of chi-square ①	746
18.6. Loglinear analysis using SPSS ②	746
18.6.1. Initial considerations ②	746
18.6.2. Running loglinear analysis ②	748
18.6.3. Output from loglinear analysis ③	750
18.6.4. Following up loglinear analysis ②	753
18.7. Effect sizes in loglinear analysis ②	755
18.8. Reporting the results of loglinear analysis ②	756
18.9. Brian's attempt to woo Jane ①	757
18.10. What next? ①	757
18.11. Key terms that I've discovered	758
18.12. Smart Alex's tasks	758
18.13. Further reading	759
<b>19 Logistic regression</b>	<b>760</b>
19.1. What will this chapter tell me? ①	760
19.2. Background to logistic regression ①	761



19.3. What are the principles behind logistic regression? ③	762
19.3.1. Assessing the model: the log-likelihood statistic ③	763
19.3.2. Assessing the model: the deviance statistic ③	763
19.3.3. Assessing the model: $R$ and $R^2$ ③	764
19.3.4. Assessing the contribution of predictors: the Wald statistic ②	766
19.3.5. The odds ratio: $\exp(B)$ ③	766
19.3.6. Model building and parsimony ②	767
19.4. Sources of bias and common problems ④	768
19.4.1. Assumptions ②	768
19.4.2. Incomplete information from the predictors ④	769
19.4.3. Complete separation ④	770
19.4.4. Overdispersion ④	772
19.5. Binary logistic regression: an example that will make you feel eel ②	773
19.5.1. Building a model ①	774
19.5.2. Logistic regression: the general procedure ①	775
19.5.3. Data entry ①	775
19.5.4. Building the models in SPSS ②	775
19.5.5. Method of regression ②	776
19.5.6. Categorical predictors ②	776
19.5.7. Comparing the models ②	778
19.5.8. Rerunning the model ①	780
19.5.9. Obtaining residuals ②	781
19.5.10. Further options ②	781
19.5.11. Bootstrapping ②	782
19.6. Interpreting logistic regression ②	783
19.6.1. Block 0 ②	783
19.6.2. Model summary ②	783
19.6.3. Listing predicted probabilities ②	789
19.6.4. Interpreting residuals ②	789
19.6.5. Calculating the effect size ②	792
19.7. How to report logistic regression ②	792
19.8. Testing assumptions: another example ②	792
19.8.1. Testing for linearity of the logit ③	794
19.8.2. Testing for multicollinearity ③	794
19.9. Predicting several categories: multinomial logistic regression ③	797
19.9.1. Running multinomial logistic regression in SPSS ③	799
19.9.2. Statistics ③	802
19.9.3. Other options ③	803
19.9.4. Interpreting the multinomial logistic regression output ②	804
19.9.5. Reporting the results ②	811
19.10. Brian's attempt to woo Jane ①	811
19.11. What next? ①	811
19.12. Key terms that I've discovered	812
19.13. Smart Alex's tasks	812
19.14. Further reading	813
<b>20 Multilevel linear models</b>	<b>814</b>
20.1. What will this chapter tell me? ①	814
20.2. Hierarchical data ②	815
20.2.1. The intraclass correlation ②	816
20.2.2. Benefits of multilevel models ②	818

20.3	Theory of multilevel linear models ③	819
20.3.1.	An example ②	819
20.3.2.	Fixed and random coefficients ③	820
20.4	The multilevel model ④	823
20.4.1.	Assessing the fit and comparing multilevel models ④	825
20.4.2.	Types of covariance structures ④	826
20.5	Some practical issues ③	827
20.5.1.	Assumptions ③	827
20.5.2.	Robust multilevel models ③	828
20.5.3.	Sample size and power ③	829
20.5.4.	Centring predictors ③	829
20.6	Multilevel modelling using SPSS ④	830
20.6.1.	Entering the data ②	831
20.6.2.	Ignoring the data structure: ANOVA ②	831
20.6.3.	Ignoring the data structure: ANCOVA ②	836
20.6.4.	Factoring in the data structure: random intercepts ③	838
20.6.5.	Factoring in the data structure: random intercepts and slopes ④	841
20.6.6.	Adding an interaction to the model ④	844
20.7.	Growth models ④	849
20.7.1.	Growth curves (polynomials) ④	850
20.7.2.	An example: the honeymoon period ②	851
20.7.3.	Restructuring the data ③	853
20.7.4.	Running a growth model on SPSS ④	854
20.7.5.	Further analysis ④	860
20.8.	How to report a multilevel model ③	862
20.9.	A message from the octopus of inescapable despair ①	863
20.10.	Brian's attempt to woo Jane ①	864
20.11.	What next? ②	864
20.12.	Key terms that I've discovered	865
20.13.	Smart Alex's tasks	865
20.14.	Further reading	866
<b>21</b>	<b>Epilogue: life after discovering statistics</b>	<b>867</b>
21.1.	Nice emails	867
21.2.	Everybody thinks that I'm a statistician	868
21.3.	Craziness on a grand scale	868
21.3.1.	Statistics	868
21.3.2.	Cult of underlying numerical truths	869
21.3.3.	And then it got really weird	869
	Glossary	870
	Appendix	887
	References	899
	Index	908

# PREFACE

*Karma Police, arrest this man, he talks in maths, he buzzes like a fridge, he's like a detuned radio*

Radiohead, 'Karma Police', *OK Computer* (1997)

## Introduction

Many behavioural and social science students (and researchers for that matter) despise statistics. Most of us have a non-mathematical background, which makes understanding complex statistical equations very difficult. Nevertheless, the evil goat-warriors of Satan force our non-mathematical brains to apply themselves to what is the very complex task of becoming a statistics expert. The end result, as you might expect, can be quite messy. The one weapon that we have is the computer, which allows us to neatly circumvent the considerable disability of not understanding mathematics. Computer programs such as IBM SPSS Statistics, SAS, R and the like provide an opportunity to teach statistics at a conceptual level without getting too bogged down in equations. The computer to a goat-warrior of Satan is like catnip to a cat: it makes them rub their heads along the ground and purr and dribble ceaselessly. The only downside of the computer is that it makes it really easy to make a complete idiot of yourself if you don't really understand what you're doing. Using a computer without any statistical knowledge at all can be a dangerous thing. Hence this book.

