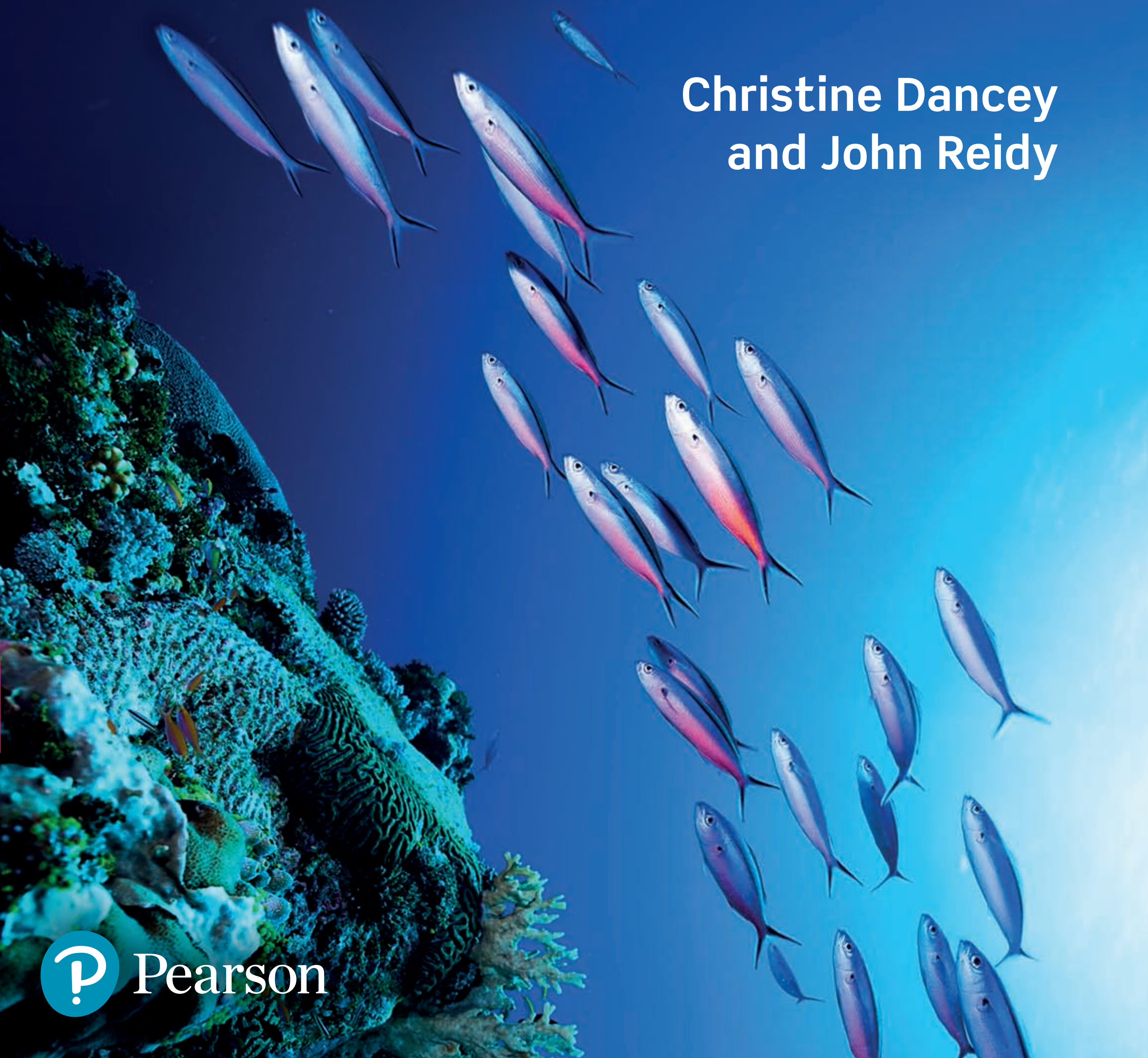
SEVENTH EDITION

Statistics without Maths for Psychology

Christine Dancey
and John Reidy



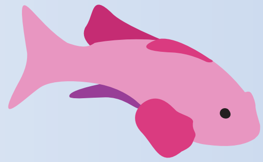
Pearson

Statistics Without Maths for Psychology

British Psychological Society standards in Quantitative Methods in Psychology

The British Psychological Society (BPS) accredits psychology degree programmes across the UK. It has set guidelines as to which major topics should be covered within quantitative methods in psychology. We have listed these topics below and indicated where in this textbook each is covered most fully.

BPS guidelines on the teaching of quantitative methods in psychology	Which chapters?
Descriptive and summary statistics	3, 4 and 5
Probability theory	4 and 5
The normal distribution	3, 4 and 5
Statistical inference	4 and 5
Confidence intervals	4
Mean and error bar graphs	4
Non-parametric alternatives to t-tests	16
Tests of proportions	9
Cramer's Phi as a measure of association in contingency tables	—
McNemar's test of change	—
Bivariate correlation and linear regression	6 and 12
The analysis of variance	10, 11 and 15
Non-parametric alternatives to one factor analyses of variance	16
The choice of an appropriate statistical analysis	5



Statistics Without Maths for Psychology

Seventh Edition

Christine P. Dancey University of East London

John Reidy Sheffield Hallam University

Pearson

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney
Dubai • 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 1999 (print)
Second edition 2002 (print)
Third edition 2004 (print)
Fourth edition 2008 (print)
Fifth edition 2011 (print)
Sixth edition 2014 (print and electronic)
Seventh edition published 2017 (print and electronic)

ISBN: 978-1-292-12885-6 (print)
978-1-292-12889-4 (PDF)
978-1-292-13027-9 (ePub)

British Library Cataloguing-in-Publication Data
A catalogue record for the print edition is available from the British Library

Library of Congress Cataloging-in-Publication Data
Dancey, Christine P., author. | Reidy, John, author.
Statistics without maths for psychology / Christine P. Dancey, University of East London, John Reidy, Sheffield Hallam University.
Seventh Edition. | New York : Pearson, [2017] | Revised edition of the authors' Statistics without maths for psychology, 2014.
LCCN 2016059329| ISBN 9781292128856 (print) | ISBN 9781292128894 (pdf) ISBN 9781292130279 (epub)
LCSH: SPSS for Windows. | Mathematical statistics. | Psychology--Statistical methods.
LCC BF39 .D26 2017 | DDC 150.1/5195--dc23
LC record available at https://lccn.loc.gov/2016059329

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

Print edition typeset in 10/12pt Times New Roman PS Pro by SPi Global
Printed in Slovakia by Neografia

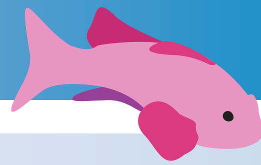
NOTE THAT ANY PAGE CROSS REFERENCES REFER TO THE PRINT EDITION

Christine would like to dedicate this book to Donna Wiles and Linda Perkins. Our close friendship and support for each other is very important to me. You are both strong, beautiful and fantastic people. Thanks a million, for everything.

John would like to dedicate this book to Ollie ... super schnotz (100% Schnauzer)

Brief contents

Contents



Companion Website

ON THE WEBSITE

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

Lecturer Resources

For password-protected online resources tailored to support the use of this textbook in teaching, please visit **www.pearsoned.co.uk/dancey**

Preface

It seems incredible to us that it is now 18 years since our book was first published. We have been amazed at how well the book has been received and thankful for the kind words tutors and students alike have said about it. In this seventh edition of the book we have kept true to our vision for the book to provide conceptual explanations of statistical concepts without making you suffer through the formulae. We have built upon the strengths of the previous editions and updated our examples from the literature, updated some of the practical exercises, provided reflections from authors of published research and responded, with revised explanations, to a number of reviewers who kindly provided feedback on the sixth edition.

