7

Analysing data

Introduction • The shape of your data • The nature of data • Managing your data • Computer-based data management and analysis • The process of analysis • Analysing documents • Analysing interviews • Analysing observations • Analysing questionnaires • Interpretation • Summary • Exercises • Further reading

Introduction

We hope that you are reading this chapter well before you have finished collecting your research data. You are likely, after all, to begin analysing your data before you have collected them all, possibly starting as soon as you have some data to work on. Analysis is an ongoing process which may occur throughout your research, with earlier analysis informing later data collection. Research is, as we have said a number of times in this book, a messy business, and the stages and processes involved do not simply follow one after the other.

> You might like to refer back at this point to the section on **Getting a favour of possibilities** in Chapter 1.

You would probably be best advised to look through this chapter before you finally decide how you are going to focus your study, and what kinds of approaches and techniques you will apply. It makes sense to have some understanding of the kinds of data analysis you might engage in, and how the kinds of data you collect will affect and limit this, before you commit yourself to a particular project.

The purpose of this chapter, then, is to help you get your data into shape, and to suggest how you might go about analysing and interpreting it. We start from those unsure, initial feelings, which are so common to both novice and more experienced researchers, of having an overwhelming or chaotic collection of research data. By the time you have finished this chapter, however, we aim to have taken you to a position where you can begin to write up your results and conclusions.

The chapter is organized in terms of the following themes:

- **The shape of your data**. The condition which your research data are in, and the facilities you have available to analyse them.
- **The nature of data**. What research data are, the meaning of numbers and words.
- **Managing your data**. Coding, reducing and summarizing your raw data.
- **Computer-based data management and analysis**. Using software packages with quantitative and qualitative data.
- **The process of analysis**. Thinking about and planning your analysis.
- **Analysing documents**. How to make sense of your notes.
- **Analysing interviews**. How to make sense of your transcripts.
- **Analysing observations**. How to make sense of your records.
- **Analysing questionnaires**. How to make sense of your replies.
- **Interpretation**. How to understand and contextualize the results of your analyses.

Hint: If you feel traumatized or terrorized by the process of analysing the data you have collected, you might like to think of it as analogous to cooking. What and how you cook depends on your taste, skills and the resources you have available. You may like your food simple and freshly prepared, or carefully blended over a long period, or fast and processed. You may mix the ingredients together using a recipe, or based on previous experience, or you may buy a packet already prepared. You may use a range of tools in your cooking, from a simple knife or spoon through to an expensive food processor. You may be preparing food just for yourself or for a banquet. See if you can find further parallels as you cook your data!

The shape of your data

Two basic issues affecting your whole approach to data analysis will be considered in this section:

- the condition which the data you have collected are in;
- where, and with what facilities, you are able to analyse them.

Order or chaos?

You will probably spend a considerable amount of time collecting your research data, and – unless you are relying entirely on secondary data sources with which you are already familiar – the shape of the data collection which you end up with will almost certainly be rather different from the way you had envisaged it when you started. While your plans for data collection may have seemed very methodical, the data you have actually collected may initially appear to be anything but. They may seem more chaotic than ordered (see Box 7.1).

Whether your data appear ordered or chaotic depends in part on your preferences, and in part upon your perceptions: one person's chaos may be another's order. The real issue is what works well for you. So long as you know where to find what you want or need to find, that's OK. If you are new to the process of research, of course, you may be finding this out as you go along. There is no single 'right' strategy for carrying out research, nor for ordering and analysing data. Much of what is said in this book can be taken to indicate a preference for planning, structure and order, but these qualities may be conceived of very differently in practice.

The condition your data are in will undoubtedly, though, change during the process of analysis. However poor, ill-organized or inadequate you may think they are at the beginning, you are likely to find strengths in them as you proceed. Similarly, even if you start from the position that you have all the data you need, you are likely to recognize deficiencies as you get into the depths of analysis.

Data analysis is about moving from chaos to order, and from order to chaos. Data which seem under control are likely to become somewhat more disorganized, at least for a while; while some semblance of order will be found, or imposed upon, even the most chaotic collection. Your data may, at times during the process of analysis, appear to be both messy and structured. Areas where you think that your data add to an understanding of the topic you are

Box 7.1 Ordered or chaotic data?

Appearance of order	Appearance of chaos
Neat notebooks	Odd notes
Card indexes	Scraps of paper
Piles of questionnaires	Baskets of cuttings
Colour-coded folders	Bulging or empty files
Labelled, transcribed tapes	Jotted down quotes
Highlighted photocopies	Half-remembered references
Clear plan and schedule	Back of an envelope
Computer database	Illegible handwriting

researching may be seen as ordered, while areas in which your work has raised more questions than answers (the normal pattern) may appear as more chaotic.

Where to analyse, and with what?

The resources you have available for your research, and how you might tailor your research plans to them, have already been considered elsewhere in this book.

> You might like to have another look at the sections on **Choosing a topic** in Chapter 2, and **Using computers** in Chapter 5.

Obviously, you are restricted in how, where and when you carry out your data analysis by the available resources. There are, however, practical issues concerned with the place, space and time in which you do your analysis which are worth further consideration. For example:

- Do you prefer working at a desk or in an armchair?
- Will you want to spread your work over a floor or a wall?
- Do you like to work with paper and pen (or pencil)? Or straight onto a computer screen?
- Does your analysis require extensive dedicated periods of time, or can it be done in smaller chunks? Or are there elements of both?
- Can you do your analysis in one place, or will it require visits to a number of separate facilities?

Clearly, your answers to these and related questions will help to determine how you go about analysing your data. You will need to reconcile your preferences with what is feasible, and with the nature of the data you have collected.

The nature of data

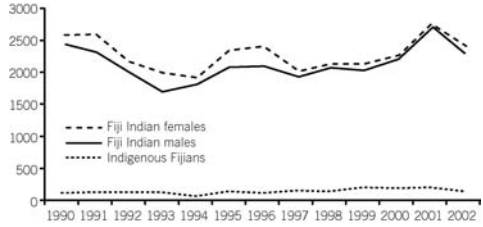
The data you have collected are likely to be in a number of forms, though it is perfectly possible to carry out interesting and valid research with just one form of data. Your data might include, for example, completed questionnaires, interview transcripts, notes on readings or observations, measurements of behaviour, Internet materials, charts, diagrams and photographs. Some may be in digital form. Now might be a good point to take some time to remind yourself about the nature of your data, the amount you have, where they have come from and how they have been produced.

Boxes 7.2 and 7.3 include a variety of examples of different sorts of quantitative and qualitative data to remind you of some of the possibilities.

