

STEP VII

PROCESSING AND DISPLAYING DATA

This operational step includes two chapters:

- [Chapter 15: Processing data](#)
- [Chapter 16: Displaying data](#)

PROCESSING DATA

In this chapter you will learn about

- **Methods for processing data in quantitative studies**
- **How to edit data and prepare for coding**
- **How to code data**
- **How to code qualitative data in quantitative studies**
- **Methods for processing data in qualitative studies**
- **Analysing data in qualitative and quantitative studies**
- **The role of computers in data analysis**
- **The role of statistics in research**

Keywords

analysis,
closed questions,
code book,
coding,
concepts,
content analysis,
cross-tabulation,
data display,
data processing,
editing,
frame of analysis,
frequency distribution,
multiple responses,
open-ended questions,
pre-test.

At the end of this chapter, you should have an understanding of

- **What data processing in quantitative studies entails**
- **How to process data in qualitative studies**
- **How to develop a code book and the process of coding in quantitative studies**
- **Developing themes from open-ended responses, descriptions or narrations**
- **The role of computers and statistics in a research study**

If you were actually doing a research study, you would by now have reached the stage of having either extracted or collected the required information for your study. But what to do with this information? What steps do you need to

take beyond data collection to find answers to your research questions? How do you make sense of the information collected? How do you draw meaning from the data gathered? If you have a hypothesis, how do you accept or reject it? How should the information be analysed to achieve the objectives of your study? To answer these questions you need to subject your data to a number of procedures that constitute the core of data processing ([Figure 15.1](#)).

These procedures are the same whether your study is quantitative or qualitative, but what you do within each procedure is different. For both types of study you need to visualise how you are going to present your findings to your readership in the light of its background and the purpose of the study. You need to decide what type of analysis would be appropriate for the readers of your report. It is in the light of the purpose of your study and your impression about the level of understanding of your readership that you decide the type of analysis you should undertake. For example, there is no point in doing a sophisticated statistical analysis if your readers are not familiar with statistical procedures. In quantitative research the main emphasis in data analysis is to decide how you are going to analyse information obtained in response to each question that you asked of your respondents. In qualitative research the focus is on what should be the basis of analysis of the information obtained; that is, is it contents, discourse, narrative or event analysis? Because of the different techniques used in processing data in quantitative and qualitative research, this chapter deals separately with data processing in quantitative and qualitative studies.

Data processing in quantitative studies

Editing

Irrespective of the method of data collection, the information collected is called *raw data* or simply *data*. The first step in processing your data is to ensure that the data is ‘clean’ – that is, free from inconsistencies and incompleteness. This cleaning process is called **editing**.