We wrote this book primarily for our students, most of whom disliked mathematics, and could not understand why they had to learn mathematical formulae when their computer software performed the calculations for them. They were not convinced by the argument that working through calculations gave them an understanding of the test – neither were we. We wanted them to have a conceptual understanding of statistics and to *enjoy* data analysis. Over the past 20 years we have had to adapt our teaching to large groups of students, many of whom have no formal training in mathematics. We found it was difficult to recommend some of the traditional statistics textbooks – either they were full of mathematical formulae, and perceived by the students as dull or boring, or they were simple, statistical cookbook recipes, which showed them how to perform calculations, but gave them no real understanding of what the statistics meant. We therefore decided to write this book, which seeks to give students a conceptual understanding of statistics while avoiding the distraction of formulae and calculations.

Another problem we found with recommending statistics textbooks was the over-reliance on the probability value in the interpretation of results. We found it difficult to convince them to take effect size, and confidence intervals, into consideration when the textbooks that were available made no mention of the debates around hypothesis testing, but simply instructed students to say $p < 0.05$ is significant and $p > 0.05$ is not significant! We hope in writing this book that students will become more aware of such issues.

We also wanted to show students how to incorporate the results of their analysis into laboratory reports, and how to interpret results sections of journal articles. Until recently, statistics books ignored this aspect of data analysis. Of course, we realise that the way we have written our example 'results sections' will be different from the way that other psychologists would write them. Students can use these sections to gain confidence in writing their own results, and hopefully they will build on them, as they progress through their course.

We have tried to simplify complex, sometimes very complex, concepts. In simplifying, there is a trade-off in accuracy. We were aware of this when writing the book, and have tried to be as accurate as possible, while giving the simplest explanation. We are also aware that some students do not use SPSS (an IBM company*) for their data analysis. IBM® SPSS® Statistics, however,

*SPSS was acquired by IBM in October 2009.

is the most commonly used statistical package for the social sciences, and this is why the text is tied so closely to SPSS. Students not using this package should find the book useful anyway. This edition of the book has been updated for use with SPSS version 23 and earlier.

As with the sixth edition of the book we have included information about the authors of articles which we have drawn upon in the writing of this book – and have included photos of them where possible – strictly with their permission, of course. We also asked them why they had chosen their particular research topic, and whether they had encountered any problems in the running of the experiment/study. We thought this would enrich the text. Although we have updated many examples from the literature, we have left in some early studies because they illustrate exactly the points made in the text. Some reviewers thought there should be more challenging activities and/or multiple choice questions. Therefore, we have added activities which are based on examples from the literature, and require students to *interpret* the material, in their own words. They can then compare their interpretation with the authors' interpretation.

We hope that students who read the book will not only learn from it, but also enjoy our explanations and examples. We also hope that as a result of reading this book students will feel confident in their ability to perform their own statistical analyses.

How to use this book

To help you get the most from this book we thought that it would be useful to provide a brief overview of the book and of the structure of the chapters. The best way to use the book if you are new to statistics in psychology or if you have been away from statistics for a long while is to work your way through the chapters from Chapter 1 onwards. The most important chapters to read and ensure that you understand fully are the first five chapters as these provide you with the core concepts for comprehending the main statistical techniques covered later in the book. If you spend the time and effort on these opening chapters then you will be rewarded by having a better understanding of what the statistical tests are able to tell us about our data. We cannot stress enough the importance of such an understanding for appropriate use of statistical techniques and for your ability to understand and critique others' use of such techniques.

The chapters that follow these opening chapters generally explain the concepts underlying specific types of tests as well as how to conduct and interpret the findings from these. We start off with the more basic tests which look at the fewest possible variables ('variables' will be explained in Chapter 1) and then using these as a basis we move on to the more complex tests later in the book. In some ways it might be better to read about a basic type of test, say simple correlations (see Chapter 6), and then move on to the more complex versions of these tests, say regression and multiple regression (see Chapter 12). As another example, start with simple tests of differences between two groups (in Chapter 7) and then move on to tests of differences between more than two groups (Chapters 10 and 11). However, often statistics modules don't follow this sort of pattern but rather cover all of the basic tests first and only then move on to the complex tests. In such a learning pattern there is the danger that to some extent some of the links between the simple and complex tests may get lost.

Rather disappointingly we have read some reviews of the book which focus entirely on the step-by-step guides we give to conducting the statistical analyses with SPSS for Windows (now called SPSS Statistics). We would like to stress that this book is not simply a 'cookbook' for how to run statistical tests. If used appropriately you should come out with a good understanding of the statistical concepts covered in the book as well as the skills necessary to conduct the analyses using SPSS Statistics. If you already have a conceptual understanding of the statistical techniques covered in the book then by all means simply follow the step-by-step guide to carrying out the analyses, but if you are relatively new to statistics you should ensure that you read the text so that you understand what the statistical analyses are telling you.

There are a number of features in this book to help you learn the concepts being covered (in technical terms these are called 'pedagogic' features). These are explained below, but before we explain these we will give you a general overview of what to expect in each chapter.

In each chapter we will highlight what is to come and then we will explain the statistical concepts underlying the particular topics for that chapter. Once we have covered the statistical concepts you will be given step-by-step guides to conducting analyses using SPSS Statistics. Towards the end of each chapter you will be provided with a means of testing your knowledge, followed by some pointers to further reading. We will now describe some of the features found in the chapters in more detail.

At the beginning of every chapter there is a **Chapter overview.** These overviews provide you with information about what is contained in each chapter and what you should have achieved from working through it. Sometimes we will also highlight what you need to know beforehand to be able to get the most from the chapter. You should make sure that you read these (it is very easy to get into the habit of not doing this) as they will set the scene for you and prepare your mind for the concepts coming up in the book.

At the end of each chapter there are **Summaries** which outline the main concepts that were covered. These are important for consolidating what you have learnt and help put the concepts learnt later in the chapter back in the context of the earlier concepts. You will also find **SPSS Statistics exercises, activities and multiple choice questions.** We cannot stress enough the importance of working through these when you finish each chapter. They are designed to test your knowledge and to help you actively work with the information that you have learnt. The best way to learn about things is to do them. The answers to the multiple choice questions are also provided at the very end of each chapter so that you can check your progress. If you have answered questions incorrectly go back and read the relevant part of the chapter to ensure that you have a good understanding of the material. The answers to the SPSS Statistics exercises are provided at the end of the book. Check these and if you have different answers go back and try to work out where you might have gone wrong. Often it might be that you have input the data incorrectly into SPSS Statistics. There are additional multiple choice questions and SPSS Statistics exercises on the companion website and so please do make use of these also.

Within each chapter there are a number of features designed to get you thinking about what you have been reading. There are **Discussion points** which help you to explore different ideas or theories in more detail. There are also a number of **Activity boxes** which provide additional opportunities for you to test your understanding of the theories and ideas being discussed. It is important to complete the activities as we have placed these to ensure that you are actively engaging with the material. Our experience has shown that actively working with material helps learning (and makes reading more enjoyable). You will also find a number of **Example boxes** where we provide a concrete example of what we are discussing. Providing such concrete examples helps students understand the concepts more easily. There are also lots of **examples from the psychological literature** which show how active psychology researchers use the statistical techniques which have been covered in the chapters.

Where appropriate we have included as many **diagrams and pictures** as we can as these will help you to understand (and remember) the text more easily. The thought of giving you endless pages of text without breaking it up is not worth thinking about. This would probably lead to a lot of Zzzzzz. On a serious note though, remember that the pictures are not there to be pretty nor just to break up the text. Please consult these along with reading the text and this will help you learn and understand the concept under discussion. Occasionally in the book you will come across **Caution boxes.** These are there to warn you of possible problems or issues related to certain techniques or statistical concepts. These are useful in many ways as they are designed to help you to understand some of the limits of statistical tests and they serve as a reminder that we have to think carefully about how we analyse our data.