My first aim is to strike a good balance between theory and practice: I want to use the computer as a tool for teaching statistical concepts in the hope that you will gain a better understanding of both theory and practice. If you want theory and you like equations then there are certainly better books: Howell (2012), Stevens (2002) and Tabachnick and Fidell (2012) have taught (and continue to teach) me more about statistics than you could possibly imagine. (I have an ambition to be cited in one of these books, but I don't think that will ever happen.) However, if you want a stats book that also discusses digital rectal stimulation then you have just spent your money wisely.

Too many books create the impression that there is a 'right' and 'wrong' way to do statistics. Data analysis is more subjective than is often made out. Therefore, although I make recommendations, within the limits imposed by the senseless destruction of rainforests, I hope to give you enough background in theory to enable you to make your own decisions about how best to conduct your analysis.

A second (ridiculously ambitious) aim is to make this the only statistics book that you'll ever need to buy. It's a book that I hope will become your friend from first year at university right through to your professorship. The start of the book is aimed at first-year undergraduates (Chapters 1–9), and then we move onto second-year undergraduate level material (Chapters 5, 8 and 10–15) before a dramatic climax that should keep postgraduates tickled

(Chapters 16–20). There should be something for everyone in each chapter also, and to help you gauge the difficulty of material, I flag the level of each section within each chapter (more on that in a moment).

My final and most important aim is to make the learning process fun. I have a sticky history with maths. This extract is from my school report at the age of 11:

MATHEMATICS ADDL. MATHS.	43	59	27	D	C	His work shows lack of discipline in thought and presentation. I don't hope it will matter next year.
CHEMISTRY						

The '27' in the report is to say that I came equal 27th with another student out of a class of 29. That's pretty much bottom of the class. The 43 is my exam mark as a percentage. Oh dear. Four years later (at 15) this was my school report:

NAME Andrew Field..... FORM 4Q SUBJECT Mathematics..

Andrew's progress in Mathematics has been remarkable. From being a weaker candidate who lacked confidence he has developed into a budding Mathematician. He should achieve a good grade.

EXAM	
ATTAINMENT	
EFFORT	

Date 27/6/88

B.A. Greate..... Subject Teacher

The catalyst of this remarkable change was having a good teacher: my brother, Paul. I owe my life as an academic to Paul's ability to teach me stuff in an engaging way – something my maths teachers failed to do. Paul's a great teacher because he cares about bringing out the best in people, and he was able to make things interesting and relevant to me. He got the 'good teaching' genes in the family, but wasted them by not becoming a teacher; however, they're a little less wasted because his approach inspires mine. I strongly believe that people appreciate the human touch, and so I try to inject a lot of my own personality and sense of humour (or lack of) into *Discovering Statistics Using ...* books. Many of the examples in this book, although inspired by some of the craziness that you find in the real world, are designed to reflect topics that play on the minds of the average student (i.e., sex, drugs, rock and roll, celebrity, people doing crazy stuff). There are also some examples that are there simply because they made me laugh. So, the examples are light-hearted (some have said 'smutty', but I prefer 'light-hearted') and by the end, for better or worse, I think you will have some idea of what goes on in my head on a daily basis. I apologize to those who think it's crass, hate it, or think that I'm undermining the seriousness of science, but, come on, what's not funny about a man putting an eel up his anus?

I never believe that I meet my aims, but previous editions have certainly been popular. I enjoy the rare luxury of having complete strangers emailing me to tell me how wonderful I am. (Admittedly, there are also emails calling me a pile of gibbon excrement, but you have to take the rough with the smooth.) The second edition of this book also won the British Psychological Society book award in 2007. However, with every new edition, I fear that the changes I make will ruin all of my previous hard work. Let's see what those changes are.

## What do you get for your money?

This book takes you on a journey (and I try my best to make it a pleasant one) not just of statistics but also of the weird and wonderful contents of the world and my brain. It's full of stupid examples, bad jokes, smut and filth. Aside from the smut, I have been forced reluctantly to include some academic content. In essence it contains everything I know about statistics (actually, more than I know ...). It also has these features:

- **Everything you'll ever need to know:** I want this book to be good value for money, so it guides you from complete ignorance (Chapter 1 tells you the basics of doing research) to being an expert on multilevel modelling (Chapter 20). Of course no book that it's physically possible to lift will contain everything, but I think this one has a fair crack. It's pretty good for developing your biceps also.
- **Stupid faces:** You'll notice that the book is riddled with stupid faces, some of them my own. You can find out more about the pedagogic function of these 'characters' in the next section, but even without any useful function they're nice to look at.
- **Data sets:** There are about 132 data files associated with this book on the companion website. Not unusual in itself for a statistics book, but my data sets contain more sperm (not literally) than other books. I'll let you judge for yourself whether this is a good thing.
- **My life story:** Each chapter is book-ended by a chronological story from my life. Does this help you to learn about statistics? Probably not, but hopefully it provides some light relief between chapters.
- **SPSS tips:** SPSS does weird things sometimes. In each chapter, there are boxes containing tips, hints and pitfalls related to SPSS.
- **Self-test questions:** Given how much students hate tests, I thought the best way to commit commercial suicide was to liberally scatter tests throughout each chapter. These range from simple questions to test what you have just learned to going back to a technique that you read about several chapters before and applying it in a new context. All of these questions have answers to them on the companion website so that you can check on your progress.
- **Companion website:** The companion website contains an absolutely insane amount of additional material, all of which is described in the section about the companion website.
- **Digital stimulation:** No, not the aforementioned type of digital stimulation, but brain stimulation. Many of the features on the companion website will be accessible from tablets and smartphones, so that when you're bored in the cinema you can read about the fascinating world of heteroscedasticity instead.
- **Reporting your analysis:** Every chapter has a guide to writing up your analysis. How you write up an analysis varies a bit from one discipline to another, but my guides should get you heading in the right direction.
- **Glossary:** Writing the glossary was so horribly painful that it made me stick a vacuum cleaner into my ear to suck out my own brain. You can find my brain in the bottom of the vacuum cleaner in my house.
- **Real-world data:** Students like to have 'real data' to play with. The trouble is that real research can be quite boring. However, just for you, I trawled the world for examples of research on really fascinating topics (in my opinion). I then stalked the authors of the research until they gave me their data. Every chapter has a real research example.