Box 7.2 Examples of quantitative data

Order	*No. of responses*	*Factor*
1	113	Higher pay
2	73	Feeling valued by stakeholders in education
3	70	Desire to help children learn
4	64	Less administration
5	59	More non-contact time for planning and preparation
6	51	More support with pupil discipline issues
7	49	A reduction in overall work load
8	48	Good working relations with managers and other staff
9	41	Good prospects of career advancement
10	35	Smaller class sizes

(*Source*: Rhodes et al. 2004: 74)



(*Source*: Chandra 2004: 185)

Profile of Returnees

Start	Finish	University	Course	College in NZ?	Scholarship	Work in NZ	PR	Twin Program
1959	1963	VUW	BA/MA (Geog)	N	Y	N	N	N
1965	1966	VUW	Accounting	N	Y	N	N	N
1963	1967	Canterbury	BE/BSc/MA (Chemical Engineering)	N	Y	N	N	N
1964	1967	Canterbury	Economics & Agriculture	N	Y	N	N	N
1961	1967	Canterbury and Otago	B.Sc & PGDip Statistics	N	Y	N	N	N
1961	1969	Canterbury	B.E./1/2 M.E.	N	Y	N	N	N
1970	1973	VUW	BA (Hons:Psychology)	N	Y	N	N	N
1974	1977	VUW	BA (History and Ed)/Hons (Ed)	N	Y	N	N	N
1988	1991	VUW	BA (Eng & Pol)/Hons (Engl Lit.)	N	N	N	N	N

(*Source*: Butcher 2004: 280)

Factor Analysis of Predictors of Identification with the Employing Organization

Predictor	Factor Loadings			
	Organizational attributes	Relationship with management	Relationship with colleagues	Positive distinctiveness
Providing opportunities to creatively solve problems	*.86*	.18	.12	.23
Keeping up to date with changes in IT	*.82*	.15	.01	.24
Providing career advancement opportunities	*.81*	.18	.11	.28
Doing high-quality work	*.77*	.28	.00	.24
Providing a work environment that is free of politics	*.77*	.16	.11	.01
I trust that this person will advance my best interests when decisions which affect me are made.	.27	*.88*	.16	.23
I have trust and confidence in that X employee regarding his/her general fairness.	.29	*.88*	.11	.21
I feel free to discuss the problems and difficulties in my job with that X employee.	.18	*.87*	.15	.18

(*Source*: Chattopadhyay 2005: 69)

Box 7.3 Examples of qualitative data

Chapter 6

1 We recommend to the Government that it should have a long term strategic aim of responding to increased demand for higher education, much of which we expect to be at sub-degree level; and that to this end, the cap on full-time undergraduate places should be lifted over the next two to three years and the cap on full-time sub-degree places should be lifted immediately.

Chapter 7

2 We recommend to the Government and the Funding Bodies that, when allocating funds for the expansion of higher education, they give priority to those institutions which can demonstrate a commitment to widening participation, and have in place a participation strategy, a mechanism for monitoring progress, and provision for review by the governing body of achievement.

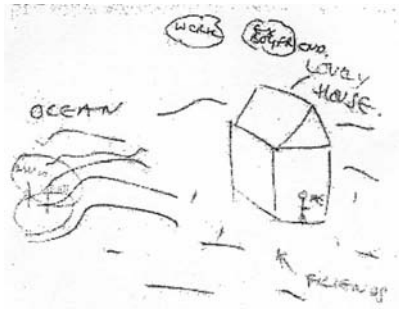
(*Source*: National Committee of Inquiry into Higher Education 1997: 42)

Cris: I remember playing dress up and I got to be the princess and you had to be the prince, you were the older sister and you had to be the prince.
Kathy: I remember that; that was a lot of fun. I remember I hated the way Mom used to always make me wear pink and you always got the blue dress.
Cris: I always hated dressing up like that anyway. Regardless of what it looked like.
Kathy: It was kind of cute.
Cris: It was so uncomfortable.

(*Source*: Davis and Salkin 2005)

Principle	Explanation	Foundational citations
1 Reformation of the professor–student relationship	A feminist pedagogy offers the professor and the students new relational roles. Individuals involved in the learning experience share knowledge and thus enact the teaching role as well as acquire knowledge and thus enact the learner role (Parry 1996).	Bowker and Dunkin 1992, Bell 1993, Bright 1993, Shrewsbury 1993, Foss and Griffin 1995, Christie 1997, Scering 1997, Stanovsky 1997
2 Empowerment	To empower a student is to enact 'a participatory, democratic process in which at least some power is shared' (Shrewsbury 1993: 9). The professor can acknowledge power as evaluator and grader, while also redefining the teaching role from knowledge leader to 'activation of multiple perspectives' (Scering 1997: 66).	Bright 1993, Shrewsbury 1993, Woodbridge 1994, Chapman 1997, Scering 1997, Middlecamp and Subramaniam 1999

(*Source*: Webb et al. 2004: 425)



(*Source*: Bagnoli 2004: 11)

The qualitative/quantitative divide

Among these different kinds of data we may recognize a basic distinction between the quantitative (i.e. numbers) and the qualitative (i.e. words and everything else). This distinction has a major influence on how data may be analysed, and also reflects the varied 'traditions', philosophies and practices of different social science disciplines or sub-disciplines. You are almost certain to have examples of both types among your data, though either the qualitative or the quantitative may predominate.

You may wish to refer back to the sections on **Which method is best?** and **Families, approaches and techniques** in Chapter 3.

However, the distinction between words and numbers is not as precise as it may appear to be at first sight. Both offer representations of what we as individuals perceive of as our 'reality'. It may be that qualitative data offer more detail about the subject under consideration, while quantitative data appear to provide more precision, but both give only a partial description. Neither are 'facts' in anything but a very subjective sense. The accuracy of the representation is also likely to be reduced further during the research process, as we attempt to summarize or draw out key points from the vastness of the data available.

The quantitative and qualitative also have a tendency to shade into each other, such that it is very rare to find reports of research which do not include both numbers and words. Qualitative data may be quantified, and quantitative data qualified. For example, it is common practice in analysing surveys to assign, sometimes arbitrarily, numerical values to qualitative data, such as, 'successful' (1), 'unsuccessful' (2).

Researchers who adopt an explicitly qualitative stance can find themselves giving prominence to numbers, and vice versa. Thus, if you conduct your research entirely through interviews, and analyse the results by searching for similarities and differences in the interview records, you are quite likely to end up using numbers or their written equivalents in your writing: e.g. 'all of the interviewees', 'most of the respondents', 'half of the women I spoke to', etc. Or, if you base your study wholly on numerical data, you will still introduce qualitative factors in your analysis, as in discussing the relative worth of different data sources, and in interpreting what your results mean for practice.

The next two sub-sections aim to make these points clearer. You may want to skip one or other of them if you are already familiar with quantitative or qualitative approaches.

What do numbers mean?

Exercise 7.1 asks you to examine the examples of quantitative data included in Box 7.2. Box 7.2 does not, of course, include examples of all the different kinds or uses of numbers which you might come across in the course of your research, but it does contain some of the most common. If you have carried out a survey or experiment as part of your research, you are quite likely to have produced figures not unlike some of them. These may include, for example:

- direct measurements, or what might be called 'raw' or 'real' numbers;
- categories, where responses have been coded or assigned a numerical value;
- percentages, a measure of proportion;
- averages, which summarize a series of measurements.

The second question posed in Exercise 7.1 highlights a key point about quantitative data (and data in general), namely that they might tell you a lot if you only knew how they were arrived at, and how to interpret them. Every data source needs to be interrogated as to its representativeness, reliability and accuracy. Researchers ideally need to know by whom they were produced, for what purpose and in what ways. Numbers, by their very seeming precision, can hide their manufacture, imprecision and subjectivity. These issues are considered further in the section later in this chapter on *Interpretation*.

The third question posed in the exercise indicates that, once you are presented with a set of quantitative data, you can usually start to do other, quantitative or qualitative things, with it. You may have found yourself calculating averages, or thinking that one item was bigger or smaller than another, or of the same value. If you have sufficient information, you can calculate percentages from raw data, or produce the raw data from the percentages reported.