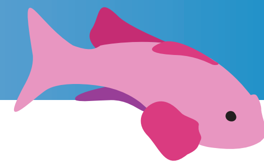
Where in a chapter we want to show you how to use SPSS Statistics we provide **annotated screenshots.** These will show you which buttons to click in SPSS Statistics as well as how and where to move information around to get the analyses that you want. Finally, at the end of each chapter there is a **Reference** section. In this we will provide details of all the other authors' works that we have mentioned within the chapter. This is pretty much what you should do when writing up your own research. Some of the references will provide the details of the examples from the literature that we have presented and some will be examples of potentially useful further reading. You can follow up these as and when you choose to. Sometimes it is good to follow up the examples from the research literature as you can then see the context to the example analyses that we present. Also, by looking at how the experts present their research you can better learn how to present your research.

Companion website

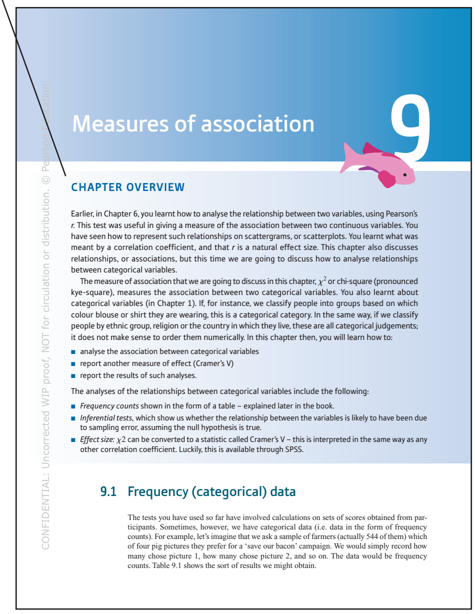
We would urge you to make as much use as possible of the resources available to you on the companion website. When you get on to the site you will see that it is broken down into resources for each chapter. For each chapter you will find **SPSS Statistics dataset files** which are simply the data for the examples that we provide in each chapter. You can access these to ensure that you have input data correctly or so that you can carry out the same analyses that we present in each chapter to make sure that you get the same results. Also, on the website you will find **additional multiple choice questions.** If you find that you have made mistakes in the multiple choice questions provided in the book you should go back through the chapter and try to make sure that you fully understand the concepts presented. It wouldn't make sense for you to then test yourself using the same multiple choice questions and so we have provided the additional ones on the companion website. As another means of testing yourself and to help you actively learn we provide **additional SPSS Statistics exercises** for each chapter and a step-by-step guide to the analysis to conduct on this data and how to interpret the output.

Finally, you will find **links to interesting and useful websites** which are relevant to the concepts being covered in each chapter.

Guided tour



The chapter overview gives you a feel for what will be covered and what you should have learnt by the end of the topic.



CONFIDENTIAL: Uncorrected WIP proof, NOT for circulation or distribution. © Pearson Education.

Measures of association 9

CHAPTER OVERVIEW

Earlier, in Chapter 6, you learnt how to analyse the relationship between two variables, using Pearson's *r*. This test was useful in giving a measure of the association between two continuous variables. You have seen how to represent such relationships on scattergrams, or scatterplots. You learnt what was meant by a correlation coefficient, and that *r* is a natural effect size. This chapter also discusses relationships, or associations, but this time we are going to discuss how to analyse relationships between categorical variables.

The measure of association that we are going to discuss in this chapter, χ^2 or chi-square (pronounced kye-square), measures the association between two categorical variables. You also learnt about categorical variables (in Chapter 1). If, for instance, we classify people into groups based on which colour blouse or shirt they are wearing, this is a categorical category. In the same way, if we classify people by ethnic group, religion or the country in which they live, these are all categorical judgements; it does not make sense to order them numerically. In this chapter then, you will learn how to:

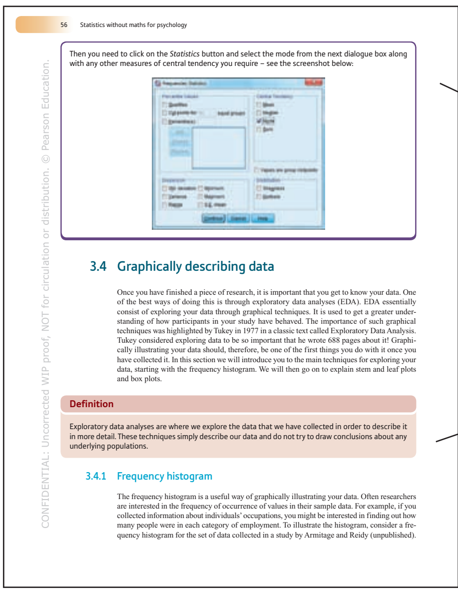
- analyse the association between categorical variables
- report another measure of effect (Cramer's V)
- report the results of such analyses.

The analyses of the relationships between categorical variables include the following:

- *Frequency counts* shown in the form of a table – explained later in the book.
- *Inferential tests*, which show us whether the relationship between the variables is likely to have been due to sampling error, assuming the null hypothesis is true.
- *Effect size:* $\chi 2$ can be converted to a statistic called Cramer's V – this is interpreted in the same way as any other correlation coefficient. Luckily, this is available through SPSS.

9.1 Frequency (categorical) data

The tests you have used so far have involved calculations on sets of scores obtained from participants. Sometimes, however, we have categorical data (i.e. data in the form of frequency counts). For example, let's imagine that we ask a sample of farmers (actually 544 of them) which of four pig pictures they prefer for a 'save our bacon' campaign. We would simply record how many chose picture 1, how many chose picture 2, and so on. The data would be frequency counts. Table 9.1 shows the sort of results we might obtain.



56 Statistics without maths for psychology

CONFIDENTIAL: Uncorrected WIP proof, NOT for circulation or distribution. © Pearson Education.

Then you need to click on the *Statistics* button and select the mode from the next dialogue box along with any other measures of central tendency you require – see the screenshot below:

3.4 Graphically describing data

Once you have finished a piece of research, it is important that you get to know your data. One of the best ways of doing this is through exploratory data analyses (EDA). EDA essentially consist of exploring your data through graphical techniques. It is used to get a greater understanding of how participants in your study have behaved. The importance of such graphical techniques was highlighted by Tukey in 1977 in a classic text called Exploratory Data Analysis. Tukey considered exploring data to be so important that he wrote 688 pages about it! Graphically illustrating your data should, therefore, be one of the first things you do with it once you have collected it. In this section we will introduce you to the main techniques for exploring your data, starting with the frequency histogram. We will then go on to explain stem and leaf plots and box plots.

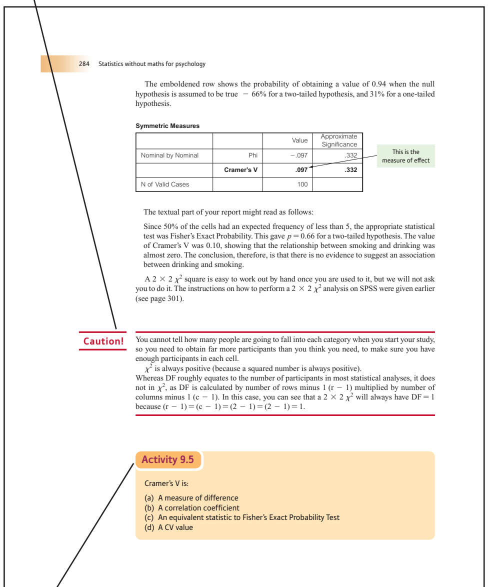
Definition

Exploratory data analyses are where we explore the data that we have collected in order to describe it in more detail. These techniques simply describe our data and do not try to draw conclusions about any underlying populations.