## What do you get that you didn't get last time?

My publishers are obsessed with me telling you how this book differs from the last edition. I suppose if you have spent your hard-earned money on the previous edition it's reasonable that you want a good reason to spend more money on this edition. In some respects it's hard to quantify all of the changes in a list: I'm a better writer than I was 4 years ago, so there is a lot of me rewriting things because I think I can do it better than before. I spent 6 months solidly on the updates, so suffice it to say that a lot has changed; but anything you might have liked about the previous edition probably hasn't changed:

- **IBM SPSS compliance:** This edition was written using versions 20 and 21 of IBM SPSS Statistics. IBM bring out a new SPSS each year and this book gets rewritten about every 4 years, so, depending on when you buy the book, it may not reflect the latest version. This shouldn't bother you because one edition of SPSS is much the same as another (see Section 3.2).
- **New! Mediation and Moderation:** Even since the first edition I have been meaning to do a chapter on mediation and moderation, because they are two very widely used techniques. With each new edition I have run out of energy. Not this time though: I wrote it in the middle of the update before I managed to completely burn myself out. Chapter 10 is brand spanking new and all about mediation and moderation.
- **New! Structure:** My publishers soiled their underwear at the thought of me changing the structure because they think lecturers who use the book don't like this sort of change. They might have a point, but I changed it anyway. So, logistic regression (a complex topic) has moved towards the end of the book, and non-parametric tests (a relatively straightforward topic) have moved towards the beginning. In my opinion this change enables the book's story to flow better.
- **New! Focus:** Statistical times are a-changing, and people are starting to appreciate the limitations of significance testing, so I have discussed this more in Chapter 2, and the points made there permeate the rest of the book. The theme of 'everything being the same model' has run through all editions of the book, but I have made this theme even more explicit this time.
- **New! Tasks:** There are 111 more Smart Alex tasks, and 8 more Labcoat Leni tasks. This, of course, means there are quite a lot more pages of answers to these tasks on the companion website.
- **New! Bootstrapping:** The SPSS bootstrapping procedure is covered in every chapter where it is relevant.
- **New! Process diagrams:** Every chapter has a diagrammatic summary of the key steps that you go through for a particular analysis.
- **New! Love story:** Every chapter has a diagrammatic summary at the end (*Brian's attempt to woo Jane*). More interesting, though, Brian Haemorrhage has fallen in love with Jane Superbrain (see next section) and these diagrams follow Brian's attempts to convince Jane to go on a date with him.
- **New! Characters:** I enjoy coming up with new characters, and this edition has a crazy hippy called Odit, and a deranged philosopher called Confusius (see the next section).
- **New-ish! Assumptions:** I've never really liked the way I dealt with assumptions, so I completely rewrote Chapter 5 to try to give more of a sense of when assumptions actually matter.

Every chapter had a serious edit/rewrite, but here is a chapter-by-chapter run-down of the more substantial changes:

- **Chapter 1 (Doing research):** I added some more material on reporting data. I added stuff about variance and standard deviations, and expanded the discussion of  $p$ -values.
- **Chapter 2 (Statistical theory):** I added material on estimating parameters, significance testing and its limitations, problems with one-tailed tests, running multiple tests (i.e., familywise error), confidence intervals and significance, sample size and significance, effect sizes (including Cohen's  $d$  and meta-analysis), and reporting basic statistics. It's changed a lot.
- **Chapter 3 (IBM SPSS):** No dramatic changes.
- **Chapter 4 (Graphs):** I moved the discussion of outliers into Chapter 5, which meant I had to rewrite one of the examples. I now include population pyramids also.
- **Chapter 5 (Assumptions):** I completely rewrote this chapter. It's still about assumptions, but I try to explain when they matter and what they bias. Rather than dealing with assumptions separately in every chapter, because everything in the book is a linear model, I deal with the assumptions of linear models here. Therefore, this chapter acts as a single reference point for all subsequent chapters. I also cover other sources of bias such as outliers (which used to be scattered about in different chapters).
- **Chapter 6 (Non-parametric models):** This is a fully updated and rewritten chapter on non-parametric statistics. It used to be later in the book, but now flows gracefully on from the discussion of assumptions.
- **Chapter 7 (Correlation):** No dramatic changes.
- **Chapter 8 (Regression):** I restructured this chapter so that most of the theory is now at the beginning and most of the SPSS is at the end. I did a fair bit of editing, too, moved categorical predictors into Chapter 10, and integrated simple and multiple regression more.
- **Chapter 9 ( $t$ -tests):** The old version of this chapter used spider examples, but someone emailed me to say that this freaked them out, so I changed the example to be about cloaks of invisibility. Hopefully that won't freak anyone out. I restructured a bit, too, so that the theory is in one place and the SPSS in another.
- **Chapter 10 (Mediation and moderation):** This chapter is completely new.
- **Chapter 11 (GLM 1):** I gave more prominence to ANOVA as a general linear model because this makes it easier to think about assumptions and bias. I moved some of the more technical bits of the SPSS interpretation into boxes so that you can ignore them if you wish.
- **Chapter 12 (GLM 2):** Again some restructuring and a bit more discussion on whether the covariate and predictor need to be independent.
- **Chapters 13–15 (GLM 3–5):** These haven't changed much. I restructured each one a bit, edited down/rewrote a lot and gave more prominence to the GLM way of thinking.
- **Chapter 16 (MANOVA):** I gave the writing a bit of a polish, but no real content changes.
- **Chapter 17 (Factor analysis):** I added some stuff to the theory to make the distinction between principal component analysis (PCA) and factor analysis (FA) clearer. The chapter used to focus on PCA, but I changed it so that the focus is on FA. I edited out 3000 words of my tedious, repetitive, superfluous drivel.