The quantitative data presented, whether you have produced them yourself or obtained them from a secondary source, are usually only the starting point of the analysis. In carrying out an analysis, the researcher inevitably gets further and further away from the original or 'real' data, producing more and more highly refined abstractions. You need to be able to trace the routes taken in this process, whether they are your own or another researcher's.

What do words mean?

Exercise 7.2, in a way analogous to the previous one, asks you to examine the examples of qualitative data included in Box 7.3. As in the case of Box 7.2, Box 7.3 does not include examples of all possible forms or types of words (or other forms of representation). It is, obviously, restricted to English language sources for a start. Nevertheless, we can recognize from the examples given some of the most common forms of written data, including:

- directly written words, and spoken words which have been transcribed either directly or in the form in which they were reported;

- written notes, put together during, soon after or long after the events which they purport to describe;
- carefully considered written words, intended for publication and broad circulation, and those not published and meant for a smaller and more ephemeral distribution.

Within these examples, we can recognize different levels of abstraction. Notes clearly only offer a partial summary of events, focusing on those aspects which the person making the notes felt to be most significant at the time for their own purposes (and could remember). Documents are usually only released after a process of drafting and re-drafting, and may be as interesting for what they don't say as for what they do say, as well as for how they say it.

Even direct speech is selective, however, spoken only after the speaker has thought, for a longer or shorter period, about what they might say; and determined in part by what the speaker thinks the listener might want to hear. It also, when compared with written English, betrays the effects of improvization in its punctuation, stumblings, repetitions and pauses. Again, as with quantitative data, there is a need to interrogate the sources and ask where the speakers or writers are coming from and why.

This brief discussion suggests that some analysis has already occurred in all of the examples given. Anything which you, as a researcher, may do to data such as these in the course of your analysis will further refine and select from the words given. Thus, you may choose (as we have) particular quotes or phrases as significant or illustrative, and will almost certainly in the end effectively ignore the great bulk of the written texts available.

As you will probably have noted by now, the last two sub-sections, on the meanings of numbers and words, have been very similar in format and approach. We have taken this approach for two main reasons:

- the processes involved in analysing these forms of data are broadly analogous;
- neither form of data is intrinsically better, more accurate or 'real' – each has to be assessed, analysed and used on its own merits.

Having assembled your data for analysis, the next stage is normally a managerial one. This typically involves sorting, coding, reducing or summarizing the data from its original form, and getting it into a shape better suited to analysis and reportage. These techniques are the subject of the next section.

Managing your data

> You might well find yourself, 6 months before the end of your study, with an alpine collection of information that might require a week just to read over carefully. A chronic problem of qualitative research is that it is done chiefly with words, not with numbers. Words are fatter than numbers and usually have multiple meanings . . . Numbers are usually less ambiguous and can be processed more economically. Small wonder, then, that many researchers prefer working with numbers alone or getting the words they collected translated into numbers as quickly as possible. We argue that although words may be more unwieldy than numbers, they can render more meaning than numbers alone and should be hung onto throughout data analysis.
>
> (Miles and Huberman 1994: 56)

You will probably have collected a substantial amount of data for the purposes of your research project. But your data in their raw state do not constitute the results of your research. You would be unlikely, for example, to simply bind together transcripts of all the interviews you have undertaken, or of all the questionnaires you have had returned, or of all the notes you have taken, and present that as your report or dissertation. That would be too long and too demanding for your readers, and it would lack insight and significance. The business of analysing the data you have collected, therefore, really involves two closely related processes:

- managing your data, by reducing their size and scope, so that you can report upon them adequately and usefully;
- analysing your managed set of data, by abtracting from it and drawing attention to what you feel is of particular importance or significance.

The first of these processes is considered in this section and the next one, the second in the remainder of the chapter. Each process is essential to research.

> **Hint**: Some of the tasks involved in analysing data are very basic and repetitive. Save these for when you are unable to do, or do not feel like doing, anything more demanding.

You may choose to manage your data in a whole series of related ways. Some of these you will already be familiar with, whether you realize it or not. Thus, the techniques described in Chapter 4, **Reading for research**, are all about management, and are used by many social science researchers. Those described in Box 7.4 are analogous, and also overlap to a considerable extent. You will probably use all or them in your own analysis.

Box 7.4 Techniques for managing data

Coding. The process by which items or groups of data are assigned codes. These may be used to simplify and standardize the data for analytical purposes, as when characteristics like sex, marital status or occupation are replaced by numbers (e.g. replacing 'male' by '1', 'female' by '2'). Or the process may involve some reduction in the quantity of the data, as when ages, locations or attitudes are categorized into a limited number of groups, with each group then assigned its own numerical identity (e.g. categorizing ages as 'under 21', '21–64' and '65 and over', and then replacing these by '1', '2' and '3' respectively).

Annotating. The process by which written (or perhaps audio or visual) material is altered by the addition of notes or comments. On books or papers, these may take the forms of marginal notes, or underlining or highlighting the text itself. The process may draw attention to what you consider to be the more significant sections, perhaps for later abstraction and quotation. Or it may serve as part of your continuing debate with your texts, a means to refine and progress your ideas further.

Labelling. Where you have an analytical scheme in mind, or are developing one, you may go through materials such as interviews or policy documents and label passages or statements with significant words (e.g. 'mother', 'conservative', 'career break', 'introvert'). These labels can then serve to direct your further analysis. A fine distinction might be drawn between the related processes of labelling and annotation, in that labelling smacks of stereotyping, of having your ideas or prejudices worked out in advance, whereas annotating seems more open or flexible.

Selection. A key process in the management of data, through which interesting, significant, unusual or representative items are chosen to illustrate your arguments. This may take the form, for example, of one member of a group, one institution, one answer to a survey, one particular quotation, one text, or a number of such selections. The point is that you are choosing, for a variety of reasons, which examples of your data collection to emphasize and discuss. There is always a good deal of subjectivity involved in such a process.

Summary. The process where, rather than choose one or more examples from a larger body of data, you opt to produce a reduced version, precis or synopsis of the whole data set. This would probably aim to retain something of the variability of the original data collected, while saying something about the generality and/or typical cases.

(*Note*: All of these techniques may be carried out, for qualitative or quantitative data, using available software as well as manually. The names given here to the techniques are often used in inter-changeable ways.)

All of the techniques outlined in Box 7.4 – coding, annotating, labelling, selection, summary – may be applied to a range of types of data, both qualitative and quantitative. All of them also result, though perhaps not initially, in a reduction in the quantity of the data which you have available for analysis. This is essential if you are going to be able to carry out a manageable analysis. All are subjective to a greater or lesser degree, and all involve the loss of some information. Given the same data set, different researchers would proceed with its management in somewhat different ways, leading to different forms of analysis and different results. That is why, if you are involved in a group research project, it can be relatively easy to each submit a different report.

Computer-based data management and analysis

It may be that your research project is sufficiently small-scale for you not to need to use sophisticated, computer-based software packages to manage and analyse your data. Or you may have made a conscious choice not to use them: much analysis can, after all, be done manually, and you may prefer to do yours in this way. However, even if you have collected a relatively modest amount of data – say, a few dozen questionnaires, or half a dozen interviews, or the records of 20 observations – there is still much to be said for computer-based analysis. Once your data has been inputted into the computer, computer-based analysis is much quicker and more accurate than anything you might do manually.

