

A step by step guide to data analysis using IBM SPSS

SPSS

Survival Manual



Worldwide
Bestseller

Julie Pallant

This is what readers from around the world say about the *SPSS Survival Manual*:

‘The book provides a no-nonsense approach to understanding SPSS. With step by step explanations to carrying out data analysis it is a welcome addition to the resources for any social science student.’

Gail Steptoe Warren, University of Coventry, UK

‘I just wanted to say how much I value Julie Pallant's *SPSS Survival Manual*. It's quite the best text on SPSS I've encountered and I recommend it to anyone who's listening!’

Professor Carolyn Hicks, Health Sciences, Birmingham University, UK

‘An excellent introduction to using SPSS for data analysis ... It provides a self-contained resource, with more than simply (detailed and clear) step by step descriptions of statistical procedures in SPSS. There is also a wealth of tips and advice, and for each statistical technique a brief, but consistently reliable explanation is provided.’

Associate Professor George Dunbar, University of Warwick, UK

‘This most recent edition of Julie Pallant's SPSS bible continues to combine a number of essential elements: clear explanations of different use cases for SPSS; guides on interpreting the (often voluminous and poorly labelled) output; and example data files (from real studies) to practice on ... I had PhD students, this would be their welcome gift on their first day. Essential.’

Dr. P. J. A. Wicks, Research Psychologist, London

‘Having perceived myself as one who was not confident in anything statistical, I worked my way through the book and with each turn of the page gained more and more confidence until I was running off analyses with (almost) glee. I now enjoy using SPSS and this book is the reason for that.’

Dr Marina Harvey, Centre for Professional Development, Macquarie University, Australia

‘I have two copies of Julie Pallant's *SPSS Survival Manual*—one for the home office and one for school— which are both well-worn from lending. I never miss a chance to recommend this useful guide to other doctoral students as a “24-hour TA” to review syntax, interpretation of output, or presentation of results. The coordinating data sets available online are useful for conducting a first run through on each technique for labs or tutoring. I highly recommend this resource.’

Doctoral student, University of California, Los Angeles, USA

‘This book really lives up to its name ... I highly recommend this book to any MBA student carrying out a dissertation project, or anyone who needs some basic help with using SPSS and data analysis techniques.’

Business student, UK

‘This book is simple to understand, easy to read and very concise. Those who have a general fear or dislike for statistics or statistics and computers should enjoy reading this book.’

Lloyd G. Waller PhD, Jamaica

‘There are several SPSS manuals published and this one really does “do what it says on the tin” ... Whether you are a beginner doing your BSc or struggling with your PhD research (or beyond!), I wholeheartedly recommend this book.’

British Journal of Occupational Therapy

‘I love the *SPSS Survival Manual* ... I can't imagine teaching without it. After seeing my copy and hearing me talk about it many of my other colleagues are also utilising it.’

Wendy Close PhD, Psychology Department, Wisconsin Lutheran College, USA

'... being an external student so much of the time is spent teaching myself. But this has been made easier with your manual as I have found much of the content very easy to follow. I only wish I had discovered it earlier.'

Anthropology student, Australia

'This book is a "must have" introduction to SPSS. Brilliant and highly recommended.'

Dr Joe, South Africa

'The strength of this book lies in the explanations that accompany the descriptions of tests and I predict great popularity for this text among teachers, lecturers and researchers.'

Roger Watson, Journal of Advanced Nursing

'I didn't think it was possible for me to love SPSS but with the help of this book I do! The step by step guide is everything I need to use this difficult software. I would recommend it to anyone!'

Alissa Johnston, Occupational Therapy student

'I love this book! I haven't touched stats or SPSS in nearly fifteen years. This book told me everything I needed to know to do my job better with clear, concise language. It's like she knew what all my questions were before I asked them! Awesome!'

T. James, Australia, Amazon.com

'For me, this is the essential guide to data analysis using SPSS. From getting started with SPSS to explaining the sometimes complex output produced by it, from deciding which analysis to use to providing practice materials to hone your skills, this guides you through the process using simple, clear instructions, appropriate to both the novice and experienced researcher.'

Julia Robertson, University of Buckinghamshire, UK

'I have four SPSS manuals and have found that this is the only manual that explains the issues clearly and is easy to follow. SPSS is evil and anything that makes it less so is fabulous.'

Helen Scott, Psychology Honours student, University of Queensland, Australia

'To any students who have found themselves faced with the horror of SPSS when they had signed up for a degree in psychology—this is a godsend.'

Psychology student, Ireland

'This is the best SPSS manual I've had. It's comprehensive and easy to follow. I really enjoy it.'

Norshidah Mohamed, Kuala Lumpur, Malaysia

'I am currently in a doctoral-level Methods and Analysis of Quantitative Research class, and I was struggling to understand SPSS. The university provided web-based tutorial videos, but the videos were too fast and skipped over important steps. I am so glad I found this book! The author breaks everything down into plain English and makes the software very easy to understand. I am now getting an A in my course, due in no small part to this book.'

L. E. Eastman, Doctoral statistics student, USA

'Simply the best book on introductory SPSS that exists. I know nothing about the author but having bought this book in the middle of a statistics open assignment I can confidently say that I love her and want to marry her. There must be dozens of books that claim to be beginners' guides to SPSS. This one actually does what it says. Totally brilliant.'