3.4.1 Frequency histogram

The frequency histogram is a useful way of graphically illustrating your data. Often researchers are interested in the frequency of occurrence of values in their sample data. For example, if you collected information about individuals' occupations, you might be interested in finding out how many people were in each category of employment. To illustrate the histogram, consider a frequency histogram for the set of data collected in a study by Armitage and Reidy (unpublished).

Caution boxes highlight possible problems you may encounter or issues for consideration.



284 Statistics without maths for psychology

The emboldened row shows the probability of obtaining a value of 0.94 when the null hypothesis is assumed to be true − 66% for a two-tailed hypothesis, and 31% for a one-tailed hypothesis.

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	−.097	.332
	Cramer's V	**.097**	**.332**
N of Valid Cases		100	

This is the measure of effect

The textual part of your report might read as follows:

Since 50% of the cells had an expected frequency of less than 5, the appropriate statistical test was Fisher's Exact Probability. This gave $p = 0.66$ for a two-tailed hypothesis. The value of Cramer's V was 0.10, showing that the relationship between smoking and drinking was almost zero. The conclusion, therefore, is that there is no evidence to suggest an association between drinking and smoking.

A 2×2 χ^2 square is easy to work out by hand once you are used to it, but we will not ask you to do it. The instructions on how to perform a 2×2 χ^2 analysis on SPSS were given earlier (see page 301).

Caution!

You cannot tell how many people are going to fall into each category when you start your study, so you need to obtain far more participants than you think you need, to make sure you have enough participants in each cell.

χ^2 is always positive (because a squared number is always positive).

Whereas DF roughly equates to the number of participants in most statistical analyses, it does not in χ^2, as DF is calculated by number of rows minus 1 $(r - 1)$ multiplied by number of columns minus 1 $(c - 1)$. In this case, you can see that a 2×2 χ^2 will always have DF = 1 because $(r - 1) = (c - 1) = (2 - 1) = (2 - 1) = 1$.

Activity 9.5

Cramer's V is:

(a) A measure of difference
(b) A correlation coefficient
(c) An equivalent statistic to Fisher's Exact Probability Test
(d) A CV value

Activity boxes provide you with opportunities to test your understanding as you go along.

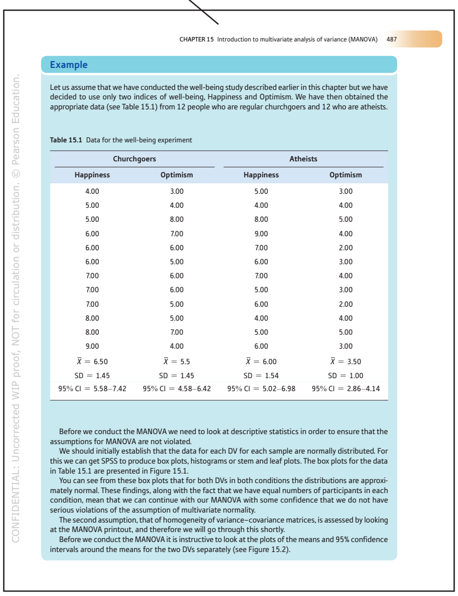
SPSS sections guide you through how to use the software for each process, with annotated, full-colour screenshots to demonstrate what should be happening on screen.

Definitions explain the key terms you need to understand statistics.

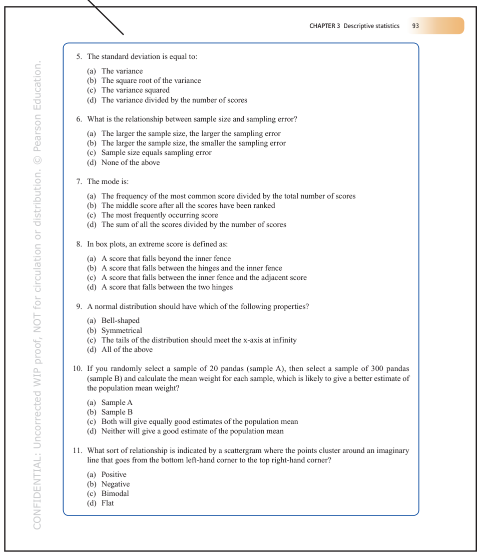
Personal reflection boxes bring statistics to life through interviews with researchers, showing their important role in psychological discoveries.

Examples from the literature highlight a key piece of research in the area.

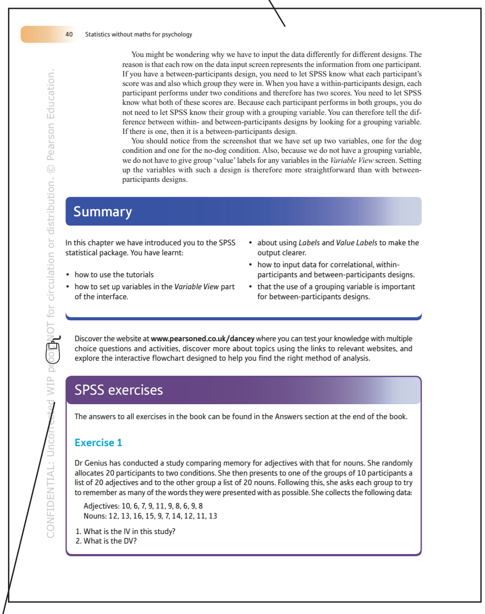
Numerous examples in each chapter illustrate the key points.



Multiple choice questions at the end of each chapter allow you to test your knowledge.



Chapter summaries enable you to revise the main points of the chapter after you've read it.



SPSS exercises at the end of each chapter give you an opportunity to test yourself using real data.

Acknowledgements

Our grateful thanks go to the reviewers of this seventh edition of the book for their time and valuable help:

Paul Warren - University of Manchester
Richard Rowe - Sheffield University
Jennifer Murray - Edinburgh Napier University

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

Screenshots

Screenshots on pages 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39,53, 54-55, 66, 70, 88, 89, 120, 124, 125, 157, 158, 159, 160, 163, 164, 165, 166, 188, 189, 197, 198, 201, 202, 228, 229, 234, 235, 235, 269, 270, 271, 273, 274, 275, 279, 280, 305, 306, 313, 314, 315, 346, 347, 348, 349, 350, 351, 359, 360, 361, 368, 369, 380, 381, 382, 391, 392, 416, 417, 418, 419, 420, 432, 433, 468, 469, 470, 473, 495, 496, 503, 504, 505, 517, 518, 523, 524, 530, 531, 537, 538, 542 from International Business Corporation, Reprint Courtesy of International Business Machines Corporation, © International Business Machines Corporation. IBM, the IBM logo, ibm.com, PASW and SPSS are trademarks or registered trademarks of International Business Machines Corporation, 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]www.ibm.com/legal/copytrade.shtml.

Tables

Table on page 259 from Health complaints and unemployment: the role of self-efficacy in a prospective cohort study, *Journal of Social and Clinical Psychology,* 32, 97–115 (Zenger, M., Berth, H., Brähler, E. and Stöbel-Richter, Y. 2013), republished with permission of Guilford Press, permission conveyed through Copyright Clearance Center, Inc.; Table on page 289 adapted from Everyday memory in children with developmental coordination disorder (DCD), *Research in Developmental Disabilities,* 34, pp. 687–94 (Chen, I. C., Tsai, P. L., Hsu, Y. W., Ma, H. I. and Lai, H. A. 2013), Copyright © 2013, with permission from Elsevier; Table on page 311 from Differential effects of age on involuntary and voluntary autobiographical memory, *Psychology and Aging,* 24, pp. 397–411 (Schlagman, S., Kliegel, M., Schulz, J. and Kvavilashvili, L. 2009), Copyright © 2009 American Psychological Association.