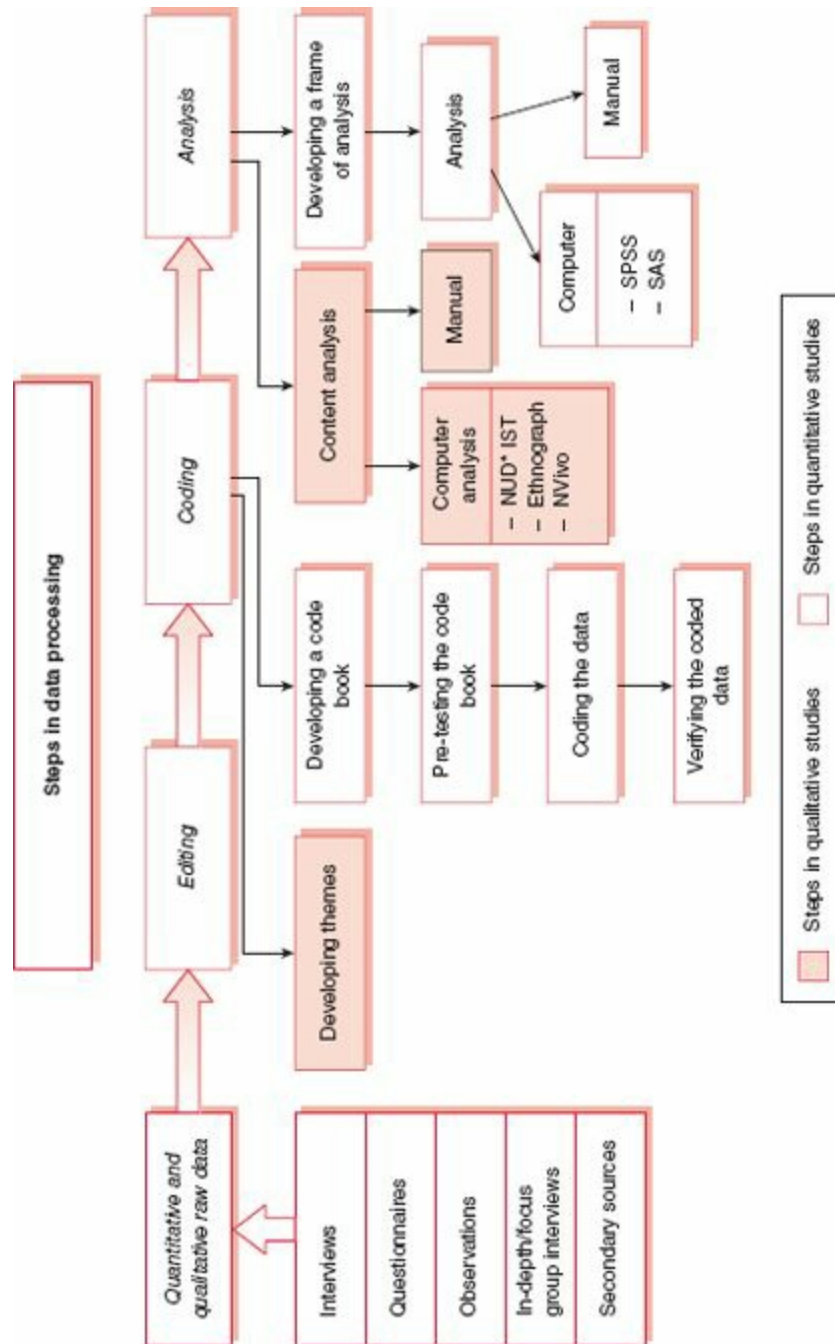


Figure 15.1 Steps in data processing

Editing consists of scrutinising the completed research instruments to identify and minimise, as far as possible, errors, incompleteness, misclassification and gaps in the information obtained from the respondents. Sometimes even the best investigators can:

- **forget to ask a question;**

- **forget to record a response;**
- **wrongly classify a response;**
- **write only half a response;**
- **write illegibly.**

Similar problems can arise in questionnaires. These problems to a great extent can be reduced simply by (1) checking the contents for completeness, and (2) checking the responses for internal consistency.

The way you check the contents for completeness depends upon the way the data has been collected. In the case of an interview, just checking the interview schedule for the above problems may improve the quality of the data. It is good practice for an interviewer to take a few moments after completing an interview to peruse responses for possible incompleteness and inconsistencies. In the case of a questionnaire, again, just by carefully checking the responses some of the problems may be reduced. There are several ways of minimising such problems:

- **By inference** – Certain questions in a research instrument may be related to one another, and it might be possible to find the answer to one question from the answer to another. Of course, you must be careful about making such inferences or you may introduce new errors into the data.
- **By recall** – If the data is collected by means of interviews, sometimes it might be possible for the interviewer to recall a respondent's answers. Again, you must be extremely careful.
- **By going back to the respondent** – If the data has been collected by means of interviews or the questionnaires contain some identifying information, it is possible to visit or phone a respondent to confirm or ascertain an answer. This is, of course, expensive and time-consuming.

There are two ways of editing the data:

- 1. examine all the answers to one question or variable at a time;**
- 2. examine all the responses given to all the questions by one respondent at a time.**

The author prefers the second method as it provides a total picture of the responses, which also helps you to assess their internal consistency.

Coding

Having cleaned the data, the next step is to code it. The method of **coding** is largely dictated by two considerations:

- 1. the way a variable has been measured (measurement scale) in your research instrument (e.g. if a response to a question is descriptive, categorical or quantitative);**
- 2. the way you want to communicate the findings about a variable to your readers.**

For coding, the first level of distinction is whether a set of data is qualitative or quantitative in nature. For qualitative data a further distinction is whether the information is descriptive in nature (e.g. a description of a service to a community, a case history) or is generated through discrete qualitative categories. For example, the following information about a respondent is in discrete qualitative categories: income (above average, average, below average); gender (male, female); religion (Christian, Hindu, Muslim, Buddhist, etc.); and attitude towards an issue (strongly favourable, favourable, uncertain, unfavourable, strongly unfavourable). Each of these variables is measured either on a **nominal scale** or an **ordinal scale**. Some of them could also have been measured on a **ratio scale** or an **interval scale**. For example, income can be measured in dollars (ratio scale), or an attitude towards an issue can be measured on an interval or a ratio scale. The way you proceed with the coding depends upon the measurement scale used in the measurement of a variable at the time of collecting information and whether a question is open-ended or closed.