If, then, you are considering using a software package to help you manage and analyse your data, you will find it worthwhile to explore the various possibilities before committing yourself. The sooner you start doing this, the better.

> **Hint**: Doing a research project provides you with a splendid opportunity to learn about what some of these software packages can do. It is much more difficult to learn about them in the abstract, without any real data or any real purpose for analysing them.

Software packages designed to carry out quantitative analysis are much better established than those for qualitative analysis. The most widely available quantitative package in social science departments in universities is probably SPSS (Statistical Package for the Social Sciences). There are, however, other common quantitative data analysis packages, such as MINITAB; while many spreadsheet and database packages also support the simpler forms of quantitative analysis.

SPSS enables you to input raw data, to modify and re-organize them once they have been inputted, and to carry out a wide range of simple, statistical

and multivariate analyses. These range from listing the frequencies of different responses and calculating means, through cross-tabulation, correlation and regression analyses, analyses of variance and covariance, to cluster and factor analysis. In the UK, the Economic and Social Research Council (ESRC) has a web site (http://tramss.data-archive.ac.uk), which offers online training for those interested in statistical data analysis.

See also the section later in this Chapter on **Analysing questionnaires**.

If your data are primarily qualitative, the choice of a software package to manage and analyse it may not seem so straightforward. Box 7.5 outlines some

Box 7.5 Issues to consider when choosing data management packages for qualitative research

- What kind(s) and amount of data do you have, and how do you want to handle it?
- What is your preferred style of working?
- What is your theoretical approach to analysis and how well developed is it at the outset?
- Do you have a well defined methodology?
- Do you want a simple to use software which will mainly help you manage *your* thinking and *thematic* coding?
- Are you more concerned with the *language*, the terminology used in the data, the comparison and occurrence of words and phrases across cases or between different variables?
- Do you wish to consider tools which offer suggestions for coding, using *Artificial Intelligence* devices?
- Do you want both thematic and quantitative content information from the data?
- Do you want a multiplicity of tools (not quite so simple) enabling many ways of handling and interrogating data?
- How much time do you have to 'learn' the software?
- How much analysis time has been built into the project?
- Are you working individually on the project or as part of a team?
- Is this just one phase of a larger project – do you already have quantitative data?
- Is there a package – and peer support – already available at your institution or place of work?

(Lewins and Silver 2005: 5)

of the questions to bear in mind. Another ESRC web site, CAQDAS (http://caqdas.soc.surrey.ac.uk) provides a key site for information and knowledge about the software (e.g. ATLAS.ti, NVivo) that is being developed to facilitate qualitative data analysis, with links to software developers and demonstrations. CAQDAS also provides transcription guidelines.

The process of analysis

Analysis can be a fearful word for the novice small-scale researcher. You will probably have started your research project with some pre-conceptions of what you would find. You have now collected a great deal of data on your topic, and have got them into a shape for analysis. But how do you get from the vast array of words and numbers that you have collected or produced to a seemingly neat set of conclusions or recommendations? What is this process called analysis? Exercise 7.3 asks you to think about its nature and meaning.

In doing Exercise 7.3, did you refer to terms like concept, explanation, theory and understanding? These, together with synonymous and related terms, are at the heart of the process of analysis. Put simply:

- *concepts* are abstract or general ideas, which are important to how we think about particular subjects or issues;
- *theories* are suppositions which explain, or seek to explain something;
- *explanations* are statements which seek to make something intelligible, about why things are the way they are;
- *understanding* is our perception of the meaning of something, in this case the subject area, the issues and/or the research questions under consideration.

Analysis is about the search for explanation and understanding, in the course of which concepts and theories will likely be advanced, considered and developed. You will find a great deal, and much more detailed, discussion of these and related ideas in some of the books listed in the **Further reading** section at the end of this chapter.

The next four sections pick up the discussion from Chapter 6, **Collecting data**, by focusing on the analysis of the four main techniques for data collection identified:

- documents
- interviews

- observations
- questionnaires

The examples and strategies discussed in these sections may seem to suggest that certain of these techniques are either predominantly qualitative or predominantly quantitative. It should be stressed, however, that each of these techniques may be applied to deal with either quantitative or qualitative data or both.

Remember: Analysis is meant to be a rigorous process, using data that has been carefully produced and managed. In the end, however, what you produce from it is your own 'document', an attempt to persuade your readers of your own interpretation.

Analysing documents

> Documents . . . do not simply reflect, but also construct social reality and versions of events. The search for documents' 'meaning' continues, but with researchers also exercising 'suspicion'. It is not then assumed that documents are neutral artefacts which independently report social reality, or that analysis must be rooted in that nebulous concept, commonsense reasoning. Documents are now viewed as media through which social power is expressed. They are approached in terms of the cultural context in which they were written and may be viewed 'as attempts at persuasion'.
>
> (May 2001: 183)

As we have already indicated, documentary analysis is akin to the processes gone through in reading for research purposes. These are discussed at some length elsewhere in this book.

See Chapter 4, especially the section on **Good enough reading** and Box 4.9; the section in Chapter 6 on **Documents**; and Chapter 8, particularly the section on **How to criticize**.

Documentary analysis involves the careful consideration of a range of related questions. These have been summarized in Box 7.6. Some examples of the process of analysing documents are given in Box 7.7 (page 209).

Box 7.6 Issues in documentary analysis

For each document you are analysing, ask yourself:

- Who is the author?
- What is their position?
- What are their biases?
- Where and when was the document produced?
- Why was the document produced?
- How was it produced? who for?
- In what context was the document produced?
- What are its underlying assumptions?
- What does the document say, and not say?
- How is the argument presented?
- How well supported and convincing is its argument?
- How does this document relate to previous ones?
- How does this document relate to later ones?
- What do other sources have to say about it?

Two key points come out the list of issues in Box 7.6, and the examples given in Box 7.7:

- Documents, whatever their nature (statistics or words, official or unofficial, public or private), cannot be taken at face value. They are artificial and partial accounts, which need to be critically assessed for research purposes.
- Much of the significance and interest in documents is revealed when they are considered in relation to each other. We develop our understanding of the ideas, issues and policies with which documents deal through a comparative analysis.

If you doubt these points, try Exercise 7.4.

Documentary analysis proceeds by abstracting from each document those elements which we consider to be important or relevant, and by grouping together these findings, or setting them alongside others which we believe to be related. What you see or read in documents will be a product of your viewpoint, discipline or focus.

Like social research in general, documentary analysis may proceed by quantitative or qualitative means, or a combination of both. A common quantitative approach is content analysis, which is outlined in Box 7.8 (page 210). An alternative, qualitative strategy, discourse analysis – commonly also used for the analysis of conversations and interviews – is outlined in the next section in Box 7.10.