- **Chapters 18 and 19 (Categorical data and logistic regression):** Because these chapters both deal with categorical outcomes, I rewrote them and put them together. The basic content is the same as before.
- **Chapter 20 (Multilevel models):** I polished the writing a bit and updated, but there are no changes that will upset anyone.

## Goodbye

The first edition of this book was the result of two years (give or take a few weeks to write up my Ph.D.) of trying to write a statistics book that I would enjoy reading. With each new edition I try not just to make superficial changes but also to rewrite and improve everything (one of the problems with getting older is you look back at your past work and think you can do things better). This fourth edition is the culmination of about 6 years of full-time work (on top of my actual job). This book has literally consumed the last 15 years or so of my life, and each time I get a nice email from someone who found it useful I am reminded that it is the most useful thing I'll ever do with my life. It began and continues to be a labour of love. It still isn't perfect, and I still love to have feedback (good or bad) from the people who matter most: you.

Andy



[www.facebook.com/profandyfield](http://www.facebook.com/profandyfield)



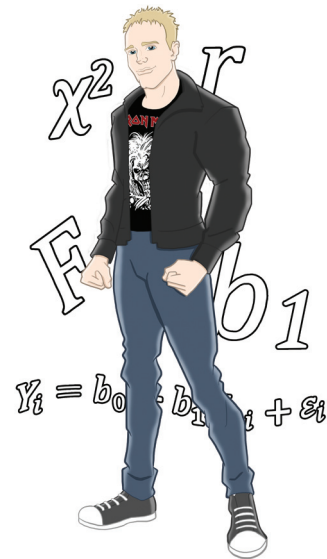
[@ProfAndyField](https://twitter.com/ProfAndyField)



[www.youtube.com/user/ProfAndyField](http://www.youtube.com/user/ProfAndyField)



[www.methodspace.com/profile/ProfessorAndyField](http://www.methodspace.com/profile/ProfessorAndyField)





# HOW TO USE THIS BOOK

When the publishers asked me to write a section on ‘How to use this book’ it was tempting to write ‘Buy a large bottle of Olay anti-wrinkle cream (which you’ll need to fend off the effects of ageing while you read), find a comfy chair, sit down, fold back the front cover, begin reading and stop when you reach the back cover.’ However, I think they wanted something more useful.☺

## What background knowledge do I need?

In essence, I assume that you know nothing about statistics, but that you have a very basic grasp of computers (I won’t be telling you how to switch them on, for example) and maths (although I have included a quick revision of some very basic concepts, so I really don’t assume much).

## Do the chapters get more difficult as I go through the book?

Yes, more or less: Chapters 1–9 are first-year degree level, Chapters 8–15 move into second-year degree level, and Chapters 16–20 discuss more technical topics. However, my main aim is to tell a statistical story rather than worrying about what level a topic is at. Many books teach different tests in isolation and never really give you a grasp of the similarities between them; this, I think, creates an unnecessary mystery. Most of the tests in this book are the same thing expressed in slightly different ways. I want the book to tell this story, and I see it as consisting of seven parts:

- Part 1 (Doing research linear models): Chapters 1–3.
- Part 2 (Exploring data): Chapters 4–6.
- Part 3 (Linear models with continuous predictors): Chapters 7 and 8.
- Part 4 (Linear models with continuous or categorical predictors): Chapters 9–15.
- Part 5 (Linear models with multiple outcomes): Chapter 16 and 17.
- Part 6 (Linear models with categorical outcomes): Chapters 18–19.
- Part 7 (Linear models with hierarchical data structures): Chapter 20.

This structure might help you to see the method in my madness. If not, to help you on your journey I’ve coded each section with an icon. These icons are designed to give you

an idea of the difficulty of the section. It doesn't mean you can skip the sections (but see Smart Alex in the next section), but it will let you know whether a section is at about your level, or whether it's going to push you. It's based on a wonderful categorization system using the letter 'I':

- ① *Introductory*, which I hope means that everyone should be able to understand these sections. These are for people just starting their undergraduate courses.
- ② *Intermediate*. Anyone with a bit of background in statistics should be able to get to grips with these sections. They are aimed at people who are perhaps in the second year of their degree, but they can still be quite challenging in places.
- ③ *In at the deep end*. These topics are difficult. I'd expect final-year undergraduates and recent postgraduate students to be able to tackle these sections.
- ④ *Incinerate your brain*. These are difficult topics. I would expect these sections to be challenging for undergraduates, but postgraduates with a reasonable background in research methods shouldn't find them too much of a problem.

## Why do I keep seeing stupid faces everywhere?



**Brian Haemorrhage:** Brian is a really nice guy, and he has a massive crush on Jane Superbrain. He's seen her around the university campus carrying her jars of brains (see below). Whenever he sees her, he gets a knot in his stomach and he imagines slipping a ring onto her finger on a beach in Hawaii, in front of their loving family and friends. Jane never even notices him; this makes him very sad. His friends have told him that the only way she'll marry him is if he becomes a statistics genius (and changes his surname). Therefore, he's on a mission to learn statistics. It's his last hope of impressing Jane, settling down and living happily ever after. At the moment he knows nothing, but he's about to embark on a journey that will take him from statistically challenged to a genius, in 900 pages. Along his journey he pops up and asks questions, and at the end of each chapter he flaunts his newly found knowledge to Jane in the hope she'll go on a date with him.



**New! Confusius:** The great philosopher Confucius had a lesser-known brother called Confusius. Jealous of his brother's great wisdom and modesty, Confusius vowed to bring confusion to the world. To this end, he built the confusion machine. He puts statistical terms into it, and out of it come different names for the same concept. When you see Confusius he will be alerting you to statistical terms that mean the same thing.



**Cramming Sam:** Samantha thinks statistics is a boring waste of time and she just wants to pass her exam and forget that she ever had to know anything about normal distributions. She appears and gives you a summary of the key points that you need to know. If, like Samantha, you're cramming for an exam, she will tell you the essential information to save you having to trawl through hundreds of pages of my drivel.