Text

Extract on page 2 from Statistics Commission Report No. 38 – Official Statistics: Value and Trust, page 38, © Crown copyright. Contains public sector information licensed under the Open Government Licence (OGL) v3.0. http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/; Extract on page 15 from Perceptual biases in relation to paranormal and conspiracy beliefs, *PLoS ONE,* 10 (van Elk M. 2015), reproduced with permission; Extract on page 101 from No regrets? Measuring the career benefits of a psychology placement year, *Assessment & Evaluation in Higher Education,* 37, Iss. 5 (Moores, E. and Reddy, P. 2012), reprinted by permission of Taylor & Francis Ltd, http://www.tandfonline.com, Copyright © 2012 Routledge; Quote on page 141 from Geoffrey R. Loftus reproduced with permission; Extract on page 149 from Politicians and trustworthiness: acting contrary to self-interest enhances trustworthiness, *Basic and Applied Social Psychology,* 32, pp. 328–39 (Combs, D. J. and Keller, P. S. 2010), Copyright © 2010 Routledge. Reprinted by permission of Taylor & Francis Ltd, http://www.tandfonline.com; Extract on page 204 from Evaluation of a new measure of mood intolerance, the Tolerance of Mood States Scale (TOMS): Psychometric properties and associations with eating disorder symptoms, *Eating Behaviors,* 13, Iss. 4 (Allen, K.L., Mclean, N.J. and Byrne, S.M. 2012), Copyright © 2012, with permission from Elsevier; Extract on page 207 from Age-related differences in recognition memory for items and associations: contribution of individual differences in working memory and metamemory, *Psychology & Aging,* 27 (Bender, R. and Raz, N 2012), reprinted by permission of Taylor & Francis Ltd, http://www.tandfonline.com; Extract on pages 203–04 from Sexting as an intervention: relationship, satisfaction and motivation considerations, *The American Journal of Family Therapy,* 41 (Parker, T.S. Blackburn, K.M., Perry, M.S. and Hawks, J.M. 2013), Copyright © 2013 Routledge, reprinted by permission of Taylor & Francis Ltd, http://www.tandfonline.com; Extract on page 233 from A comparative study on the attitudes and uses of music by adults with visual impairments and those who are sighted — JVIB Abstract, *Journal of Visual Impairment & Blindness,* 109 (Park, H.Y., Chong, H.J. and Kim, S.J. 2015), republished with permission of American Foundation for the Blind; permission conveyed through Copyright Clearance Center, Inc.; Extract on page 247 from Beyond the null ritual: Formal modeling of psychological processes, *Journal of Psychology,* 217 (Marewski, J. N., and Olsson, H. 2009), Copyright © 2009 American Psychological Association; Quote on page 247 from Statistical procedures and the justification of knowledge in psychological science, *American Psychologist,* 44 (Rosnow, R.J. and Rosenthal, R. 1989), Copyright © 1989 American Psychological Association; Extract on page 247 from On statistical testing in psychology, *British Journal of Psychology,* 88 (Macdonald, R.R. 1997), British journal of psychology by BRITISH PSYCHOLOGICAL SOCIETY Reproduced with permission of CAMBRIDGE UNIVERSITY PRESS. in the format Book via Copyright Clearance Center; Extract on page 288 from Collectivism and the meaning of suffering, *Journal of Personality and Social Psychology,* 103 (Sullivan, D., Landau, M. J., Kay, A. C. and Rothschild, Z. K. 2012), Copyright © 2012 *American Psychological Association*; Extract on page 358 from A quick eye to anger: An investigation of a differential effect of facial features in detecting angry and happy expressions, *International Journal of Psychology* (Lo, L. Y., and Cheng, M. Y. 2015), International journal of psychology by INTERNATIONAL UNION OF PSYCHOLOGICAL SCIENCE Reproduced with permission of PSYCHOLOGY PRESS in the format Book via Copyright Clearance Center; Extract on page 367 from Smartphone applications utilizing biofeedback can aid stress reduction, *Frontiers in Psychology,* 7, 832 (Dillon, A., Kelly, M., Robertson, I. H., & Robertson, D. A. 2016), Copyright © 2016 Dillon, Kelly, Robertson and Robertson.

Picture Credits

The publisher would like to thank the following for their kind permission to reproduce their photographs:

(Key: b-bottom; c-centre; l-left; r-right; t-top)

Dr Karina Allen: 204; **Manna Alma:** 318tl; **Ellen Boddington:** 464l; **Jonathan Lent:** 405r; **Geoffrey Loftus:** 247b; **Professor Mark McDermott:** 464r; **Dr Liz Moores:** 100br; **Peter Reddy:** 100r; **Robert Rosenthal:** 249br; **Sarah Partington:** 84b; **Daniel Sullivan:** 287tr.

Cover images: *Front:* **Getty Images**

Variables and research design

1

CHAPTER OVERVIEW

In trying to explain how to use and understand statistics it is perhaps best to start by outlining the principal factors in designing research. We will therefore describe the most important aspects of research design with a view to explaining how they influence the use of statistics. In this chapter, therefore, we aim to teach you about the following:

- variables: continuous, discrete and categorical
- independent and dependent variables
- correlational, experimental and quasi-experimental designs
- between-participant and within-participant designs.

1.1 Why teach statistics without mathematical formulae?

Statistics as a topic tends to strike fear into the hearts and minds of most social science students and a good many lecturers too. Understanding statistical concepts should, however, be no more difficult than understanding any other theoretical concept (for example, the concept of intelligence). In fact, one would think that understanding a very concrete concept such as the arithmetical mean would be a good deal easier than understanding a rather vague psychological concept such as ‘an attitude’. Yet, every year, it seems that the majority of students, who apparently grasp many non-statistical concepts with consummate ease, struggle to understand statistics. Our view is that most people are fearful of statistics because the concepts are lost in the mathematical formulae. We therefore seek to explain statistics in a conceptual way without confusing students with unnecessary mathematical formulae – unnecessary, that is, in these days of statistical computer packages. If students wish to learn these formulae to enhance their knowledge, what better platform to have than a conceptual understanding of statistics?

Statistics tend to have a bad reputation, as this quote often attributed to former British Prime Minister Benjamin Disraeli illustrates: ‘There are three sorts of lies: lies, damned lies and statistics.’ It is not the statistics that are at fault, however, but rather the way they are used. After all, we do not usually blame the gun for killing someone but the person who pulled the trigger. All too often, particularly in politics, statistics are quoted out of context or even used selectively. This problem is clearly illustrated in a letter from Ed Humpherson, the Director

General for regulations at the UK Statistics Authority to Siobhan Carey who was Head of Profession for Statistics at the UK Government Department for Business, Innovation and Skills, sent on 16 February 2016 (you can find this letter on the site by typing 'Carey' in the search box on the homepage). In this letter Ed Humpherson is seen to reprimand the Minister of State Joseph Johnson for the use of complex statistics relating to poor performing UK universities which were not clearly defined and had not been previously published. Ed Humpherson notes that because there was a lack of clarity with these statistics that it was not clear that the proportion of poorly performing universities was high as was implied by Joseph Johnson. The letter concludes with the following: 'The Authority would ask that you raise this with your colleagues and take steps to ensure that future such references to statistics are supported by publication with sufficient commentary and guidance as to enable informed debate.' This clearly indicates an expectation that statistics be used within an appropriate context and be clearly defined and explained. The letter from Ed Humpherson, along with other letters relating to the official use of statistics in the UK, can be found at the UK Statistics Authority website (www.statisticsauthority.gov.uk). This is a really good website as it provides insights into how politicians use and often misuse statistics. Another good website about statistics and research is 'Sense About Science' (www.senseaboutscience.org). This site provides lots of useful information with the intention of helping people better understand science and scientific findings. One part of the site, the 'For the record' section, highlights examples of poor representations of scientific research in the news. A recent example of this was a study reported in the national UK media (e.g. *Daily Mail* and *The Daily Telegraph*). The findings from the original unpublished study were presented at an academic conference in the US and found differences between mice born to mothers exposed to the vapours from e-cigarettes and those born to mothers exposed just to clean air. The study was reported in the media as providing evidence that using e-cigarettes during pregnancy is as bad as, or even worse than, smoking cigarettes. On the 'Sense About Science' website, Professor Peter Hajek clearly outlines the problems with the reporting of this study. He states that this was an unpublished study and so the data cannot be checked and verified and, more fundamentally, the study did not compare the mice exposed to e-cigarette vapour with those exposed to tobacco smoke, and so the comparisons with smoking cigarettes used in the headlines and the newspaper articles themselves were unjustified.