In addition, the types of statistical procedures that can be applied to a set of information to a large extent depend upon the measurement scale on which a variable was measured in the research instrument. For example, you can calculate different statistical descriptors such as mean, mode and median if income is measured on a ratio scale, but not if it is measured on an ordinal or a nominal scale. It is extremely important to understand that the way you are able to analyse a set of information is dependent upon the measurement scale used in the research instrument for measuring a variable. It is therefore important to visualise – particularly at the planning stage when constructing the research instrument – the way you are going to communicate your findings.

How you can analyse information obtained in response to a question depends upon how a question was asked, and how a respondent answered it. In other words, it depends upon the measurement scale on which a response can be measured or classified. If you study answers given by your respondents in reply to a question, you will realise that almost all responses can be classified into one of the following three categories:

- 1. quantitative responses;**
- 2. categorical responses (which may be quantitative or qualitative);**
- 3. descriptive responses (which are invariably qualitative – keep in mind that this is qualitative data collected as part of quantitative research and not qualitative research).**

For the purpose of analysis, quantitative and categorical responses need to be dealt with differently from descriptive ones. Both quantitative and categorical information go through a process that is primarily aimed at transforming the information into numerical values, called **codes**, so that the information can be easily analysed, either manually or by computer. On the other hand, descriptive information first goes through a process called **content analysis**, whereby you identify the main themes that emerge from the descriptions given by respondents in answer to questions. Having identified the main themes, there are three ways that you can deal with them: (1) you can examine verbatim responses and integrate them into the text of your report to either support or contradict your argument; (2) you can assign a code to each theme and count how frequently each has occurred; and (3) you can combine both methods to communicate your findings. This is your choice, and it is based on your impression of the preference of your readers.

For coding quantitative and qualitative data in quantitative studies you need to go through the following steps:

- Stage I developing a code book;**
- Stage II pre-testing the code book;**
- Stage III coding the data;**
- Stage IV verifying the coded data.**

Stage I: Developing a code book

A **code book** provides a set of rules for assigning numerical values to answers obtained from respondents. Let us take an example. [Figure 15.2](#) lists some questions taken from a questionnaire used in a survey conducted by the author to ascertain the impact of occupational redeployment on an individual. The questions selected should be sufficient to serve as a prototype for developing a code book, as they cover the various issues involved in the process.

There are two formats for data entry: 'fixed' and 'free'. In this chapter we will be using the fixed format to illustrate how to develop a code book. The fixed format stipulates that a piece of information obtained from a respondent is entered in a specific column. Each column has a number and the 'Col. no.' in the code book refers to the column in which a specific type of information is to be entered. The information about an individual is thus entered in a row(s) comprising these columns.

For a beginner it is important to understand the structure of a code book. The example in [Table 15.1](#) is based on the responses given to the questions listed in [Figure 15.2](#).

In [Table 15.1](#), *column 1* refers to the columns in which a particular piece of information is to be entered. Allocation of columns in a fixed format is extremely important because, when you write a program, you need to specify the column in which a particular piece of information is entered so that the computer can perform the required procedures.

Column 2 identifies the question number in the research instrument for which the information is being coded. This is primarily to identify coding with the question number in the instrument.

Column 3 refers to the name of the variable. Each variable in a program is given a unique name so that the program can carry out the requested statistical procedures. Usually there are restrictions on the way you can name a variable (e.g. the number of characters you can use to name a variable and whether you use the alphabet or numerals). You need to check the program you are using for this. It is advisable to name a variable in such a way that you can easily recognise it from its name.

Column 4 lists the possible responses to the various questions. Developing a response pattern for the questions is the most important, difficult and time-consuming part of developing a code book. The degree of difficulty in developing a response pattern differs with the types of questions in your research instrument (open-ended or closed). If a question is closed, the

response pattern has already been developed as part of the instrument construction and all you need to do at this stage is to assign a numerical value to each response category. In terms of analysis, this is one of the main advantages of closed questions. If a closed question includes 'other' as one of the response categories, to accommodate any response that you may not have listed when developing the instrument, you should analyse the responses and assign them to non-overlapping categories in the same way as you would do for open-ended questions. Add these to the already developed response categories and assign each a numerical value.