J. Sutherland, Amazon.co.uk

**A step by step guide to data analysis using IBM
SPSS**

5th edition

Julie Pallant



Open University Press
McGraw-Hill Education
McGraw-Hill House
Shoppenhangers Road
Maidenhead
Berkshire
England
SL6 2QL

email: enquiries@openup.co.uk
world wide web: www.openup.co.uk

and Two Penn Plaza, New York, NY 10121-2289, USA

First published 2013

Copyright © Julie Pallant, 2013

All rights reserved. Except for the quotation of short passages for the purposes of criticism and review, no part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publisher or a licence from the Copyright Licensing Agency Limited. Details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Ltd of Saffron House, 6-10 Kirby Street, London, EC1N 8TS.

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

ISBN-13: 9780335262588

ISBN-10: 03355262589

eISBN: 9780335262595

Library of Congress Cataloging-in-Publication Data
CIP data applied for

Typeset by Midland Typesetters, Australia
Printed in China at Everbest Printing Co.

Fictitious names of companies, products, people, characters and/or data that may be used herein (in case studies or in examples) are not intended to represent any real individual, company, product or event.

The McGraw-Hill Companies

Contents

[Preface](#)

[Data files and website](#)

[Introduction and overview](#)

[Part One Getting started](#)

[1 Designing a study](#)

[2 Preparing a codebook](#)

[3 Getting to know IBM SPSS](#)

[Part Two Preparing the data file](#)

[4 Creating a data file and entering data](#)

[5 Screening and cleaning the data](#)

[Part Three Preliminary analyses](#)

[6 Descriptive statistics](#)

[7 Using graphs to describe and explore the data](#)

[8 Manipulating the data](#)

[9 Checking the reliability of a scale](#)

[10 Choosing the right statistic](#)

[Part Four Statistical techniques to explore relationships among variables](#)

[11 Correlation](#)

[12 Partial correlation](#)

[13 Multiple regression](#)

[14 Logistic regression](#)

[15 Factor analysis](#)

[Part Five Statistical techniques to compare groups](#)

[16 Non-parametric statistics](#)

[17 T-tests](#)

[18 One-way analysis of variance](#)

[19 Two-way between-groups ANOVA](#)

[20 Mixed between-within subjects analysis of variance](#)

[21 Multivariate analysis of variance](#)

[22 Analysis of covariance](#)

[Appendix: Details of data files](#)

[Recommended reading](#)

[References](#)

[Index](#)

Preface

For many students, the thought of completing a statistics subject, or using statistics in their research is a major source of stress and frustration. The aim of the original *SPSS Survival Manual* (published in 2000) was to provide a simple, step-by-step guide to the process of data analysis using IBM SPSS. Unlike other statistical titles it did not focus on the mathematical underpinnings of the techniques, but rather on the appropriate use of IBM SPSS as a tool. Since the publication of the four editions of the *SPSS Survival Manual*, I have received many hundreds of emails from students who have been grateful for the helping hand (or lifeline).

The same simple approach has been incorporated in this fifth edition. Over the last few years SPSS has undergone a number of changes—including a brief period when it changed name. During 2009 version 18 of the program was renamed **PASW Statistics**, which stands for Predictive Analytic Software. The name was changed again in 2010 to **IBM SPSS**. All chapters in this current edition have been updated to suit version 21 of the package (although most of the material is also suitable for use of earlier versions).

I have resisted urges from students, instructors and reviewers to add too many extra topics, but instead have upgraded and expanded the existing material. This book is not intended to cover all possible statistical procedures available in IBM SPSS, or to answer all questions researchers might have about statistics. Instead, it is designed to get you started with your research and to help you gain confidence in the use of the program to analyse your data. There are many other excellent statistical texts available that you should refer to—suggestions are made throughout each chapter in the book. Additional material is also available on the book's website (details in the next section).

Data files and website

Throughout the book, you will see examples of research that are taken from a number of data files included on the website that accompanies this book: www.allenandunwin.com/spss.

From this site you can download the data files to your hard drive or memory stick by following the instructions on screen. These files can be opened only in IBM SPSS.

The **survey5ED.sav** data file is a 'real' data file, based on a research project that was conducted by one of my graduate diploma classes. So that you can get a feel for the research process from start to finish, I have also included in the Appendix a copy of part of the questionnaire that was used to generate this data and the codebook used to code the data. This will allow you to follow along with the analyses that are presented in the book, and to experiment further using other variables. The full questionnaire can be downloaded from the website.

The second data file (**error5ED.sav**) is the same file as the **survey5ED.sav**, but I have deliberately added some errors to give you practice in [Chapter 5](#) at screening and cleaning your data file.

The third data file (**experim5ED.sav**) is a manufactured (fake) data file, constructed and manipulated to illustrate the use of a number of techniques covered in Part Five of the book (e.g. Paired Samples t-test, Repeated Measures ANOVA). This file also includes additional variables that will allow you to practise the skills learnt throughout the book. Just don't get too excited about the results you obtain and attempt to replicate them in your own research!

The fourth file used in the examples in the book is **depress5ED.sav**. This is used in [Chapter 16](#), on non-parametric techniques, to illustrate some techniques used in health and medical research.

Two other data files have been included, giving you the opportunity to complete some additional activities with data from different discipline areas. The **sleep5ED.sav** file is a real data file from a study conducted to explore the prevalence and impact of sleep problems on aspects of people's lives. The **staffsurvey5ED.sav** file comes from a staff satisfaction survey conducted for a large national educational institution.

See the Appendix for further details of these files (and associated materials). Apart from the data files, the *SPSS Survival Manual* website also contains a number of useful items for students and instructors, including:

- guidelines for preparing a research report
- practice exercises
- updates on changes to IBM SPSS as new versions are released
- useful links to other websites
- additional reading

Introduction and overview

This book is designed for students completing research design and statistics courses and for those involved in planning and executing research of their own. Hopefully this guide will give you the confidence to tackle statistical analyses calmly and sensibly, or at least without too much stress!