These examples show some of the problems with understanding and reporting of research based upon statistics. Yet politicians and the national media are happy to rely on poorly reported statistics to help colour our judgments about a whole range of issues for their own purposes. We should point out that this is not just a problem for politicians actually in government, it is widespread among politicians. This is even acknowledged in a report by the UK's Statistics Commission which was the forerunner to the UK Statistics Authority. In this report (2008) the Commission states:

> Statistics have been, and always will be, used selectively by politicians and commentators in the course of public debate. The selection and emphasis of particular statistical information to favour, or contest, a policy argument has to be tolerated as part of the political process. It is essential however that, to balance the politically selective use of statistics, the figures themselves, with full explanation, should be equally accessible and understandable to everyone. There should also be public corrections of manifestly misleading interpretations.

These examples clearly illustrate the importance of viewing statistics in the correct context. If we say to you, for example, that the average (mean) height of the adult male is 5 ft 8 in (173 cm), this may be meaningful for British men but not necessarily for men from African pygmy tribes where the average height can be as low as 4 ft 9 in (145 cm). We believe that being able to interpret statistics and whether or not they have been used appropriately is a very important life skill, particularly in the age of the internet and the widespread availability of information (good and bad in quality) about every aspect of life.

1.2 Variables

We have explained a very important aspect of statistics: that they are only meaningful in a context. But what is it that statistics actually do? Essentially, statistics give us information about factors that we can measure. In research the things that we measure are called *variables*.

Variables are the main focus of research in science. A variable is simply something that can vary: that is, it can take on many different values or categories. Examples of variables are gender, typing speed, top speed of a car, number of reported symptoms of an illness, temperature, attendances at rock festivals (e.g. the Download festival), level of anxiety, number of goals scored in football matches, intelligence, number of social encounters while walking your dog, amount of violence on television, occupation, number of cars owned, number of children per family and favourite colours. These are all things that we can measure and record and that vary from one situation or person to another.

But why are we interested in variables? We are generally interested in variables because we want to understand why they vary as they do. In order to achieve such understanding we need to be able to measure and record the changes in these variables in any given situation.

1.2.1 Characteristics of variables

You will notice from the examples of variables above that they have different characteristics. Whereas you can measure temperature in terms of Fahrenheit or Celsius and put a number to it, you cannot meaningfully do this for type of occupation. This represents one important characteristic of variables: that is, how they actually change. At one end of the spectrum we have variables that are said to be *continuous*: that is, they can take any value within a given range. Or, more accurately, the variable itself doesn't change in discrete jumps. A good example of a continuous variable is temperature. This is because you could measure the temperature as, say, 40 °C or you could measure it more accurately as, say, 40.2558 °C. Another less obvious example is the measurement of the amount of violence on television. We could measure this in terms of the amount of time that violence appears on screen per day. If measured in this way, in terms of time, the variable could take on any value in terms of seconds or parts of seconds (e.g. 1000 s or 1000.1235672 s per day). The only limitation in the precision of measurement of such variables is the accuracy of the measuring instrument. With continuous variables there is an assumption that the underlying variable itself is continuous, even if the way in which we measure it is not. Of the examples given earlier, temperature, level of anxiety, top speed of a car, typing speed and intelligence could be regarded as continuous whereas the rest could not (see Table 1.1).

Table 1.1 Examples of continuous, discrete and categorical variables

Continuous	Discrete	Categorical
■ Temperature ■ A car's top speed ■ Typing speed ■ Intelligence ■ Level of anxiety	■ Number of reported symptoms of an illness ■ Number of cars owned ■ Number of goals scored in a football match ■ Number of social encounters while walking your dog ■ Attendances at heavy rock festivals ■ Number of children in a family	■ Gender ■ Occupation ■ Favourite colour ■ Type of fast food restaurant

A variable could also be *discrete*: that is, it can take on only certain discrete values within the range. An example of such a variable is the reported number of symptoms of an illness that a person has. These can only be recorded in terms of presence or absence of symptoms and therefore in terms of whole symptoms present. Another example would be if we chose to measure the amount of violence on television in terms of the number of violent incidents per week. In such a case, we could only report the number of discrete violent incidents. We could not use it to measure in terms of fractions of a violent incident; therefore violence on television measured this way is termed a discrete variable. Of the examples given earlier, the most obvious discrete variables are number of reported symptoms of an illness, number of social encounters while walking your dog, attendance at a rock festival, number of cars owned, number of children per family and number of goals scored in a game of football.

One problem that arises when thinking about continuous and discrete variables is confusing the underlying variable with how it is measured. A variable may in theory be continuous, but the way we measure it will always be discrete, no matter how accurate we are. We could measure anxiety (a theoretically continuous variable) using a questionnaire (e.g. the State–Trait Anxiety Inventory; Spielberger *et al*., 1983) where the total score on the questionnaire gives an indication of a person's level of anxiety. Total scores on this questionnaire can only increase in whole units, say from 38 to 39 or from 61 to 62. Thus, the way we have measured anxiety is discrete whereas the underlying variable is assumed to be continuous.

Additionally, often when analysing discrete variables they are treated as if they were continuous. Many of the statistical tests that we use assume that we have continuous variables. Often when a discrete variable can take on many different values within a range (e.g. attendances at heavy rock festivals) they can reasonably be treated as if they were continuous for the sake of statistical testing.

Another type of variable is a *categorical* variable. This is where the values that the variables can take are categories. A good example is gender, which has only two values that it can take: male or female. Categorical variables can also sometimes have many possible values, as in type of occupation (e.g. judges, teachers, miners, grocers, civil servants). When dealing with categorical data we have an infinite number of variables that we might wish to investigate. We could, if we wished to, categorise people on the basis of whether or not they ate chocolate sponge with tomato ketchup at 6.30 this morning. The only obvious examples of categorical variables given in our list of variables described at the beginning of this section are occupation, gender and favourite colour.

Try to ensure that you understand the different types of variable that you are measuring, as this is important when deciding how to analyse data.

Definitions

Continuous variables can take on absolutely any value within a given range.

Discrete variables can only take on certain discrete values in a range.

Categorical variables are those in which we simply allocate people to categories.

1.2.2 Dichotomising continuous and discrete variables

It is often the case that researchers convert continuous or discrete variables into categorical variables. For example, we might wish to compare the spatial ability of tall and short people. We could do this by comparing people who are over 6 ft 4 in (193 cm) with those under 4 ft 10 in (147 cm) on a spatial ability test. Thus, we have chosen points on the continuous scale (height) and decided to compare those participants who score above and below these points (see Figure 1.1).