1 Please indicate:

- (a) Your current age in completed years: _____
- (c) Your marital status: (Please tick)
 - Currently married _____
 - Living in a de facto relationship _____
 - Separated _____
 - Divorced _____
 - Never married _____

2 (b) If tertiary/university, please specify the level achieved and area of study. (Please specify all postgraduate qualifications.)

Level of achievement	Area of study: e.g. engineering, accounting
Associate diploma	
Diploma	
Bachelor	
Graduate diploma	
Masters	
PhD	

11 What, in your opinion, are the main differences between your jobs prior to and after redeployment?

12 We would like to know your perception of the two jobs *before* and *after* redeployment with respect to the following aspects of your job. Please rate them on a five-point scale using the following guide: 5 = extremely satisfied, 4 = very satisfied, 3 = satisfied, 2 = dissatisfied, 1 = extremely dissatisfied

Before redeployment					Areas	After redeployment				
1	2	3	4	5		1	2	3	4	5
					Job status					
					Job satisfaction					
					Motivation to work					
					Interest in the job					
					Self-esteem					
					Professional competence					
					Peer interaction					
					Morale					
					Work environment					
					Social interaction					

Figure 15.2 Example of questions from a survey

If the number of responses to a question is less than nine, you need only one column to code the responses, and if it is more than nine but less than 99, you need two columns (column 1 in the code book). But if a question asks respondents to give more than one response, the number of columns assigned should be in accordance with the number of responses to be coded. If there

are, say, eight possible responses to a particular question and a respondent is asked to give three responses, you need three columns to code the responses to the question. Such questions are called multiple response questions. Let us assume there are 12 possible responses to a question. To code each response you need two columns and, therefore, to code three responses you need six columns.

The coding of open-ended questions is more difficult. Coding of open-ended questions requires the response categories to be developed first through a process called content analysis. One of the easier ways of analysing open-ended questions is to select a number of interview schedules/questionnaires randomly from all the completed interview schedules/questionnaires received. Select an open-ended question from one of these schedules/questionnaires and write down the response(s) on a sheet of paper. If the person has given more than one response, write them separately on the same sheet. Similarly, from the same questionnaire/schedule select another open-ended question and write down the responses given on a separate sheet. In the same way you can select other open-ended questions and write down the response(s). Remember that the response to each question should be written on a separate sheet. Now select another questionnaire/interview schedule and go through the same process, adding response(s) given for the same question on the sheet for that question. Continue the process until you feel that the responses are being repeated and you are getting no or very few new ones – that is, when you have reached a saturation point.

Now, one by one, examine the responses to each question to ascertain the similarities and differences. If two or more responses are similar in meaning though not necessarily in language, try to combine them under one category. Give a name to the category that is descriptive of the responses. Remember, when you code the data you code categories, *not* responses per se. It is advisable to write down the different responses under each category in the code book so that, while coding, you know the type of responses you have grouped under a category. In developing these categories there are three important considerations:

- 1. The categories should be *mutually exclusive*. Develop non-overlapping categories. It should not be possible to place a response in two categories.**

- 2. The categories should be *exhaustive*; that is, it should be possible to place almost every response in one of the categories. If too many responses cannot be so categorised, it is an indication of ineffective categorisation. In such a situation you should examine your categories again.**
- 3. The use of the ‘other’ category, effectively a ‘waste basket’ for those odd responses that cannot be put into any category, must be kept to the absolute minimum because, as mentioned, it reflects the failure of the classification system. This category should not include more than 5 per cent of the total responses and should not contain more responses than any other category.**

Column 5 lists the actual codes of the code book that you decide to assign to a response. You can assign any numerical value to any response so long as you do not repeat it for another response within the same question. Two responses to questions are commonly repeated: ‘not applicable’ and ‘no response’. You should select a number that can be used for these responses for all or most questions. For example, responses such as ‘not applicable’ and ‘no response’ could be given a code of 8 and 9 respectively, even though the responses to a question may be limited to only 2 or 3. In other words, suppose you want to code the gender of a respondent and you have decided to code female = 1 and male = 2. For ‘no response’, instead of assigning a code of 3, assign a code of 9. This suggestion helps in remembering codes, which will help to increase your speed in coding.