Many of the problems that students experience with statistical analysis are due to anxiety and confusion from dealing with strange jargon, complex underlying theories and too many choices. Unfortunately, most statistics courses and textbooks encourage both of these sensations! In this book we try to translate statistics into a language that can be more easily understood and digested.

The *SPSS Survival Manual* is presented in a structured format, setting out step by step what you need to do to prepare and analyse your data. Think of your data as the raw ingredients in a recipe. You can choose to cook your 'ingredients' in different ways—a first course, main course, dessert. Depending on what ingredients you have available, different options may, or may not, be suitable. (There is no point planning to make beef stroganoff if all you have is chicken.) Planning and preparation are an important part of the process (both in cooking and in data analysis). Some things you will need to consider are:

- Do you have the correct ingredients in the right amounts?
- What preparation is needed to get the ingredients ready to cook?
- What type of cooking approach will you use (boil, bake, stir-fry)?
- Do you have a picture in your mind of how the end result (e.g. chocolate cake) is supposed to look?
- How will you tell when it is cooked?
- Once it is cooked, how should you serve it so that it looks appetising?

The same questions apply equally well to the process of analysing your data. You must plan your experiment or survey so that it provides the information you need, in the correct format. You must prepare your data file properly and enter your data carefully. You should have a clear idea of your research questions and how you might go about addressing them. You need to know what statistical techniques are available, what sort of variables are suitable and what are not. You must be able to perform your chosen statistical technique (e.g. t-test) correctly and interpret the output. Finally, you need to relate this 'output' back to your original research question and know how to present this in your report (or in cooking terms, should you serve your chocolate cake with cream or ice-cream, perhaps some berries and a sprinkle of icing sugar on top?).

In both cooking and data analysis, you can't just throw all your ingredients in together, shove it in the oven (or IBM SPSS, as the case may be) and hope for the best. Hopefully this book will help you understand the data analysis process a little better and give you the confidence and skills to be a better 'cook'.

STRUCTURE OF THIS BOOK

This *SPSS Survival Manual* consists of 22 chapters, covering the research process from designing a study through to the analysis of the data and presentation of the results. It is broken into five main parts. Part One (Getting started) covers the preliminaries: designing a study, preparing a codebook and becoming familiar with IBM SPSS. In Part Two (Preparing the data file) you will be shown how to prepare a data file, enter your data and check for errors. Preliminary analyses are covered in Part

Three, which includes chapters on the use of descriptive statistics and graphs; the manipulation of data; and the procedures for checking the reliability of scales. You will also be guided, step by step, through the sometimes difficult task of choosing which statistical technique is suitable for your data.

In Part Four the major statistical techniques that can be used to explore relationships are presented (e.g. correlation, partial correlation, multiple regression, logistic regression and factor analysis). These chapters summarise the purpose of each technique, the underlying assumptions, how to obtain results, how to interpret the output, and how to present these results in your thesis or report.

Part Five discusses the statistical techniques that can be used to compare groups. These include non-parametric techniques, t-tests, analysis of variance, multivariate analysis of variance and analysis of covariance.

USING THIS BOOK

To use this book effectively as a guide to IBM SPSS, you need some basic computer skills. In the instructions and examples provided throughout the text I assume that you are already familiar with using a personal computer, particularly the Windows functions. I have listed below some of the skills you will need. Seek help if you have difficulty with any of these operations. You will need to be able to:

- use the Windows drop-down menus
- use the left and right buttons on the mouse
- use the click and drag technique for highlighting text
- minimise and maximise windows
- start and exit programs from the Start menu or from Windows Explorer
- move between programs that are running simultaneously
- open, save, rename, move and close files
- work with more than one file at a time, and move between files that are open
- use Windows Explorer to copy files from a memory stick to the hard drive, and back again
- use Windows Explorer to create folders and to move files between folders.

This book is not designed to 'stand alone'. It is assumed that you have been exposed to the fundamentals of statistics and have access to a statistics text. It is important that you understand some of what goes on 'below the surface' when using IBM SPSS. It is an enormously powerful data analysis package that can handle very complex statistical procedures. This manual does not attempt to cover all the different statistical techniques available in the program. Only the most commonly used statistical techniques are covered. It is designed to get you started and to develop your confidence in using the program.

Depending on your research questions and your data, it may be necessary to tackle some of the more complex analyses available in IBM SPSS. There are many good books available covering the various statistical techniques in more detail. Read as widely as you can. Browse the shelves in your library, look for books that explain statistics in a language that you understand (well, at least some of it anyway!). Collect this material together to form a resource to be used throughout your statistics classes and your research project. It is also useful to collect examples of journal articles where statistical analyses are explained and results are presented. You can use these as models for your final write-up.

The *SPSS Survival Manual* is suitable for use as both an in-class text, where you have an instructor taking you through the various aspects of the research process, and as a self-instruction book for those conducting an individual research project. If you are teaching yourself, be sure to actually practise using IBM SPSS by analysing the data that is included on the website accompanying this book (see

viii for details). The best way to learn is by actually doing, rather than just reading. 'Play' with the data files from which the examples in the book are taken before you start using your own data files. This will improve your confidence and also allow you to check that you are performing the analyses correctly.

