JULIE PALLANT

SPSS

SURVIVAL MANUAL

A STEP BY STEP GUIDE TO DATA ANALYSIS USING IBM SPSS

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

Contents

PART ONE
Getting started

Data analysis is only one part of the research process. Before you can use IBM SPSS Statistics to analyse your data, there are several things that need to happen. First, you must design your study and choose appropriate data collection instruments. Once you have conducted your study, the information obtained needs to be prepared for entry into IBM SPSS Statistics using something called a 'codebook'. To enter the data you must understand how the program 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 Statistics. Chapter 3 takes you on a guided tour of the program, and some of the basic skills that you need are discussed. If this is your first time using the program, 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 Statistics, 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 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 various 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 to 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 all 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, and 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 need. See, e.g. Stangor (2006).
- 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 of 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 variety of 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 (e.g. 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 types 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 to 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 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 several different ways. The most commonly used statistic is Cronbach's coefficient alpha (available using IBM SPSS Statistics; 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 in the literature 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 (2015).

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 (2015) 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 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 De Vaus (2014).

Question types

Most questions can be classified into two groups: closed and open-ended. A closed question involves offering respondents a set of defined response choices. They are

asked to mark their response using a tick, cross, circle and so on. The choices may be a simple 'yes' or 'no', 'male' or '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 Statistics. For example, no can be coded as a 0, yes can be coded as a 1; 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 and are 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 different categories for entry into IBM SPSS Statistics. These categories are usually identified after scanning through the range of responses 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 value (e.g. work = 1, finances = 2, relationships = 3), and this number is entered into IBM SPSS Statistics. 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 set of defined responses and an additional category ('other') that they can tick if the response they wish to give is not listed. A line or two are 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 a correlational analysis. So, if you intend to explore the correlation between age and, say, self-esteem, you 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 (or 'collapse') 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 need to make a decision concerning the number of response steps (e.g. 1 to 6) that you use. DeVellis (2012) 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, provide an example if appropriate, and always pilot-test on the types of people that will make up your sample. Iron out any sources of confusion before distributing hundreds of your questionnaires. In 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 pilot-testing 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 (2014) and Kline (2005).

2

Preparing a codebook

Before you can enter the information from your questionnaire, interviews or experiment into IBM SPSS Statistics, 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 the program can understand. The steps involved are 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 Statistics 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 in the codebook (see Table 2.1) you write the abbreviated name for that variable that will appear in the data file (see conventions below), and in the second column you can describe this variable in more detail so you can identify it later.