Box 7.7 Examples of documentary analysis

> The original analysis of Summerland [the fire at the Summerland Leisure Centre in the Isle of Man] was based upon data drawn entirely from the official public inquiry into the incident . . . I worked paragraph by paragraph through this report, as I did for all of the accident reports published between 1965 and 1975. I asked, for each paragraph, what names or 'labels for ideas' I needed to identify those elements, events or notions which were of interest to me in my broad and initially very unfocused concern to develop a theory of disaster preconditions. I then recorded each name or concept label on the top of a 5″ by 8″ file card, together with a note of the source paragraph, and added further paragraph references to the card as I encountered additional instances of the concept identified. Eventually for my whole study I ended up with 182 of these cards, which had to be sifted, sorted and juggled into a coherent theoretical model. I produced general definitions for each of the categories which recurred, looking for causal and other links and moved gradually towards a theoretical pattern which helped to explain the range of data which I had about accidents.
>
> (Turner 1994: 198)

Tight (2003) studied 406 articles published in 17 specialist higher education journals in the year 2000, and 284 books that were in print in that year, restricting his sample to non-North American sources and to English language publications. These materials were analysed in terms of the themes or issues they addressed (eight key themes were identified), the methods and/or methodologies they employed (again eight categories), the levels on which they focused (seven categories), and the characteristics of their authors.

Analysing interviews

> Working with a long transcript – on average of 60 pages – and various pages of field notes is not an easy task . . . In addition to the volume of data produced, focus group transcripts have multiple meanings and several different interpretations. Moreover, researchers have different assumptions and principles of analysis – about systematicity, verification, accessibility and so on. It is therefore important that the analysis is as focused as possible: key or primary questions are of the utmost importance for

Box 7.8 Content analysis

. . . *content* analysis . . . comprises three stages: stating the research problem, retrieving the text and employing sampling methods and interpretation and analysis. This focus considers the frequency with which certain words or particular phrases occur in the text as a means of identifying its characteristics . . . Words or phrases in the document are transformed into numbers. The number of times in which a word occurs in the text is taken as an indicator of its significance . . .

In considering the problems of a quantitative count . . . [five issues are raised]. First, this method considers product and says little of process . . . Second, an empiricist problem is raised for it deals only with information which can be measured and standardized and for this reason considers only data which can be simplified into categories. Third, in this preoccupation, it reproduces the meanings used by authors in the first instance, as opposed to subjecting them to critical analysis . . . Fourth, from an ethnomethodological perspective, it fails to understand the common-sense context of their production and interpretation as part of the methods by which people make sense of their social world. Fifth, it assumes that the audiences who receive the methods must translate it as the analyst does. By default, it therefore negates the idea that a text is open to a number of possible readings . . .

(May 2001: 191–2)

analysis, some questions do not deserve analysis at the same level, while others may be eliminated, as they simply set the background for discussion.

(Litosseliti 2003: 91)

Two examples of the process of analysing interviews are given in Box 7.9. These two examples usefully illustrate both some of the different approaches possible, and some of the commonalities, in the analysis of interview data.

The first example, involving a more experienced researcher, applied an established methodology developed within the phenomenological tradition, explicitation, organized in five stages. This could be carried out by individual researchers, alone, comparatively or together. Like most methods of qualitative data analysis, explicitation works by steadily extracting from the data collected a series of themes.

The second example in Box 7.9 involved a novice researcher who was not consciously following any particular approach to the analysis of the data she had collected. Nevertheless, the account of her analysis shows strong similarities to the other example. The examination of interview transcripts question

Box 7.9 Examples of interview analysis

Explicitation of the data:
This explicitation process has five 'steps' or phases, which are:

1 Bracketing and phenomenological reduction.
2 Delineating units of meaning.
3 Clustering of units of meaning to form themes.
4 Summarising each interview, validating and where necessary modifying it.
5 Extracting general and unique themes from all the interviews and making a composite summary.

(Groenewald 2004)

An orderly person spreads out his or her interview records in the garden:
Hester was working on the records of the interviews she had carried out with a sample of students. Each record contained her typed up shorthand notes made during the interview, and a summary of the student's background. Each consisted of several pages, including direct quotations. She first went through the interview notes, analysing them 'question by question'. This meant having all of the records spread out at once. She wanted her analysis to be both 'professional' and 'scientific', without losing the personal touch. She preferred an orderly approach: 'I tried breaking up all of the scripts, question by question. I sat with the scripts and got out my pad, and went through each script and each question and noted down the similarities and dissimilarities. First of all I looked for common themes, and then I went through each script again noting which themes had come up.'

by question, and the comparison of the answers to specific questions given by a range of interviewees, is also analogous to the general approach to documentary analysis outlined in the previous section.

There are, of course, other approaches to the analysis of interviews. You may not produce a transcript, but analyse the recordings direct. You may not have recorded the interviews, but be working from your notes. You may input your data to a computer and use a software program for analysis (see the previous section in this chapter on **Computer-based data management and analysis**). You might use a particular strategy for your analysis, such as explicitation or critical discourse analysis (see Box 7.10 – discourse analysis might also be employed in analysing documents). The process of looking for significant statements, and comparing what was said in different interviews, will, however, be similar.

Box 7.10 Critical discourse analysis

- Views a prevailing social order as historically situated and therefore relative, socially constructed and changeable.
- Views a prevailing social order and social processes as constituted and sustained less by the will of individuals than by the pervasiveness of particular constructions or versions of reality – often referred to as discourses.
- Views discourse as coloured by and productive of ideology.
- Views power in society not so much as imposed on individual subjects as an inevitable *effect* of a way particular discursive configurations or arrangements privilege the status and positions of some people over others.
- Views human subjectivity as at least in part constructed or inscribed by discourse, and discourse as manifested in the various ways people *are* and *enact* the sorts of people they are.
- Views reality as textually and intertextually mediated via verbal and non-verbal language systems, and texts as sites for both the inculcation and the contestation of discourses.
- Views the systematic analysis and interpretation of texts as potentially revelatory of ways in which discourses consolidate power and colonize human subjects through often covert position calls.

(Locke 2004: 1–2)

Analysing observations

A small sales and marketing team from a shoe manufacturing company were sent on a tour of the Pacific region to assess market potential. The marketing manager received two early reports. One read: 'The majority of the population are not wearing shoes: excellent marketing opportunity!' The other read: 'Most of the people do not wear shoes: poor marketing opportunity.'

As this (apocryphal) anecdote suggests, it is possible for two people to analyse the same observation data and come to very different, indeed diametrically opposed, conclusions. It is also possible, as the examples of observational studies given in Box 7.11 and in the previous chapter indicate, to focus on either a more quantitative or a more qualitative form of analysis.

You might like to refer back to Box 6.15 at this point.

Box 7.11 Examples of observation analysis

> The first stage of analysis involved transcribing and importing each episode of observation into the QSR NUD*IST program. The transcriptions were read and re-read to form impressions of emerging themes and categories. A set of analytic categories were identified: inverted comma criticism, direct criticism and indirect criticism. In the second stage, data were quantified by counting instances which showed palliative care nurses doing criticism of other professionals who worked outside their organizations, for example, GPs [general practitioners] and hospital doctors. Key phrases spoken by palliative care nurses were identified. The number of times when collective pronouns 'they' or 'them' appeared in talk were counted and the number of times hospital doctors or general practitioners occurred in the nurses' talk was also counted. The constant comparative method helped to reveal systematic differences or similarities in doing criticism in each of the three palliative care settings. It also helped to identify how palliative care nurses constituted their own, their patients and other professionals' moral character. Application of the tools of CA [conversation analysis] helped to deepen analysis so as to reveal and make visible participants' local activities in palliative care nurses' talk . . .
>
> (Li 2005: 1953)

> . . . data were obtained from 7 months of participant observation conducted in a university-affiliated government hospital in Tel Aviv where alternative practitioners were working with hospitalized patients. Observations were supplemented by informal interviews with biomedical and alternative practitioners working together in the hospital.
>
> (Mizrachi and Shuval 2005: 1652)

A number of key points may be made about the analysis of observations in social science research:

- quantified forms of observation lend themselves to fairly routinized forms of data collection and analysis, which can be very powerful in getting across particular issues in tabular or diagrammatic form;
- the collection and analysis of observation data, as with that of other research techniques, occurs as much in parallel as in sequence;
- observation, again like other research techniques, is very often used in conjunction with other methods, both to contextualize and extend the analysis being carried out.