**Curious Cat:** He also pops up and asks questions (because he's curious). The only reason he's here is because I wanted a cat in the book ... and preferably one that looks like mine. Of course the educational specialists think he needs a specific role, and so his role is to look cute and make bad cat-related jokes.

**Jane Superbrain:** Jane is the cleverest person in the whole universe. A mistress of osmosis, she acquired vast statistical knowledge by stealing the brains of statisticians and eating them. Apparently they taste of sweaty tank tops. Having devoured some top statistics brains and absorbed their knowledge, she knows all of the really hard stuff. She appears in boxes to tell you advanced things that are a bit tangential to the main text. Her friends tell her that a half-whit called Brian is in love with her, but she doesn't know who he is.



**Labcoat Leni:** Leni is a budding young scientist and he's fascinated by real research. He says, 'Andy, man, I like an example about using an eel as a cure for constipation as much as the next guy, but all of your data are made up. We need some real examples, dude!' So off Leni went: he walked the globe, a lone data warrior in a thankless quest for real data. He turned up at universities, cornered academics, kidnapped their families and threatened to put them in a bath of crayfish unless he was given real data. The generous ones relented, but others? Well, let's just say their families are sore. So, when you see Leni you know that you will get some real data, from a real research study to analyse. Keep it real.



**New! Oditi's Lantern:** Oditi believes that the secret to life is hidden in numbers and that only by large-scale analysis of those numbers shall the secrets be found. He didn't have time to enter, analyse and interpret all of the data in the world, so he established the cult of undiscovered numerical truths. Working on the principle that if you gave a million monkeys typewriters, one of them would re-create Shakespeare, members of the cult sit at their computers crunching numbers in the hope that one of them will unearth the hidden meaning of life. To help his cult Oditi has set up a visual vortex called 'Oditi's Lantern'. When Oditi appears it is to implore you to stare into the lantern, which basically means there is a video tutorial to guide you.



**Oliver Twisted:** With apologies to Charles Dickens, Oliver, like the more famous fictional London urchin, is always asking 'Please, Sir, can I have some more?' Unlike Master Twist though, our young Master Twisted wants more statistics information. Of course he does, who wouldn't? Let us not be the ones to disappoint a young, dirty, slightly smelly boy who dines on gruel. When Oliver appears he's telling you that there is additional information to be found on the companion website. (It took a long time to write, so someone please actually read it.)



**Satan's Personal Statistics Slave:** Satan is a busy boy – he has all of the lost souls to torture in hell; then there are the fires to keep fuelled, not to mention organizing enough carnage on the planet's surface to keep Norwegian black metal bands inspired. Like many of us, this leaves little time for him to analyse data, and this makes him very sad. So, he has his own personal slave, who, also like some of us, spends all day dressed in a gimp mask and tight leather pants in front of IBM SPSS analysing Satan's data. Consequently, he knows a thing or two about SPSS, and when Satan's busy spanking a goat, he pops up in a box with SPSS tips.



**Smart Alex:** Alex is a very important character because he appears when things get particularly difficult. He's basically a bit of a smart Alec, and so whenever you see his face you know that something scary is about to be explained. When the hard stuff is over he reappears to let you know that it's safe to continue. You'll also find that Alex gives you tasks to do at the end of each chapter to see whether you're as smart as he is.





## Why do I keep seeing QR codes?



**Study on the Go:** QR stands for ‘quantum reality’, and if you scan one of these funny little barcode things into your mobile device (smartphone etc.) it will transport you and your device into a quantum reality in which left is right, time runs backwards, drinks pour themselves out of your mouth into bottles, and statistics is interesting. Scanning these codes will be your gateway to mobile-friendly versions of the materials on the companion website. It means you can have stats anytime, anywhere, even on the toilet.

## What is on the companion website?

In this age of downloading, CD-ROMs are for losers (at least that’s what the ‘kids’ tell me), so I’ve put my cornucopia of additional funk on that worldwide interweb thing. To enter my world of delights, go to [www.sagepub.co.uk/field4e](http://www.sagepub.co.uk/field4e). The website contains resources for students and lecturers alike, organized by chapter and character (unlike last time):

- **Testbank:** There is a (hopefully) comprehensive testbank of multiple-choice and numeracy-based/algorithmic questions for instructors to use. It comes in two flavours: (1) a file that you can upload into your online teaching system (with answers separated out), or (2) preloaded into an online homework and grading system called WebAssign® ([www.webassign.net](http://www.webassign.net)). WebAssign enables you to assign questions for exams and assignments. Students can receive instant feedback on correct/incorrect answers, including pointers to areas in the book where the right answer can be found. 
- **Data files:** You need data files to work through the examples in the book and they are all on the companion website. We did this so that you’re forced to go there and once you’re there Sage will flash up subliminal messages to make you buy more of their books.
- **Resources for different subject areas:** I am a psychologist and although I tend to base my examples around the weird and wonderful, I do have a nasty habit of resorting to psychology when I don’t have any better ideas. I realize that not everyone is as psychologically oriented as me, so my publishers have recruited some non-psychologists to provide data files and an instructor’s testbank of multiple-choice questions for those studying or teaching in **business and management**, **education**, **sport sciences** and **health sciences**. You have no idea how happy I am that I didn’t have to write those.
- **Webcasts:** Whenever you see Odit in the book it means that there is a webcast to accompany the chapter. These are hosted on my YouTube channel ([www.youtube.com/user/ProfAndyField](http://www.youtube.com/user/ProfAndyField)), which I have amusingly called  $\mu$ -Tube (see what I did there?). You can also get to them via the companion website. 
- **Self-assessment multiple-choice questions:** Organized by chapter, these will allow you to test whether wasting your life reading this book has paid off so that you can annoy your friends by walking with an air of confidence into the examination. If you fail said exam, please don’t sue me.
- **Flashcard glossary:** As if a printed glossary wasn’t enough, my publishers insisted that you’d like an electronic one too. Have fun here flipping through terms and definitions covered in the textbook; it’s better than actually learning something.