Sometimes you may find that the output you obtain is different from that presented in the book. This is likely to occur if you are using a different version of IBM SPSS from that used throughout this book (IBM SPSS Statistics v21). IBM SPSS regularly updates its products, which is great in terms of improving the program, but it can lead to confusion for students who find that what is on the screen differs from what is in the book. Usually the difference is not too dramatic, so stay calm and play detective. The information may be there, but just in a different form. For information on changes to the IBM SPSS products you may like to go to the IBM SPSS website (www.spss.com).

RESEARCH TIPS

If you are using this book to guide you through your own research project, there are a few additional tips I would like to recommend.

- **Plan your project carefully.** Draw on existing theories and research to guide the design of your project. Know what you are trying to achieve and why.
- **Think ahead.** Anticipate potential problems and hiccups—every project has them! Know what statistics you intend to employ and use this information to guide the formulation of data collection materials. Make sure that you will have the right sort of data to use when you are ready to do your statistical analyses.
- **Get organised.** Keep careful notes of all relevant research, references etc. Work out an effective filing system for the mountain of journal articles you will acquire and, later on, the output from IBM SPSS. It is easy to become disorganised, overwhelmed and confused.
- **Keep good records.** When using IBM SPSS to conduct your analyses, keep careful records of what you do. I recommend to all my students that they buy a spiral-bound exercise book to record every session they spend on IBM SPSS. You should record the date, new variables you create, analyses you perform and the names of the files where you have saved the output. If you have a problem or something goes horribly wrong with your data file, this information can be used by your supervisor to help rescue you!
- **Stay calm!** If this is your first exposure to IBM SPSS and data analysis, there may be times when you feel yourself becoming overwhelmed. Take some deep breaths and use some positive self-talk. Just take things step by step—give yourself permission to make mistakes and become confused sometimes. If it all gets too much then stop, take a walk and clear your head before you tackle it again. Most students find IBM SPSS quite easy to use, once they get the hang of it. Like learning any new skill, you just need to get past that first feeling of confusion and lack of confidence.
- **Give yourself plenty of time.** The research process, particularly the data entry and data analysis stages, always takes longer than expected, so allow plenty of time for this.
- **Work with a friend.** Make use of other students for emotional and practical support during the data analysis process. Social support is a great buffer against stress!

ADDITIONAL RESOURCES

There are a number of different topic areas covered throughout this book, from the initial design of a study, questionnaire construction, basic statistical techniques (t-tests, correlation), through to advanced statistics (multivariate analysis of variance, factor analysis). Further reading and resources

material is recommended throughout the different chapters in the book. You should try to read as broadly as you can, particularly if tackling some of the more complex statistical procedures.



PART ONE

Getting started

Data analysis is only one part of the research process. Before you can use IBM SPSS to analyse your data, there are a number of things that need to happen. First, you have to design your study and choose appropriate data collection instruments. Once you have conducted your study, the information obtained must be prepared for entry into IBM SPSS (using something called a 'codebook'). To enter the data you must understand how IBM SPSS works and how to talk to it appropriately. Each of these

steps is discussed in Part One.

~~Chapter 1~~ provides some tips and suggestions for designing a study, with the aim of obtaining good quality data. ~~Chapter 2~~ covers the preparation of a codebook to translate the information obtained from your study into a format suitable for IBM SPSS. ~~Chapter 3~~ takes you on a guided tour of the program, and some of the basic skills that you will need are discussed. If this is your first time using IBM SPSS, it is important that you read the material presented in ~~Chapter 3~~ before attempting any of the analyses presented later in the book.



1

Designing a study

Although it might seem a bit strange to discuss research design in a book on IBM SPSS, it is an essential part of the research process that has implications for the quality of the data collected and analysed. The data you enter must come from somewhere—responses to a questionnaire, information collected from interviews, coded observations of actual behaviour, or objective measurements of output or performance. The data are only as good as the instrument that you used to collect them and the research framework that guided their collection.

In this chapter a number of aspects of the research process are discussed that have an impact on the potential quality of the data. First, the overall design of the study is considered; this is followed by a discussion of some of the issues to consider when choosing scales and measures; and finally, some guidelines for preparing a questionnaire are presented.

PLANNING THE STUDY

Good research depends on the careful planning and execution of the study. There are many excellent

books written on the topic of research design to help you with this process—from a review of the literature, formulation of hypotheses, choice of study design, selection and allocation of participants, recording of observations and collection of data. Decisions made at each of these stages can affect the quality of the data you have to analyse and the way you address your research questions. In designing your own study I would recommend that you take your time working through the design process to make it the best study that you can produce. Reading a variety of texts on the topic will help. A few good, easy-to-follow titles are listed in the recommended reading section at the back of the book. To get you started, consider these tips when designing your study:

- Consider what type of research design (e.g. experiment, survey, observation) is the best way to address your research question. There are advantages and disadvantages to all types of research approaches; choose the most appropriate approach for your particular research question. Have a good understanding of the research that has already been conducted in your topic area.
- If you choose to use an experiment, decide whether a between-groups design (different cases in each experimental condition) or a repeated measures design (same cases tested under different conditions) is the more appropriate for your research question. There are advantages and disadvantages to each approach, so weigh up each approach carefully.
- In experimental studies, make sure you include enough levels in your independent variable. Using only two levels (or groups) means fewer participants are required, but it limits the conclusions that you can draw. Is a control group necessary or desirable? Will the lack of control group limit the conclusions that you can draw?
- Always select more participants than you need, particularly if you are using a sample of humans. People are notoriously unreliable—they don't turn up when they are supposed to, they get sick, drop out and don't fill out questionnaires properly! So plan accordingly. Err on the side of pessimism rather than optimism.
- In experimental studies, check that you have enough participants in each of your groups (and try to keep them equal when possible). With small groups, it is difficult to detect statistically significant differences between groups (an issue of power, discussed in the introduction to Part Five). There are calculations you can perform to determine the sample size that you will need. See, for example, Stangor (2006), or consult other statistical texts under the heading 'power'.
- Wherever possible, randomly assign participants to each of your experimental conditions rather than using existing groups. This reduces the problem associated with non-equivalent groups in between-groups designs. Also worth considering is taking additional measurements on the groups to ensure that they don't differ substantially from one another. You may be able to statistically control for differences that you identify (e.g. using analysis of covariance).
- Choose appropriate dependent variables that are valid and reliable (see discussion on this point later in this chapter). It is a good idea to include a number of different measures—some measures are more sensitive than others. Don't put all your eggs in one basket.
- Try to anticipate the possible influence of extraneous or confounding variables. These are variables that could provide an alternative explanation for your results. Sometimes they are hard to spot when you are immersed in designing the study yourself. Always have someone else (supervisor, fellow researcher) check over your design before conducting the study. Do whatever you can to control for these potential confounding variables. Knowing your topic area well can also help you identify possible confounding variables. If there are additional variables that you cannot control, can you measure them? By measuring them, you may be able to control for them statistically (e.g. using analysis of covariance).

- If you are distributing a survey, pilot-test it first to ensure that the instructions, questions and scale items are clear. Wherever possible, pilot-test on the same type of people who will be used in the main study (e.g. adolescents, unemployed youth, prison inmates). You need to ensure that your respondents can understand the survey or questionnaire items and respond appropriately. Pilot-testing should also pick up any questions or items that may offend potential respondents.
- If you are conducting an experiment, it is a good idea to have a full dress rehearsal and pilot-test both the experimental manipulation and the dependent measures you intend to use. If you are using equipment, make sure it works properly. If you are using different experimenters or interviewers, make sure they are properly trained and know what to do. If different observers are required to rate behaviours, make sure they know how to appropriately code what they see. Have a practice run and check for inter-rater reliability (i.e. how consistent scores are from different raters). Pilot-testing of the procedures and measures helps you identify anything that might go wrong on the day and any additional contaminating factors that might influence the results. Some of these you may not be able to predict (e.g. workers doing noisy construction work just outside the lab's window), but try to control those factors that you can.

CHOOSING APPROPRIATE SCALES AND MEASURES

There are many different ways of collecting 'data', depending on the nature of your research. This might involve measuring output or performance on some objective criteria, or rating behaviour according to a set of specified criteria. It might also involve the use of scales that have been designed to 'operationalise' some underlying construct or attribute that is not directly measurable (e.g. self-esteem). There are many thousands of validated scales that can be used in research. Finding the right one for your purpose is sometimes difficult. A thorough review of the literature in your topic area is the first place to start. What measures have been used by other researchers in the area? Sometimes the actual items that make up the scales are included in the appendix to a journal article; otherwise you may need to trace back to the original article describing the design and validation of the scale you are interested in. Some scales have been copyrighted, meaning that to use them you need to purchase 'official' copies from the publisher. Other scales, which have been published in their entirety in journal articles, are considered to be 'in the public domain', meaning that they can be used by researchers without charge. It is very important, however, to properly acknowledge each of the scales you use, giving full reference details.

In choosing appropriate scales there are two characteristics that you need to be aware of: reliability and validity. Both of these factors can influence the quality of the data you obtain. When reviewing possible scales to use, you should collect information on the reliability and validity of each of the scales. You will need this information for the 'Method' section of your research report. No matter how good the reports are concerning the reliability and validity of your scales, it is important to pilot-test them with your intended sample. Sometimes scales are reliable with some groups (e.g. adults with an English-speaking background), but are totally unreliable when used with other groups (e.g. children from non-English-speaking backgrounds).

Reliability

The reliability of a scale indicates how free it is from random error. Two frequently used indicators of a scale's reliability are test-retest reliability (also referred to as 'temporal stability') and internal consistency. The test-retest reliability of a scale is assessed by administering it to the same people on two different occasions, and calculating the correlation between the two scores obtained. High test-retest correlations indicate a more reliable scale. You need to take into account the nature of the construct that the scale is measuring when considering this type of reliability. A scale designed to

measure current mood states is not likely to remain stable over a period of a few weeks. The test-retest reliability of a mood scale, therefore, is likely to be low. You would, however, hope that measures of stable personality characteristics would stay much the same, showing quite high test-retest correlations.

The second aspect of reliability that can be assessed is internal consistency. This is the degree to which the items that make up the scale are all measuring the same underlying attribute (i.e. the extent to which the items 'hang together'). Internal consistency can be measured in a number of ways. The most commonly used statistic is Cronbach's coefficient alpha (available using IBM SPSS, see [Chapter 9](#)). This statistic provides an indication of the average correlation among all of the items that make up the scale. Values range from 0 to 1, with higher values indicating greater reliability.

While different levels of reliability are required, depending on the nature and purpose of the scale, Nunnally (1978) recommends a minimum level of .7. Cronbach alpha values are dependent on the number of items in the scale. When there are a small number of items in the scale (fewer than 10) Cronbach alpha values can be quite small. In this situation it may be better to calculate and report the mean inter-item correlation for the items. Optimal mean inter-item correlation values range from .2 to .4 (as recommended by Briggs & Cheek 1986).

Validity

The validity of a scale refers to the degree to which it measures what it is supposed to measure. Unfortunately, there is no one clear-cut indicator of a scale's validity. The validation of a scale involves the collection of empirical evidence concerning its use. The main types of validity you will see discussed are content validity, criterion validity and construct validity.