The studies briefly reported in Box 7.11 make clear one further point about

the analysis of observation data, namely that it is selective. This is, however, true of all social research. While this may seem more or less obvious in any particular example, and indeed may be made manifest by the researchers concerned, these characteristics are present in other research projects as well. Being selective and, indeed, partisan is inevitable in research, and it is healthy to recognize and discuss this within your project report or dissertation.

Analysing questionnaires

Some examples of the process of analysing questionnaires are given in Box 7.12.

Box 7.12 Examples of questionnaire analysis

> Research on livelihoods and land use patterns in southern Belize used semi-structured questionnaires to generate qualitative and quantitative data from about 100 respondents in three villages. The data was analysed by hand because the team had no computer, and also both members of the research team could do the work together. Tables were drawn up on paper to contain the answers to each question. All the data was then entered onto these sheets and added up accordingly. Questions included, for example, enquiries about problems faced in agricultural production, producing a range of answers around: limited markets for specific products, lack of credit, and limited access to land in some places.
>
> (Laws et al. 2003: 381)

> A sample of 7318 rating forms from the Universidad del Pais Vasco . . . (UPV) and another sample of 90,905 rating forms from the Universidad Autonoma de Madrid (UAM) were analysed. In both cases, students filled out a rating form [questionnaire] for each teacher from whom they received classes . . . Both questionnaires shared a focus on teacher performance in lecturing . . . The rating form applied at the UPV included 50 items. Sixteen items were dropped for these analyses, because they reflect the dimensions of fulfilment of teachers' formal duties and exercises, as well as those items with a non-response rate higher than 10%. The overall rating items were also dropped from the analyses, as we considered that they would favour unidimensional solutions . . . Applying the same criteria, we analysed 13 of the 17 items in the UAM rating form . . . The statistical analysis was carried out by means of the structural equations model (confirmatory factor analysis, CFA) of the AMOS software and of similarity structure analysis (SSA) non-parametric multidimensional scaling.
>
> (Apodaca and Grad 2005: 733–4)

The data collected by questionnaires may, of course, be either qualitative or quantitative. Alternative strategies for analysing qualitative data have been suggested and discussed in the preceding sections. Questionnaires do, however, lend themselves more to quantitative forms of analysis. This is partly because they are designed to collect mainly discrete items of information, either numbers or words which can be coded and represented as numbers. This emphasis is also partly due to the larger scale of many questionnaire surveys, and their common focus on representation, which encourages a numerical or quasi-numerical summary of the results.

The discussion in this section will focus, therefore, on quantitative forms of analysis. This necessarily calls for some consideration of statistics, which is another of those terms which some readers may find very off-putting or threatening. Quantitative analysis may be used, however, at a number of levels, and the simplest of these may be the most useful in your case (see Box 7.13).

Box 7.13 Levels of quantitative analysis

Descriptive statistics
Variable frequencies, averages, ranges.

Inferential statistics
Assessing the significance of your data and results.

Simple inter-relationships
Cross-tabulation or correlation between two variables.

Multivariate analysis
Studying the linkages between more than two variables.

Many small-scale research studies which use questionnaires as a form of data collection will not need to go beyond the use of descriptive statistics and the exploration of the inter-relationships between pairs of variables (using, for example, cross-tabulations). It will be adequate to say that so many respondents (either the number or the proportion of the total) answered given questions in a certain way; and that the answers given to particular questions appear to be related. Such an analysis will make wide use of proportions and percentages, and of the various measures of central tendency ('averages') and of dispersion ('ranges') (see Box 7.14).

You may, however, wish or need to go beyond this level of analysis, and make use of inferential statistics or multivariate methods of analysis. There are dozens of inferential statistics available: three commonly used examples are outlined in Box 7.15. The functions of these statistics vary, but they are typically used to compare the measurements you have collected from your sample

Box 7.14 Descriptive statistics

For nominal or ordinal data
- proportions
- percentages
- ratios

For interval or ratio data
Measures of central tendency:
- mean: total sum of values divided by the number of cases
- median: the value of the middle case
- mode: the most frequently occurring value

Measures of dispersion:
- range: the difference between the highest and lowest values
- standard deviation: the square root of the mean of the squared deviations from the mean

Box 7.15 Examples of inferential statistics

Chi-square
Function: to compare sets of values
Assumptions: random sampling, nominal data

Kolmogorov-Smirnov
Function: to compare two samples
Assumptions: random sampling, ordinal data

Student's t test
Function: (a) single sample test of mean; (b) two sample test of means
Assumptions: random sampling, interval data, normal distribution

for a particular variable with another sample or a population, in order that a judgement may be made on how similar or dissimilar they are. It is important to note that all of these inferential statistics make certain assumptions about both the nature of your data (see Box 7.16) and about how they were collected, and should not be used if these assumptions do not hold.

Multivariate methods of analysis may be used to explore the inter-relationships among three or more variables simultaneously. Commonly used examples of these are outlined in Box 7.17. While you do not need to have an extensive mathematical knowledge to apply these techniques, as they

Box 7.16 Types of quantitative data

Nominal
Numerical values are assigned to categories as codes. For example, in coding a questionnaire for computer analysis, the response 'male' might be coded as '1', and 'female' as '2'. No mathematical operations can be performed on the resulting codes. No ordering is implied.

Ordinal
Numerical values are assigned in accordance with a qualitative scale. For example, in coding a questionnaire, the responses 'very satisfactory', 'satisfactory', 'neither satisfactory nor unsatisfactory', 'unsatisfactory' and 'very unsatisfactory' are coded '5', '4', '3', '2' and '1' respectively. The ordering of the responses is retained in the coding.

Interval
Measurements are made on a quantitative scale, in which the differences between points are consistently of the same size but the base point is arbitrary. For example, dates. The year 2000 AD occurs 1500 years after the year 500 AD. The ordering of, and distance between, values is given. Addition and subtraction can be used, but not multiplication or division.

Ratio
Measurements are made on a quantitative scale, in which the differences between points are consistently of the same size and there is a 'true zero'. For example, people's ages, countries' populations. All basic mathematical operations – addition, subtraction, multiplication and division – may be applied.

are all available as part of computer software packages, you should at least have an understanding of their principles and purposes.

See also the earlier section in this chapter on **Computer-based data management and analysis**.