Another example might be to compare the memory ability of anxious and non-anxious individuals. We could measure anxiety levels using a questionnaire; this is a continuous variable measured on a discrete scale. For example, the Hospital Anxiety and Depression Scale has an

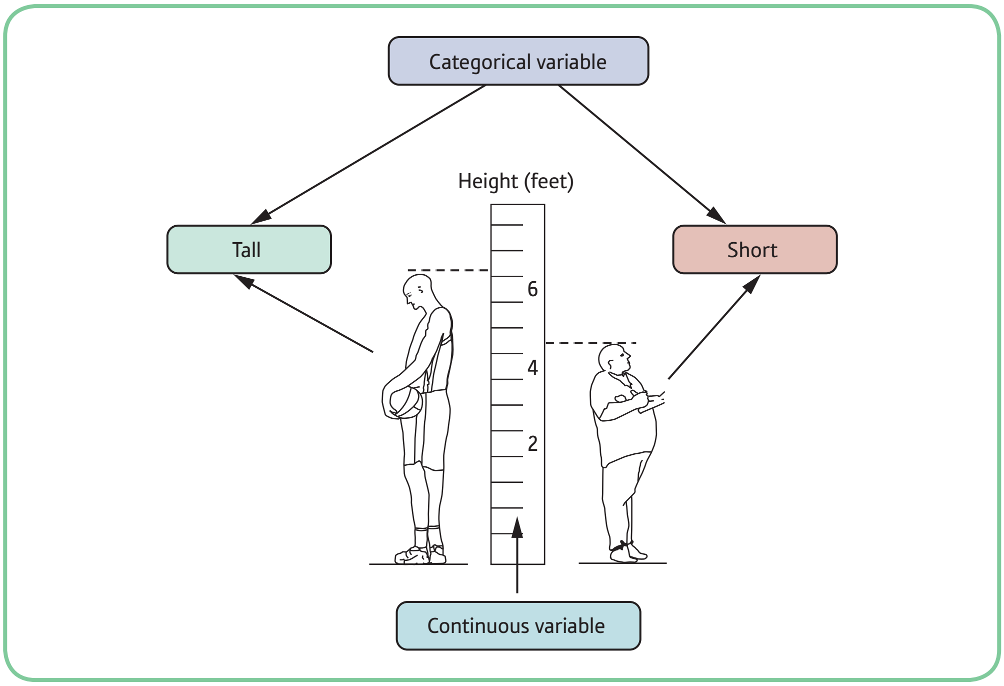


Figure 1.1 Illustration of the conversion of continuous variables into categorical variables

anxiety scale that ranges from 0 to 21. To convert this to a categorical variable we would simply compare those who score above a certain value (say, 11) with those who score below this value.

This dichotomising (dividing into two categories) of continuous and discrete variables is quite common in psychology as it enables us to find out if there are differences between groups who may be at the extremes of the continuous or discrete variables (e.g. tall and short people). We do not, however, recommend such a practice as it reduces the sensitivity of your statistical analyses. There is a good discussion of such problems in Streiner (2002), in Maxwell and Delaney (1993) and more recently in Altman and Royston (2007). We mention this here only so that you are aware that it happens in the research literature and so that you will understand what the researchers have done.

Discussion point

Dichotomising continuous variables

Why do researchers dichotomise variables? Streiner (2002) highlights the point that many decisions in psychology, psychiatry and medicine are binary decisions. Binary decisions are those where there are two choices, such as whether or not a person has a mental disorder, whether or not a person has a specific disease, whether a person should be hospitalised or whether a person should be released from hospital. It is often argued that because clinicians have to make such binary decisions, it is legitimate to investigate variables in a binary way. Such reasoning is used to support the widespread practice of dichotomising continuous variables.

Streiner argues that we do not have to view the sorts of decision that clinicians make as binary. He suggests that it would be better to think of mental illness, for example, as being on a continuum: the more symptoms you have, the more affected you are. We should then measure such constructs on continua rather than dichotomising them. That is, rather than using questionnaires to categorise individuals we could use the questionnaires to get a measure of where they fall on a continuum. Such information can then be utilised in our decisions for treating individuals, etc. It is interesting to note that the latest version of the *Diagnostic and Statistical Manual of Mental Disorders (DSM-V)* has moved much more to seeing mental disorders on a continuum rather than as categorical.

An example may illustrate dichotomisation better. We suggested earlier that we could categorise individuals as anxious or non-anxious on the basis of their scores on a questionnaire. Researchers investigating anxiety sometimes utilise questionnaires in this way. Those participants who score high on the questionnaire are classed as high in anxiety whereas those who have low scores are classed as low in anxiety. The 'median-split' method is often used in this regard, where those participants who score above the median are categorised as anxious and those who score below the median as non-anxious (e.g. Takács *et al*., 2015).

Streiner argues that the practice of dichotomising continuous variables tends to lead to research that is low in power (we cover power further in Chapters 5 and 8). The reason for this is that it results in us losing a lot of information about participants. For example, suppose two individuals score 20 and 38 on an anxiety inventory and that we come to classify them both as low in anxiety (they both fall below the median). In any subsequent analyses based upon this categorisation, both of these participants are treated as being identical in terms of their anxiety levels (i.e. they are both non-anxious). According to our questionnaire, however, there is a very large difference between them in terms of their actual anxiety levels. Treating these two individuals as the same in terms of anxiety level does not seem to make sense. It would be much more sensible to try to include their actual anxiety scores in any statistical analyses that we conduct.

Additionally, we may find that there is a larger difference in terms of anxiety between the two participants classed as non-anxious than there is between two participants where one is classed as anxious and one is not. For example, suppose our median is 39: all those scoring above 39 are classed as anxious and those who score below 39 are non-anxious. We can see here that the non-anxious person who has a score of 38 has much more in common with an anxious person whose score is 41 than they do with another non-anxious person who has a score of 20. Yet in any subsequent analyses the participants with scores of 20 and 38 are classified as identical in terms of anxiety and these are classed as equally different from the person who has a score of 41. This just does not make any sense.

Streiner also highlights research that has shown that analyses using dichotomous variables are about 67% as efficient as analyses using the original continuous/discrete measures. This is an incredible loss of sensitivity in the study. It means that you are only two-thirds as likely to detect relationships among variables if you dichotomise continuous variables. This is a serious handicap to conducting research. Moreover, loss of power is not the only problem that arises when dichotomising variables. Maxwell and Delaney (1993) have shown that such a practice can actually lead to spurious findings arising from statistical analyses.

Therefore, we advise you against dichotomising continuous variables.

Activity 1.1

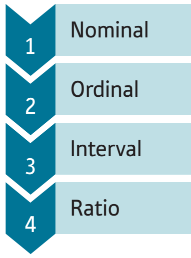
Which of the following are continuous, which are discrete and which are categorical?

- Wind speed
- Types of degree offered by a university
- Level of extroversion
- Makes of car
- Division in which football teams play
- Number of chess pieces 'captured' in a chess game
- Weight of giant pandas
- Number of paintings hanging in art galleries

The correct answers can be found at the end of the book.

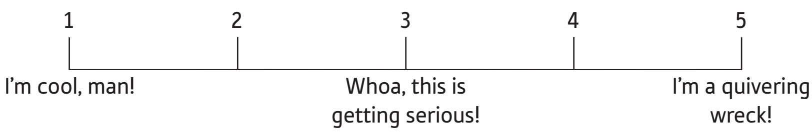
1.3 Levels of measurement

Another way of distinguishing between variables or scales is in terms of the *level of measurement*. There are four levels of measurement and these vary as a function of the way in which the variables are measured. The four different levels are:



At the lowest level of measurement are *nominal scales*. These are in effect categorical variables in that they represent different categories, but they also have the characteristic that there is no particular order that can be given to the categories. A good example of a nominal scale is gender, which has two categories, *male* and *female*. You should be able to see that there is no logical way of ordering these two categories in terms of magnitude. Another example would be ethnic group: again we can categorise people in terms of their ethnic group but we could not put these groups in any particular order – they are simply different categories. When dealing with nominal-level measures, we are simply assigning people to categories and the data we obtain are in the form of *frequency counts*. Frequency counts simply tell us how many people we have in each category.