- **Oliver Twisted's pot of gruel:** Oliver Twisted will draw your attention to the 300 pages or so of more technical information that we have put online so that (1) the planet suffers a little less, and (2) you won't die when the book falls off of your bookshelf onto your head.
- **Labcoat Leni solutions:** For all of the Labcoat Leni tasks in the book there are full and detailed answers on the companion website.
- **Smart Alex answers:** Each chapter ends with a set of tasks for you to test your newly acquired expertise. The chapters are also littered with self-test questions. The companion website contains around 300 pages (that's a different 300 pages to the 300 above) of detailed answers. Will I ever stop writing?
- **PowerPoint slides:** I can't come and teach you all in person (although you can watch my lectures on YouTube). Instead I rely on a crack team of highly skilled and super-intelligent pan-dimensional beings called 'lecturers'. I have personally grown each and every one of them in a greenhouse in my garden. To assist in their mission to spread the joy of statistics I have provided them with PowerPoint slides for each chapter. If you see a picture of an anus on their slides, then remember that's probably my fault.
- **Links:** Every website has to have links to other useful websites, and the companion website is no exception.
- **Cyberworms of knowledge:** I have used nanotechnology to create cyberworms that crawl down your broadband connection, pop out of the USB port of your computer and fly through space into your brain. They rearrange your neurons so that you understand statistics. You don't believe me? Well, you'll never know for sure unless you visit the companion website ....

Happy reading, and don't get distracted by Facebook and Twitter.

Companion Website To come

# ACKNOWLEDGEMENTS

This book (in all its SPSS, SAS and R versions) wouldn't have happened if not for Dan Wright's unwarranted faith in the ability of a then postgraduate to write the first SPSS edition. Numerous other people have contributed to previous editions of this book. I don't have room to list them all, but particular thanks are due to Dan (again), David Hitchin, Laura Murray, Gareth Williams, Lynne Slocombe and Kate Lester, who gave me significant amounts of feedback on various incarnations of this text. For this edition, lots of people sent me very helpful emails, but in particular Maria de Ridder (for a very helpful document of suggestions) and Thom Baguley (for feedback on Chapter 5 at very short notice).

Special thanks to Jeremy Miles for his help with various versions of this book over the years. Part of his 'help' involves ranting on at me about things I've written being, and I quote, 'bollocks'. Nevertheless, working on the SAS and R versions of this book with him has influenced me enormously. He's also been a very nice person to know over the past few years (apart from when he's ranting on at me about ...).

Thanks to the following for sending me their raw data – it's an honour for me to include their fascinating research in my book: Rebecca Ang, Philippe Bernard, Hakan Çetinkaya, Tomas Chamorro-Premuzic, Graham Davey, Mike Domjan, Gordon Gallup, Nicolas Guéguen, Sarah Johns, Eric Lacourse, Nate Lambert, Sarah Marzillier, Karlijn Massar, Geoffrey Miller, Peter Muris, Laura Nichols, Nick Perham, Achim Schüetzwohl, Mirjam Tuk, and Lara Zibarras.

I appreciate everyone who has taken time to write nice reviews of this book on the various Amazon (and other) websites around the world; the success of this book has been in no small part due to these people being so positive and constructive in their feedback. Thanks also to everyone who contributes so enthusiastically to my Facebook page: I hit some motivational dark times during this edition, but feeling the positive vibes from readers always got me back on track (especially the photos of cats, dogs, parrots and lizards with this book ☺). I continue to be amazed and bowled over by the nice things that people say about the book.

Not all contributions are as tangible as those above. Very early in my career Graham Hole made me realize that teaching research methods didn't have to be dull. My whole approach to teaching has been to steal all of his good ideas, and he has had the good grace not to ask for them back! He is a rarity in being brilliant, funny *and* nice.

This book wouldn't exist without the generous support of IBM who allow me to beta-test SPSS Statistics and keep me up to date with the software ([www-01.ibm.com/software/analytics/spss](http://www-01.ibm.com/software/analytics/spss)). I wrote this edition on a Mac but used Windows for the screen shots. Mac and Mac OS are trademarks of Apple Inc., registered in the United States and other countries; Windows is a registered trademark of Microsoft Corporation in the United States and other countries. Thanks to Jess Knott at TechSmith ([www.techsmith.com](http://www.techsmith.com)) who provided support for Camtasia (which I use to record and edit my webcasts) and Snagit (which I used for screen shots) for Mac; she also breathed new life into some of my old webcasts. I

created most diagrams and flowcharts in this book using OmniGraffle ([www.omnigroup.com](http://www.omnigroup.com)). Although it is unhealthy to love a piece of software, I love OmniGraffle and if I ever meet the authors I will buy them beer, and plenty of it.

My publishers, Sage, are rare in being a large, successful company that manages to maintain a family feel. Like any family, I don't always see eye-to-eye with them, and sometimes we drive each other mad, but we wouldn't want to be without each other. They generously co-funded my wife to help update the web materials for this book. My editor Mike takes his fair share of crap from me (what does he expect, he supports Tottenham?), but I always appreciate his unfaltering enthusiasm, support and willingness to make things happen. Thanks to everyone at Sage, but especially Ian, Karen (who supports a proper football team) and Ziyad (who doesn't know what football is) for their help and support over many years.

You'll notice the book is riddled with nicely drawn characters. I didn't draw them. The ones that survived from the previous edition are thanks to Alex Lee. Special thanks to Laura-Jane at Anelina Illustrations ([www.anelinaillustrations.com](http://www.anelinaillustrations.com)) for my caricature, Oditi and Confusius. It was great working with you.

I always write listening to music. For this edition I predominantly enjoyed (my neighbours less so): Absu, Anathema, Anthrax, Animals as Leaders, Audrey Horne, The Beyond, Black Breath, Black Tusk, Black Sabbath, Blue Öyster Cult, Blut Aus Nord, Deathspell Omega, Deep Purple, Foo Fighters, Genesis, Graveyard, Ihsahn, Iron Maiden, Jethro Tull, Kiss, Manowar, Marillion, Meshuggah (a lot), Metallica, Mastodon, Motörhead, Opeth (a lot), Primal Rock Rebellion, Rainbow, Rush, Secrets of the Moon, Status Quo, Steve Wilson, Storm Corrosion, Sylosis, Torche, Uriah Heep, Watain, and Wolves in the Throne Room.