One key point to be aware of when carrying out quantitative analyses is the question of causality. One of the purposes of analysis, we have argued, is to seek explanation and understanding. We would like to be able to say that something is so because of something else. However, just because two variables of which you have measurements appear to be related, this does not mean that they are. Statistical associations between two variables may be a matter of chance, or due to the effect of some third variable. In order to demonstrate

Box 7.17 Commonly used multivariate analysis techniques

- *Correlation analysis* – measures the degree and direction of relationships between variables.
- *Regression analysis* – fits a model to a data set, enabling the prediction of the value of one (dependent) variable in terms of one or more other (independent) variables.
- *Analysis of variance (ANOVA)* – measures how independent variables interact with each other and impact upon the dependent variable. *Multivariate analysis of variance (MANOVA)* is used where there is more than one dependent variable.
- *Cluster analysis* – groups cases together into clusters on the basis of their similarity in terms of the variables measured.
- *Factor analysis* – reduces a large number of variables to a limited number of factors, so that the underlying relationships within the data may be more easily assessed.
- *Discriminant analysis* – enables the discrimination between groups on the basis of predictive variables.

causality, you also have to find, or at least suggest, a mechanism linking the variables together.

Interpretation

After presenting a lecture on the book's findings, I was approached by a member of the public who quietly and authoritatively explained:

'It's all bollocks, no offence mind, but it's bollocks. You make us all like fucking wallies, they must be them dopey ones who fuck up everything, but us no. Like me, I'm a face, East End face. I own two houses. I'm her landlord, yeah, the student she pays me rent. I could pull up £250k if I had to. I'm a face. No offence, but all these people in this book they must be fucking backward. I know a good champagne, Bollinger I always drink. I'm not the only one, there's lots like me, all my mates we're all like it, all got a few bob. The Pakis they come in, all this about capitalism and the docks, we moved out 'cos the Pakis. That's why we all vote for Maggie, fucking Labour won't let you buy your Council house. We got money all of us.'

DH: 'Have you read the final section in the chapter on entrepreneurship?'

> 'No, I just had a look at some of her notes, all these silly fucking stories so I thought I'd come and front you with it.'
>
> (Hobbs 1993: 60–1)

Interpretation is the process by which you put your own meaning on the data you have collected and analysed, and compare that meaning with those advanced by others.

Your own perspective

We have emphasized at a number of places in this book how important it is to recognize, and make explicit, your own role and position within your research. This is partly about asserting ownership, and partly about recognizing the possible limitations, influences and biases of your own perspective. A critical element of the data analysis process is arriving at your own assessment of what the results mean, and how these relate to other relevant research and writing in your subject area. What do you think is significant? What do you think this suggests? Where and how do you think this kind of study might be developed further? These are the kinds of questions you should be asking yourself, and doing so, at least initially, without any direct reference to other authorities.

> **Hint**: Try explaining it to a non-specialist again. See the section on **Focusing** in Chapter 2.

Distancing yourself from your data sources

At the same time as recognizing and asserting your own perspective on your data and their analysis, it is important to not get too embedded and bound up in this view. Researchers generally have a commitment to their projects, their methods, their data and their interpretations. It is healthy, therefore, to stand back for a time and attempt to view your research from the more dispassionate perspective of an outsider. Of course, it is impossible to do this in any absolute sense, given the personal commitment which any researcher makes to their research. But it is possible to achieve some distance, though the ways in which you might do this will probably vary. Possible strategies include:

- The management of your data, through the processes of coding, annotating, labelling and so on, as discussed in an earlier section in this chapter, can reduce their immediacy and make them appear as if they have been put together by somebody else.
- Taking some time out, perhaps a week or two, before you come back to your analysis can increase the strangeness or foreign-ness of your data, and

lead to a livelier interpretation (this is not a bad idea, if you have the time and are not afraid that you will loose your purchase, at any stage of your research).
- Analysing your data alongside a similar set may lead you to focus on the similarities and dissimilarities, rather than just on your own findings and interpretations.

Shared understandings

Having recognized, and begun to develop, your own perspective on what your research indicates, it also becomes important to review these views in the light of those of others. To what extent do your findings, and your interpretation of them, agree or disagree with those of other authorities or researchers? Confirmatory or supportive results can be extremely useful in advancing general understanding. Such shared understandings can also be generated, by, for example, reporting on your findings in a seminar, workshop, conference or paper, and debating with others on their significance or interpretation.

How to handle different accounts

As well as recognizing and building upon shared understandings, you will also need to be able to accept and work from alternative perspectives. This can occur in at least two major ways: different accounts within your own data, and differences between your interpretations and those of others. Both are to be expected, welcomed and acknowledged. There is no reason, given our lack of comprehensive understanding of the world we live in, together with the varied perspectives held by different individuals, why our views and behaviours should always be common and shared. An important part of the interpretation of research is, therefore, the recognition of the diverging patterns within the data collected, and the attempted explanation of these. Similarly, you should not be unduly concerned if your findings appear to diverge from those of other researchers in your field; but you should look for reasons why this might be so, and/or argue the relevance of your interpretation against those of others.

The value of data that doesn't fit

The preceding discussion suggests the importance of the observation that doesn't fit your general interpretation, or 'the exception that proves the rule'. This saying may, of course, be taken at least two ways. One, the most literal reading, suggests that a single exception is a rogue piece of data which should in effect be ignored. The other reading, perhaps the more relevant to the research process, would be that data which don't fit should not be ignored, but accepted, reported and cherished. It is not uncommon for accepted

interpretations to be challenged and eventually demolished. Do not cast aside pieces of data which may be the basis for doing this!

What does it all mean?

Unless your interpretation is to be a one-off and wholly personal exercise, you will have to engage in a more general consideration of the relevance and usefulness of your work. Such a consideration will bring you into touch with four related concepts: significance, generalizability, reliability and validity. All competent researchers need to have an understanding of what these concepts mean, and need to be able to review and defend their own work in this light (see Box 7.18).

Box 7.18 Significance, generalizability, reliability and validity

Significance
The concept of significance has both a specific, statistical meaning and a more general, common-sense interpretation. In statistical parlance, it refers to the likelihood that a result derived from a sample could have been found by chance. The more significant a result, the more likely that it represents something genuine. In more general terms, significance has to do with how important a particular finding is judged to be.

Generalizability
The concept of generalizability, or representativeness, has particular relevance to small-scale research. It relates to whether your findings are likely to have broader applicability beyond the focus of your study. Thus, if you have carried out a detailed study of a specific institution, group or even individual, are your findings of any relevance beyond that institution, group or individual? Do they have anything to say about the behaviour or experience of other institutions, groups or individuals, and, if so, how do you know that this is the case?

Reliability
The concept of reliability has to do with how well you have carried out your research project. Have you carried it out in such a way that, if another researcher were to look into the same questions in the same setting, they would come up with essentially the same results (though not necessarily an identical interpretation). If so, then your work might be judged reliable.

Validity
Validity has to do with whether your methods, approaches and techniques actually relate to, or measure, the issues you have been exploring.

In the end, your interpretation of your findings is, however, limited by the methods you have used and the sample you have studied:

> some of the seeming polarity in the debate around home-based work can be explained by the differing methodologies and sampling procedures. For instance, the evidence collected by local homeworking research projects, officers and campaigns on the incidence and persistence of extremely low-paid, arduous, manufacturing, home-based work in the UK had often only been possible after building of trust between project workers and home-based workers . . . but because these surveys have been conducted largely in inner cities they have had little to say about non-manual homeworkers and whether the latter face particular problems.
>
> (Phizacklea and Wolkowitz 1995: 19)

Small-scale research has its limitations, therefore, but is also able to make a significant contribution in under-studied areas.