At the next level of measurement we have *ordinal scales*. Quite often in psychology we use ratings scales to measure participants' responses. For example, we might want to know how nervous a person is just before they take part in a study we are running. We could use a scale like that presented below to gauge how nervous they are.



Using such a scale we can place participants in some sort of order in terms of how nervous they are prior to the study (hence *ordinal scale*). I would be able to say that someone who put a circle around the '1' was less nervous than someone who put a circle around the '3' or around the '5'. One of the drawbacks with such scales is that we cannot say that the difference between '1' and '2' on the scale is the same as the difference between '3' and '4' on the scale or that the difference between 'I'm cool, man!' and 'Whoa, this is getting serious!' is the same as the difference between 'Whoa, this is getting serious!' and 'I'm a quivering wreck!' Thus we do not have equal intervals on the scale.

At the interval level of measurement, we are able to put scores in some sort of order of magnitude and we also have equal intervals between adjacent points on the scale (hence *interval scale*). A good example of an interval scale is one of the commonly used scales to measure temperature, such as Centigrade or Fahrenheit. On such scales we can say that the difference between 1 and 2 degrees is the same as the difference between 9 and 10 degrees or between 99 and 100 degrees. We have equal intervals between adjacent points on the scales. The disadvantage of such scales is that there is no absolute zero on them. Thus whilst there are zero points on both the Centigrade and Fahrenheit scales these are arbitrary zero points – they do not equate to zero temperature. The zero point on the Centigrade scale was chosen as it was the point at which water freezes, and the zero point on the Fahrenheit scale is equally arbitrary. When we reach zero on these scales we cannot say that there is no heat or no temperature.

Because of this we cannot say that 4 °C is half as warm as 8 °C or that 40 °C is twice as hot as 20 °C. In order to make such statements we would need a measuring scale that had an absolute rather than an arbitrary zero point. A good example from the psychological literature is anxiety which is usually measured through questionnaires such as the Spielberger State-Trait Anxiety Inventory. A zero score on this questionnaire doesn't mean that a person has absolutely no anxiety and we cannot say that a person with a score of 40 is twice as anxious as a person with a score of 20.

The final level of measurement is the *ratio scale*. Ratio scales have all the features of interval-level data but with the addition of having an absolute zero point. For example, if I wanted to measure how long it took you to read this paragraph, I would start the timer going when you started at the beginning of the paragraph and then stop it when you had read the last word of the paragraph. Here we have a scale where the intervals between adjacent points are equal: that is, the difference between 1 and 2 seconds is the same as that between 79 and 80 seconds. We also have a zero point which is an absolute zero. The point where you are just preparing to start reading the paragraph is zero in terms of time spent reading the paragraph. Another example of a ratio scale is speed of a car. When the car is not moving it has zero speed (an absolute zero point) and the difference between 9 and 10 k.p.h. is the same as that between 29 and 30 k.p.h. The useful point about having an absolute zero is that we can form ratios using such scales (hence *ratio scales*). Thus, I can say that a car moving at 100 k.p.h. is moving twice as fast as one moving at 50 k.p.h. Or a person who read this paragraph in 30 seconds read it twice as fast as someone who read it in 60 seconds.

Levels of measurement are important as they can have an influence on what sorts of statistical test we can use to analyse our data. Usually, we can only use the most sensitive statistical techniques (called parametric tests) when we have either interval- or ratio-level data. If we have nominal- or ordinal-level data, we have to make do with the less sensitive non-parametric tests (we cover the conditions for using different types of test in more detail in Chapter 5).

Definitions

Ratio scales have equal intervals between adjacent scores on the scale and an absolute zero.

Interval scales have equal intervals between adjacent scores but do not have an absolute zero.

Ordinal scales have some sort of order to the categories (e.g. in terms of magnitude) but the intervals between adjacent points on the scale are not necessarily equal.

Nominal scales consist of categories that are not ordered in any particular way.

1.4 Research designs

There are many different statistical techniques that we use to analyse the data we have collected in research. We will be introducing you to some of the most widely used in this book as well as providing you with an understanding of the factors which determine which statistical technique should be used in a given situation.

One of the biggest factors in determining which statistical tests you can use to analyse your data is the way you have designed your study. There are several ways to design a study and the way you do so can have a great influence on the sorts of statistical procedure that are available to you. Sometimes researchers wish to look for differences between two groups of participants on a particular variable and at other times they might want to see if two variables are related in some way. An example of a study which investigated differences between conditions is the research reported by Guéguen and Ciccotti (2008). In this study the researchers were interested in whether or not dogs facilitate social interactions and helping behaviours among adults. The researchers ran four different studies where male and female researchers walked with and without dogs. In two studies the researcher approached people and asked for some money, in another study the researcher dropped some coins to see if people would help to pick them up and in a final study a male researcher approached females in the street and asked them for their phone numbers. In each study the researcher did the tasks both with and without dogs. In all four studies they found that helping behaviours were higher when the researcher had a dog than when they didn't have a dog. An example of research looking for relationships would be the study reported by Antonacopoulos and Pychyl (2014). In this research they were interested in the relationship between dog walking and mental health. Through an online questionnaire they discovered that talking with others whilst walking a dog was related to how lonely people felt such that increases in talking to others was associated with decreased loneliness. The statistical tests that we would use in these examples are called *difference tests* and *correlational tests* respectively. The way you design your study will influence which of these sorts of test you can use. In the following sections we will take you through several ways of designing studies and indicate which sorts of test are available to the researcher conducting such studies.

1.4.1 Extraneous and confounding variables

Above we described a study by Guéguen and Ciccotti (2008) about the effects of walking with a dog on social interactions and helping behaviours. If you think about this study you may realise that there are factors other than owning a dog that could also affect the social encounters people have when they are out with their dogs. Other factors might include shyness of the walker, attractiveness of the walker, gender of the walker, breed of dog and a whole host of other variables. These are all factors that the researcher might not have accounted for but which may have influenced the social interactions; they are called *extraneous variables*. In any research situation, whether in chemistry, physics or psychology, account has to be taken of extraneous variables. If extraneous variables are overlooked, the conclusions that may be drawn from the studies may be unreliable. Thus, in the dog-walking example, if the extraneous variables just described had not been controlled, we would not be able to say for certain that any differences in social interactions were due to the ownership of a dog. The differences may have been due to any one or a combination of the extraneous variables just described. The main reason for conducting research under laboratory conditions is to try to control extraneous variables as much as possible. You will find that many of the research issues that we describe in this chapter are designed to reduce extraneous variables.

You have to be aware that for any given variable that you measure there will be a number of other variables that may be related to it (see Figure 1.2, for example). When we conduct a study such as the dog and social interaction one, we cannot be certain that it is being with (or without) a dog that has led to a change in social interactions. Thus we need to try to eliminate the other variables (extraneous variables) as possible reasons for our observed changes in social interactions. We do this by trying to control these other variables: for example, by trying to match our dog and no dog participants as much as possible on shyness, attractiveness and gender. Also, we could ensure that all participants are out with the same type of dog and that they are out at the same time and on the same day of the week. Once we have controlled these other variables then we may be more confident in our conclusions that being out with a dog influences the number of social interactions a person will have.

Definition

Extraneous variables are those variables that might have an impact on the other variables that we are interested in but we may have failed to take these into account when designing our study.