Content validity refers to the adequacy with which a measure or scale has sampled from the intended universe or domain of content. Criterion validity concerns the relationship between scale scores and some specified, measurable criterion. Construct validity involves testing a scale not against a single criterion but in terms of theoretically derived hypotheses concerning the nature of the underlying variable or construct. The construct validity is explored by investigating its relationship with other constructs, both related (convergent validity) and unrelated (discriminant validity). An easy-to-follow summary of the various types of validity is provided in Streiner and Norman (2008).

If you intend to use scales in your research, it would be a good idea to read further on this topic: see Kline (2005) for information on psychological tests, and Streiner and Norman (2008) for health measurement scales. Bowling also has some great books on health and medical scales.

PREPARING A QUESTIONNAIRE

In many studies it is necessary to collect information from your participants or respondents. This may involve obtaining demographic information from participants prior to exposing them to some form of experimental manipulation. Alternatively, it may involve the design of an extensive survey to be distributed to a selected sample of the population. A poorly planned and designed questionnaire will not give good data with which to address your research questions. In preparing a questionnaire, you must consider how you intend to use the information; you must know what statistics you intend to use. Depending on the statistical technique you have in mind, you may need to ask the question in a particular way, or provide different response formats. Some of the factors you need to consider in the design and construction of a questionnaire are outlined in the sections that follow.

This section only briefly skims the surface of questionnaire design, so I would suggest that you read further on the topic if you are designing your own study. A really great book for this purpose is DeVellis & Vaus (2002).

Question types

Most questions can be classified into two groups: closed or open-ended. A closed question involves offering respondents a number of defined response choices. They are asked to mark their response using a tick, cross, circle, etc. The choices may be a simple Yes/No, Male/Female, or may involve a range of different choices. For example:

What is the highest level of education you have completed (please tick)?

- 1. Primary school
- 2. Some secondary school
- 3. Completed secondary school
- 4. Trade training
- 5. Undergraduate university
- 6. Postgraduate university

Closed questions are usually quite easy to convert to the numerical format required for IBM SPSS. For example, Yes can be coded as a 1, No can be coded as a 2; Males as 1, Females as 2. In the education question shown above, the number corresponding to the response ticked by the respondent would be entered. For example, if the respondent ticked Undergraduate university, this would be coded as a 5. Numbering each of the possible responses helps with the coding process. For data entry purposes, decide on a convention for the numbering (e.g. in order across the page, and then down), and stick with it throughout the questionnaire.

Sometimes you cannot guess all the possible responses that respondents might make—it is therefore necessary to use open-ended questions. The advantage here is that respondents have the freedom to respond in their own way, not restricted to the choices provided by the researcher. For example:

What is the major source of stress in your life at the moment?

Responses to open-ended questions can be summarised into a number of different categories for entry into IBM SPSS. These categories are usually identified after looking through the range of responses actually received from the respondents. Some possibilities could also be raised from an understanding of previous research in the area. Each of these response categories is assigned a number (e.g. work=1, finances=2, relationships=3), and this number is entered into IBM SPSS. More details on this are provided in the section on preparing a codebook in [Chapter 2](#).

Sometimes a combination of both closed and open-ended questions works best. This involves providing respondents with a number of defined responses, and also an additional category (other) that they can tick if the response they wish to give is not listed. A line or two is provided so that they can write the response they wish to give. This combination of closed and open-ended questions is particularly useful in the early stages of research in an area, as it gives an indication of whether the defined response categories adequately cover all the responses that respondents wish to give.

Response format

In asking respondents a question, you also need to decide on a response format. The type of response format you choose can have implications when you come to do your statistical analysis. Some analyses (e.g. correlation) require scores that are continuous, from low through to high, with a wide range of scores. If you had asked respondents to indicate their age by giving them a category to tick (e.g. less than 30, between 31 and 50 and over 50), these data would not be suitable to use in

correlational analysis. So, if you intend to explore the correlation between age and, say, self-esteem you will need to ensure that you ask respondents for their actual age in years. Be warned though, some people don't like giving their exact age (e.g. women over 30!).

Try to provide as wide a choice of responses to your questions as possible. You can always condense things later if you need to (see [Chapter 8](#)). Don't just ask respondents whether they agree or disagree with a statement—use a Likert-type scale, which can range from strongly disagree to strongly agree:

strongly disagree 1 2 3 4 5 6 strongly agree

This type of response scale gives you a wider range of possible scores, and increases the statistical analyses that are available to you. You will need to make a decision concerning the number of response steps (e.g. 1 to 6) that you use. DeVellis (2003) has a good discussion concerning the advantages and disadvantages of different response scales. Whatever type of response format you choose, you must provide clear instructions. Do you want your respondents to tick a box, circle a number, make a mark on a line? For some respondents, this may be the first questionnaire that they have completed. Don't assume they know how to respond appropriately. Give clear instructions and provide an example if appropriate, and always pilot-test on the type of people that will make up your sample. Iron out any sources of confusion before distributing hundreds of your questionnaires. When designing your questions, always consider how a respondent might interpret the question and all the possible responses a person might want to make. For example, you may want to know whether people smoke or not. You might ask the question:

Do you smoke? (please tick) Yes No

In trialling this questionnaire, your respondent might ask whether you mean cigarettes, cigars or marijuana. Is knowing whether they smoke enough? Should you also find out how much they smoke (two or three cigarettes, versus two or three packs), and/or how often they smoke (every day or only on social occasions)? The message here is to consider each of your questions, what information they will give you and what information might be missing.