All this book-writing nonsense requires many lonely hours of typing. Without some wonderful friends to drag me out of my dimly lit room from time to time I'd be even more of a gibbering cabbage than I already am. Peter Muris, Birgit Mayer, and especially Leonora Wilkinson gave me a lot of support while writing previous editions of this book. For this edition, and in most cases all of the previous ones, my eternal gratitude goes to Graham Davey, Ben Dyson, Martin Watts, Sam Cartwright-Hatton, Mark Franklin and their lovely families for reminding me that there is more to life than work. You'd also be hard pushed to find more supportive, loving and proud parents than mine. I also throw out a robust set of horns to my brothers of metal Doug Martin and Rob Mephram for letting me deafen them with my drumming ([www.myspace.com/fracturepattern](http://www.myspace.com/fracturepattern)). I also became an uncle while writing this update, so thanks to Melody for being small and cute.

For someone who spends his life writing, I'm constantly surprised at how incapable I am of finding words to express how wonderful my wife Zoë is. She has a never-ending supply of patience, love, support and optimism (even when her husband is a grumpy, sleep-deprived, withered, self-doubting husk). As if that wasn't enough, she also did the lion's share of the updating of the accompanying web materials and SPSS screen shots for this edition. I never forget, not even for a nanosecond, how lucky I am.

With thanks to the following people for their feedback on the fourth edition

Mahmood Ali, University of Greenwich  
 Rory Allen, Goldsmiths, University of London  
 Derek Ashford, Manchester Metropolitan University  
 Chris Askew, Kingston University  
 Mirjam Baars, University of Amsterdam  
 Theodoros Bampouras, University of Cumbria  
 Kirsten Bartlett, Sheffield Hallam University  
 Roger Bennett, London Metropolitan University  
 Kanishka Bhattacharya, Oxford University  
 Kate Black, University of Chester  
 Ian Boardley, University of Birmingham  
 Nicola Brown, St Mary's University College  
 Douglas Bryson, ESC Rennes School of Business  
 Hannah Buchanan-Smith, University of Stirling  
 Louise Bunce, London Metropolitan University  
 Scott Burnet, Southampton Solent University  
 Robert Busching, University of Potsdam  
 Katherine Cagney, Waterford Institute of Technology  
 Ian Charity, Newcastle Business School  
 Sam Chenery-Morris, University Campus Suffolk  
 I-Chant Chiang, Quest University Canada  
 Nikki Coghill, University of Bristol  
 Graham Cookson, King's College London  
 Julie Davies, Bangor University  
 Matthew Davis, University of Leeds  
 Rutger de Graaf, University of Amsterdam  
 Peter de Waal, Utrecht University  
 Pedro Dias, Catholic University of Portugal  
 Gerlof Donga, Hogeschool van Amsterdam  
 Michelle Ellefson, University of Cambridge  
 Chris Fife-Schaw, University of Surrey  
 Liam Foster, University of Sheffield  
 Pauline Fox, University of West London  
 Robert Francis, King's College London  
 Benjamin Gardner, University College London  
 Lise Georgeson, St Mary's University College  
 Elisabeth Götze, Vienna University of Economics and Business  
 Martijn Goudbeek, Tilburg University  
 Hannah Greatwood, Leeds Metropolitan University  
 Suzanne Hacking, University of Central Lancashire  
 Sebastian Hagen, University of Leipzig  
 Carol Haigh, Manchester Metropolitan University  
 Julie Hall, Birmingham City University  
 Karen Hambly, University of Kent  
 Karolina Hansen, Friedrich Schiller University Jena  
 Azmi Hassali, Universiti Sains Malaysia  
 Erica Hepper, University of Southampton  
 Beryl Hilberink-Schulpen, Radboud University Nijmegen  
 Henriette Hogh, University of Surrey  
 Lee Hulbert-Williams, University of Wolverhampton  
 Gerry Humphris, University of St Andrews  
 Martyn Jarvis, University of Glamorgan  
 Domhnall Jennings, Newcastle University  
 Paul Jepson, University of Birmingham  
 Stephen Johnston, Swansea University  
 Marie Juanchich, Kingston University  
 Serdar Karabati, Bilgi University  
 Roshni Khatri, University of Northampton  
 Paul Kiff, University of East London  
 Thomas Koch, LMU Munich  
 Wolfgang Kotowski, University of Zurich  
 Wander Lowie, University of Groningen  
 Ruth Lowry, University of Chichester  
 John Mallett, University of Ulster  
 Chris Mamo, Cork Institute of Technology  
 Irene Manaras, University of Hertfordshire  
 Anne Manyande, University of West London  
 Jesse Martin, Bangor University  
 Carla Martins, University of Minho  
 Paul McCarthy, Glasgow Caledonian University  
 Philip McDonald, Leeds Trinity University College  
 Margaret McGrath, National University of Ireland, Galway  
 Gary Mckenna, University of the West of Scotland  
 Alistair McMillan, University of Sheffield  
 Juanjo Medina, University of Manchester  
 Remo Mombarg, Hanze University of Applied Sciences, Groningen  
 Maria Mos, Tilburg University  
 Eyob Mulat-Weldemeskel, London Metropolitan University  
 Drew Munn, University of Northampton  
 Tony Myers, Newman University College  
 Rich Neil, Cardiff Metropolitan University  
 Dennis Nigbur, Canterbury Christ Church University  
 Louis Passfield, University of Kent  
 Monique Pollmann, Tilburg University  
 Michelle Pyer, University of Northampton  
 Shahid Qureshi, Institute of Business Administration, Karachi  
 Eric Rietzschel, University of Groningen  
 Claire-Marie Roberts, University of Worcester  
 Janet Robertson, Lancaster University  
 Petros Roussos, University of Athens  
 Tom Scherndl, University of Salzburg  
 Kim Schildkamp, University of Twente  
 Shivani Sharma, University of Hertfordshire  
 Luke Sloan, Cardiff University  
 Sandrino Smeets, Radboud University Nijmegen  
 Andy Smith, Leeds Metropolitan University  
 Martin Southam, Canterbury Christ Church University  
 Jayne Tidd, Teesside University  
 Alastair Tomlinson, Cardiff Metropolitan University  
 Alison Tresidder, University of Bedfordshire  
 Ivo van der Lans, Wageningen University  
 Frans Van der Slik, Radboud University Nijmegen  
 Marcel van Egmond, University of Amsterdam  
 Nel Verhoeven, University College Roosevelt Academy  
 Paresch Wankhade, Liverpool Hope University  
 Maggie Whittaker, University of Essex  
 Bob Williams, Durham University