Summary

Having read this chapter, you should:

- have an appreciation of the different forms of data, and the kinds of analysis appropriate to them;
- be aware of the interleaving processes of data management, analysis and interpretation which are involved in making sense of your data collection;
- have an understanding of the different approaches which you might take to the analysis of documents, interviews, observations and questionnaires;
- be able to assess the significance, generalizability, reliability and validity of your research and findings.

Exercises

7.1 Box 7.2 contains some examples of quantitative data. What kinds of numbers are included? How do you think they were collected or produced? What might you do with or say about these numbers?

7.2 Box 7.3 contains some examples of qualitative data. What kinds of words are included? How do you think they were collected or produced? What might you do with or say about these words?

7.3 What do you understand by the term 'analysis'? Try to write an explanatory definition in your own words. You probably won't find it very helpful to turn to a dictionary, as these tend to give only brief definitions and do not have a research focus.

7.4 Take notes of a meeting you have to attend, or, if you rarely attend meetings, a television programme. After a few days, take a careful look at your notes. How full a summary are they? What has been left out, and why? What biases are there? Why are they organized in this particular way? What other documents would allow you to better assess the value of your notes?

Further reading

In this section, we list a selection of books that are of particular relevance to the topics discussed in this chapter, together with an indication of their contents.

Argyrous, G. (2000) *Statistics for Social and Health Research, With a Guide to SPSS*. London: Sage.
The six sections of this comprehensive text cover univariate descriptive statistics, bivariate descriptive statistics, inferential statistics (for one sample, two or more independent samples, and two dependent samples), and multivariate descriptive statistics.

Babbie, E. R. and Halley, F. (2005) *Adventures in Social Research: Data Analysis Using SPSS for Windows*, 5th edn. London: Pine Forge.
Designed for students, this text introduces SPSS through Windows. The text includes activities to aid learning.

Bazeley, P. and Richards, L. (2000) *The NVivo Qualitative Project Book*. London: Sage.
How to use qualitative data analysis software. Includes demonstration software on a CD-ROM.

Blaiklie, N. (2003) *Analysing Quantitative Data: From Description to Explanation*. London: Sage.
Aims to demystify quantitative analysis and help the reader overcome symbol phobia and figure blindness.

Bryman, A. and Cramer, D. (2004) *Quantitative Data Analysis with SPSS 12 and 13*, 2nd edn. London: Routledge.
Designed as a non-mathematical introduction for social scientists, explains the use of statistical tests in non-technical language.

Burn, A. and Parker, D. (2003) *Analysing Media Texts*. London: Continuum.
A variety of forms of texts and their analysis are explored in this book, including moving images, web sites, computer games and interview data.

Champney, L. (1995) *Introduction to Quantitative Political Science*. New York: HarperCollins.
An introduction to the computation and interpretation of statistics, and to key concepts of social scientific enquiry. Includes both academic and

practical applications, together with examples of research articles and reports with accompanying critiques.

Coxon, A. P. M. (1999) *Sorting Data: Collection and Analysis*. Thousand Oaks, CA: Sage.

Part of an extensive series of short books on 'quantitative applications in the social sciences'. Includes chapters on the collection, description and comparison, and analysis of free sorting data.

de Vaus, D. (2002) *Analyzing Social Science Data: 50 Key Problems in Data Analysis*. London: Sage.

A substantial text organized in seven parts: data preparation, variable preparation, data reduction, generalization, single variable analysis, two variable analysis, multivariate analysis.

Fairclough, N. (2003) *Analysing Discourse: Textual Analysis for Social Research*. London: Routledge.

Standard guide to these analytical techniques.

Field, A. (2005) *Discovering Statistics Using SPSS for Windows: Advanced Techniques for the Beginner*, 2nd edn. London: Sage.

Covers data exploration, correlation and regression, logistic regression, comparing means, analysis of variance and factor analysis.

Fielding, J. L. and Gilbert, G. N. (2000) *Understanding Social Statistics*. London: Sage.

The three sections of the book focus on preliminary issues (including the use of computers), univariate analysis (e.g. frequencies, percentages, measures of central tendency, the normal curve) and bivariate analysis (e.g. correlation and regression, sampling and inference, modelling data).

Foster, J. J. (2001) *Data Analysis Using SPSS for Windows*, new edition. London: Sage.

Following an introductory overview of statistical analysis, 20 chapters set out what can be done with SPSS, illustrated stage by stage. Topics covered include t-tests, analysis of variance, correlation and regression, non-parametric techniques, reliability analysis and factor analysis.

Gahan, C. and Hannibal, M. (1998) *Doing Qualitative Research Using QSR NUD*IST*. London: Sage.

A practical guide to using the NUD*IST package for the analysis of unstructured data such as text from interviews, historical or legal documents, or non-textual material such as videotapes.

Gibbs, G. (2002) *Qualitative Data Analysis: Explorations with NVivo*. Buckingham: Open University Press.

Chapters cover data preparation, coding, memos and attributes, searching for text, developing an analytic scheme, visualizing the data and communicating.

Hinde, A. (1998) *Demographic Methods*. London: Arnold.

Describes and explains the methods used to analyse population data. Covers basic methods, as well as parity progression ratios, survival analysis and birth interval analysis.

Hinton, P. R. (2004) *Statistics Explained: A Guide for Social Science Students*, 2nd edn. London: Routledge.
Written for psychology and other social science students, this text takes the reader through the principles of statistical analysis. Contents include descriptive statistics, hypothesis testing, sampling, significance, variance, chi-square and using statistics programs on computers.

Howarth, D. (2000) *Discourse*. Buckingham: Open University Press.
A comprehensive overview of the different conceptions and methods of discourse analysis, and of the traditions of thinking (structuralist, post-structuralist, post-Marxist) from which these have emerged.

Hoyle, R. H. (ed.) (1999) *Statistical Strategies for Small Sample Research*. Thousand Oaks, CA: Sage.
The methods and issues considered include randomized designs, bootstrapping, categorical and nonparametric data, dynamic factor analysis and structural equation modelling.

Kinnear, P. and Gray, C. (2004) *SPSS 12 Made Simple*. London: Routledge.
Covers all the facilities available in SPSS, from creating and manipulating files through comparing averages to exploratory factor analysis.

Locke, T. (2004) *Critical Discourse Analysis*. London: Continuum.
This slim volume covers theory and practice, and the analysis of both print and oral texts.

Miller, R., Acton, C., Fullerton, D. and Maltby, J. (2002) *SPSS for Social Scientists*. Basingstoke: Palgrave Macmillan.
Covers hypothesis testing, cross-tabulation, analysis of variance, correlation and regression, factor analysis and loglinear analysis.

Pallant, J. (2001) *SPSS Survival Manual: A Step by Step Guide to Data Analysis using SPSS*. Buckingham: Open University Press.
Text and screen illustrations take the reader through data preparation, preliminary analyses, and the use of a range of statistical techniques to explore relationships and compare groups.

Salkind, N. (2004) *Statistics for People who (think they) hate Statistics*, 2nd edn. Thousand Oaks, CA: Sage.
Uses a conversational tone to guide the reader through simple and more advanced statistical techniques.

Spicer, J. (2004) *Making Sense of Multivariate Data Analysis*. Thousand Oaks, CA: Sage.
Discusses philosophy, theory and practice.

Wright, D. B. (2002) *First Steps in Statistics*. London: Sage.
Chapters consider graphing, distributions, sampling and allocation, inference and confidence intervals, hypothesis testing, analysis of variance, regression and correlation, and contingency tables.