Wording the questions

There is a real art to designing clear, well-written questionnaire items. Although there are no clear-cut rules that can guide this process, there are some things you can do to improve the quality of your questions, and therefore your data. Try to avoid:

- long complex questions
- double negatives
- double-barrelled questions
- jargon or abbreviations
- culture-specific terms
- words with double meanings
- leading questions
- emotionally loaded words.

When appropriate, you should consider including a response category for 'Don't know' or 'Not applicable'. For further suggestions on writing questions, see De Vaus (2002) and Kline (2005).



2

Preparing a codebook

Before you can enter the information from your questionnaire, interviews or experiment into IBM SPSS, it is necessary to prepare a 'codebook'. This is a summary of the instructions you will use to convert the information obtained from each subject or case into a format that IBM SPSS can understand. The steps involved will be demonstrated in this chapter using a data file that was developed by a group of my graduate diploma students. A copy of the questionnaire, and the codebook that was developed for this questionnaire, can be found in the Appendix. The data file is provided on the website that accompanies this book. The provision of this material allows you to see the whole process, from questionnaire development through to the creation of the final data file ready for analysis. Although I have used a questionnaire to illustrate the steps involved in the development of a codebook, a similar process is also necessary in experimental studies, or when retrieving information from existing records (e.g. hospital medical records).

Preparing the codebook involves deciding (and documenting) how you will go about:

- defining and labelling each of the variables

- assigning numbers to each of the possible responses.

All this information should be recorded in a book or computer file. Keep this somewhere safe; there is nothing worse than coming back to a data file that you haven't used for a while and wondering what the abbreviations and numbers refer to.

In your codebook you should list all of the variables in your questionnaire, the abbreviated variable names that you will use in IBM SPSS and the way in which you will code the responses. In this chapter simplified examples are given to illustrate the various steps. In the first column of Table 2.1 you have the name of the variable (in English, rather than in computer talk). In the second column

Table 2.1 Example of a codebook

Variable	SPSS variable name	Coding instructions
Identification number	ID	Number assigned to each survey
Sex	Sex	1 = Males 2 = Females
Age	Age	Age in years
Marital status	Marital	1 = single 2 = steady relationship 3 = married for the first time 4 = remarried 5 = divorced/separated 6 = widowed
Optimism Scale items 1 to 6	op1 to op6	Enter the number circled from 1 (strongly disagree) to 5 (strongly agree)

Example of a codebook you write the abbreviated name for that variable that will appear in IBM SPSS (see conventions below), and in the third column you detail how you will code each of the responses obtained.

Variable names

Each question or item in your questionnaire must have a unique variable name. Some of these names will clearly identify the information (e.g. sex, age). Other questions, such as the items that make up a scale, may be identified using an abbreviation (e.g. op1, op2, op3 is used to identify the items that make up the Optimism Scale).

There are a number of conventions you must follow in assigning names to your variables in IBM SPSS. These are set out in the 'Rules for naming of variables' box. In earlier versions of IBM SPSS (prior to Version 12), you could use only eight characters for your variable names. The later versions of the program allow you longer variable names, but very long names can make the output rather hard to read so keep them as concise as possible.

Rules for naming of variables

Variable names:

- must be unique (i.e. each variable in a data set must have a different name)
- must begin with a letter (not a number)
- cannot include full stops, spaces or symbols (! , ? * “)
- cannot include words used as commands by SPSS (all, ne, eq, to, le, lt, by, or, gt, and, not, go, with)
- cannot exceed 64 characters.

The first variable in any data set should be ID—that is, a unique number that identifies each case. ~~Before beginning the data entry process, go through and assign a number to each of the questionnaires or data records. Write the number clearly on the front cover. Later, if you find an error in the data set, having the questionnaires or data records numbered allows you to check back and find where the error occurred.~~

CODING RESPONSES

Each response must be assigned a numerical code before it can be entered into IBM SPSS. Some of the information will already be in this format (e.g. age in years); other variables such as sex will need to be converted to numbers (e.g. 1=males, 2=females). If you have used numbers in your questions, label your responses (see, for example, the education question in [Chapter 1](#)), this is relatively straightforward. If not, decide on a convention and stick to it. For example, code the first listed response as 1, the second as 2 and so on across the page.

What is your current marital status? (please tick)			
<input type="checkbox"/> single	<input type="checkbox"/> in a relationship	<input type="checkbox"/> married	<input type="checkbox"/> divorced