## Dedication

---

Like the previous editions, this book is dedicated to my brother Paul and my cat Fuzzy, because one of them is a constant source of intellectual inspiration and the other wakes me up in the morning by sitting on me and purring in my face until I give him cat food: mornings will be considerably more pleasant when my brother gets over his love of cat food for breakfast. 😊

# SYMBOLS USED IN THIS BOOK

## Mathematical operators

$\Sigma$	This symbol (called sigma) means 'add everything up'. So, if you see something like $\Sigma x_i$ it just means 'add up all of the scores you've collected'.
$\Pi$	This symbol means 'multiply everything'. So, if you see something like $\Pi x_i$ it just means 'multiply all of the scores you've collected'.
$\sqrt{x}$	This means 'take the square root of x'.

## Greek symbols

$\alpha$	The probability of making a Type I error
$\beta$	The probability of making a Type II error
$\beta_i$	Standardized regression coefficient
$\varepsilon$	Usually stands for 'error'
$\eta^2$	Eta squared
$\mu$	The mean of a population of scores
$\rho$	The correlation in the population
$\sigma$	The standard deviation in a population of data
$\sigma^2$	The variance in a population of data
$\sigma_{\bar{x}}$	The standard error of the mean
$\tau$	Kendall's tau (non-parametric correlation coefficient)
$\chi^2$	Chi-square test statistic
$\chi^2_F$	Friedman's ANOVA test statistic
$\omega^2$	Omega squared (an effect size measure). This symbol also means 'expel the contents of your intestine immediately into your trousers'; you will understand why in due course.

## English symbols

$b_i$	The regression coefficient (unstandardized), I tend to use it for any coefficient in a linear model.
$df$	Degrees of freedom
$e_i$	The error associated with the $i$ th person
$F$	$F$ -ratio (test statistic used in ANOVA)
$H$	Kruskal–Wallis test statistic
$k$	The number of levels of a variable (i.e., the number of treatment conditions), or the number of predictors in a regression model
$\ln$	Natural logarithm
MS	The mean squared error: the average variability in the data.
$N, n, n_i$	The sample size. $N$ usually denotes the total sample size, whereas $n$ usually denotes the size of a particular group
$P$	Probability (the probability value, p-value or significance of a test are usually denoted by $p$ )
$r$	Pearson's correlation coefficient
$r_s$	Spearman's rank correlation coefficient
$r_b, r_{pb}$	Biserial correlation coefficient and point-biserial correlation coefficient, respectively
$R$	The multiple correlation coefficient
$R^2$	The coefficient of determination (i.e., the proportion of data explained by the model)
$s$	The standard deviation of a sample of data
$s^2$	The variance of a sample of data
SS	The sum of squares, or sum of squared errors, to give it its full title
$SS_A$	The sum of squares for variable $A$
$SS_M$	The model sum of squares (i.e., the variability explained by the model fitted to the data)
$SS_R$	The residual sum of squares (i.e., the variability that the model can't explain – the error in the model)
$SS_T$	The total sum of squares (i.e., the total variability within the data)
$t$	Test statistic for Student's $t$ -test
$T$	Test statistic for Wilcoxon's matched-pairs signed-rank test
$U$	Test statistic for the Mann–Whitney test
$W_s$	Test statistic for Wilcoxon's rank-sum test
$\bar{X}$	The mean of a sample of scores
$z$	A data point expressed in standard deviation units

## SOME MATHS REVISION

**Two negatives make a positive:** Although in life two wrongs don't make a right, in mathematics they do! When we multiply a negative number by another negative number, the result is a positive number. For example,  $-2 \times -4 = 8$ .

**A negative number multiplied by a positive one makes a negative number:** If you multiply a positive number by a negative number then the result is another negative number. For example,  $2 \times -4 = -8$ , or  $-2 \times 6 = -12$ .

**BODMAS:** This is an acronym for the order in which mathematical operations are performed: Brackets, Order, Division, Multiplication, Addition, and Subtraction. Mostly these operations are self-explanatory (e.g., always calculate things within brackets first) except for 'order', which refers to power terms such as squares. Four squared, or  $4^2$ , used to be called four raised to the order of 2, hence the word 'order' in BODMAS (also, if we used 'power', we'd end up with BPDMAS, which doesn't roll off the tongue quite so nicely). Let's look at an example of BODMAS: what would be the result of  $1 + 3 \times 5^2$ ? The answer is 76 (not 100 as some of you might have thought). There are no brackets so the first thing is to deal with the order term:  $5^2$  is 25, so the equation becomes  $1 + 3 \times 25$ . There is no division, so we can move on to multiplication:  $3 \times 25$ , which gives us 75. BODMAS tells us to deal with addition next:  $1 + 75$ , which gives us 76 and the equation is solved. If I'd written the original equation as  $(1 + 3) \times 5^2$ , then the answer would have been 100 because we deal with the brackets first:  $(1 + 3) = 4$ , so the equation becomes  $4 \times 5^2$ . We then deal with the order term, so the equation becomes  $4 \times 25 = 100$ .

<http://www.bbc.co.uk/schools/gcsebitesize/maths/> is a good site for revising basic maths.