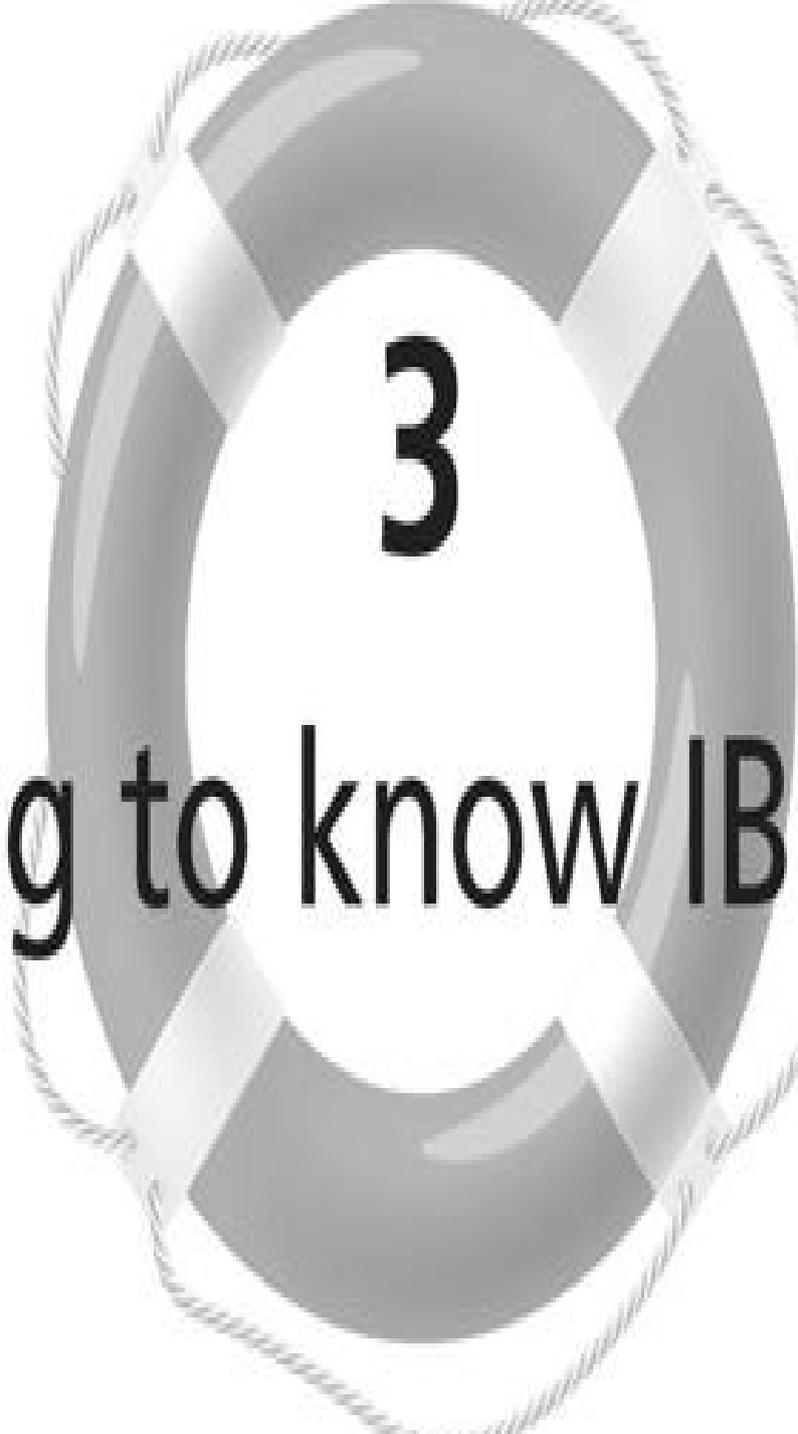
To code responses to the question above: if a person ticked single, they would be coded as 1; if in a relationship, they would be coded 2; if married, 3; and if divorced, 4.

CODING OPEN-ENDED QUESTIONS

For open-ended questions (where respondents can provide their own answers), coding is slightly more complicated. Take, for example, the question: What is the major source of stress in your life at the moment? To code responses to this, you will need to scan through the questionnaires and look for common themes. You might notice a lot of respondents listing their source of stress as related to work, finances, relationships, health or lack of time. In your codebook you list these major groups of responses under the variable name stress, and assign a number to each (work=1, spouse/partner=2 and so on). You also need to add another numerical code for responses that did not fall into these listed categories (other=99). When entering the data for each respondent, you compare his/her response with those listed in the codebook and enter the appropriate number into the data set under the variable stress.

Once you have drawn up your codebook, you are almost ready to enter your data. First you need to get to know IBM SPSS ([Chapter 3](#)), and then you need to set up a data file and enter your data ([Chapter 4](#)).

Getting to know IBM SPSS



3

Getting to know IBM SPSS

IBM SPSS operates using a number of different screens, or 'windows', designed to do different things. Before you can access these windows, you need to either open an existing data file or create one of your own. So, in this chapter we will cover how to open and close IBM SPSS; how to open and close existing data files; and how to create a data file from scratch. We will then go on to look at the different windows IBM SPSS uses.

STARTING IBM SPSS

There are a number of different ways to start IBM SPSS:

- The simplest way is to look for an SPSS icon on your desktop. Place your cursor on the icon and double-click.
- You can also start IBM SPSS by clicking on **Start**, move your cursor to **All Programs**, and then across to the list of programs available. See if you have a folder labelled IBM SPSS.

- **read [The Abduction](#)**
- **[download online Dutton's Orthopaedic Examination Evaluation and Intervention \(3rd Edition\)](#)**
- [download online Mortal Dictata \(Halo: Kilo-Five Trilogy, Book 3\) here](#)
- [read Proto-slavic Inflectional Morphology: A Comparative Handbook \(Brill's Studies in Indo-European Languages & Linguistics, Volume 14\) pdf, azw \(kindle\)](#)
- [download Making Meaning: Inference and Rhetoric in the Interpretation of Cinema \(Harvard Film Studies\)](#)

- <http://korplast.gr/lib/Effortless-E-Commerce-with-PHP-and-MySQL--2nd-Edition-.pdf>
- <http://jaythebody.com/freebooks/The-Pursuit-of-the-Ideal--Jewish-Writings-of-Steven-Schwarzschild.pdf>
- <http://patrickvincitore.com/?ebooks/Seven-Dirty-Words--The-Life-and-Crimes-of-George-Carlin.pdf>
- <http://ramazotti.ru/library/Progress-in-Food-Preservation.pdf>
- <http://growingsomeroots.com/ebooks/Making-Meaning--Inference-and-Rhetoric-in-the-Interpretation-of-Cinema--Harvard-Film-Studies-.pdf>