Table 15.1 An example of a code book

Col. no.	Q. no.	Variable name	Response pattern	Code
1	2	3	4	5
1-3	S. no.	ID	Actual serial number	Code actual
4	Record no.	RNO	First record	1
			Second record	2
			Third record	3
5	1(a)	Age	20-24	1
			25-29	2
			30-34	3
			35-39	4
			40-44	5
			45-49	6
			No response	9
6	1(c)	MS	Currently married	1
			Living in a de facto relationship	2
			Separated	3
			Divorced	4
			Never married	5
			No response	9
	2(b)	TEDU	Assoc. Dip.	1
			Diploma	2
			Bachelors	3
			Grad. Dip.	4
			Masters	5
			PhD	6
			Not applicable	8
			No response	9
7		TEDU1	Same as in TEDU	Code
8		TEDU2	Same as in TEDU	as in
9		TEDU3	Same as in TEDU	TEDU
		STUDY	Behavioural Sciences	1
			Business	2
			Economics/Commerce	3
			Communication	4
			Engineering	5
			Geography	6
			History	7
			Graphics	8
			Librarianship	9
			Nursing	10
			Performing Arts	11

Table 15.1 (Continued)

Col. no.	Q. no.	Variable name	Response pattern	Code
			Secretarial	12
			Social Work	13
			Psychology	14
			Education	15
			Chartered Acct.	16
			Zoology	17
			Anthropology	18
			Social Sciences	19
			Public/ Health Admin.	20
			English	21
			Audio-visual Aids Education	22
			Not applicable	88
			No response	99
10-11	2(b)	STUDY 1	Same as in STUDY	Code
12-13		STUDY 2	Same as in STUDY	as in STUDY
14-15		STUDY3	Same as in STUDY	
	11	DFWK	Lack of job satisfaction	
			- current job	01
			- previous job	02
			Less responsibility	
			- current job	03
			- previous job	04
			Low morale	
			- current job	05
			- previous job	06
			Lack of recognition of hard work	
			- current job	07
			- previous job	08
			Skills irrelevant	
			- current job	09
			- previous job	10
			Repetitious nature	
			- current job	11
			- previous job	12
			Job more segmented	
			- current job	13
			- previous job	14
			Reduction in occup. status	15
			Greater involvement in total process	
			- current job	16

Col. no.	Q. no.	Variable name	Response pattern	Code
			- previous job	17
			More restricted duties	
			- current job	18
			- previous job	19
			Less flexibility	
			- current job	20
			- previous job	21
			More people contact	
			- current job	22
			- previous job	23
			Less people contact	
			- current job	24
			- previous job	25
			More responsibility	
			- current job	26
			- previous job	27
			Better work environment	
			- current job	28
			- previous job	29
			Better work morale	
			- current job	30
			- previous job	31
			Working in a team	
			- current job	32
			- previous job	33
			Work as an individual	
			- current job	34
			- previous job	35
			Lack of job security	
			- current job	36
			- previous job	37
			Current job full-time	38
			Better career prospects	
			- current job	39
			- previous job	40
			No commonality	41
			Different working conditions	42
			Change in occup. status (not specified)	43
			High occup. level	
			- current job	44
			- previous job	45

Table 15.1 (Continued)

Col. no.	Q. no.	Variable name	Response pattern	Code
			More job security in present job	46
			More professional environment in	
			- current job	47
			- previous job	48
			Different skills to learn in current job	49
			Higher workload	
			- current job	50
			- previous job	51
			Less supervision in	
			- current job	52
			- previous job	53
			Current job more relevant to my training	54
			Current job more rewarding	55
			Better morale in current job	56
			Greater variety of duties in	
			- current job	57
			- previous job	58
			Current job more rewarding	59
			Better morale in	
			- current job	60
			Less variety of tasks	
			in current job	61
			Fewer skills required in	
			- current job	62
			- previous job	63
			Change in job title	64
			Previous job very demanding	65
			Opportunity to use initiative in previous job	66
			Current job does not require initiative	67
			Different skills required in current job	68
			No career prospects	
			- current job	69
			- previous job	70
			More professional environment in current job	71
			Different skills to learn in current job	72
			Different working conditions	73
			Change in occup. status (change not specified)	74
			Change in job resp. (change not specified)	75
			Current job is more structured	76
			Now part of the team	77

Col. no.	Q. no.	Variable name	Response pattern	Code
			No change	78
			Irrelevant response	79
			Other	80
			Not applicable	88
			No response	99
16–17	11	DIFWK1	Same as in DIFWK	Code as
18–19		DIFWK2	Same as in DIFWK	in
20–21		DIFWK3	Same as in DIFWK	DIFWK
22–23		DIFWK4		
24–25		DIFWK5		
			Extremely dissatisfied	1
			Dissatisfied	2
	12	BEAFTER	Satisfied	3
			Very satisfied	4
			Extremely satisfied	5
			Same as in BEAFTER	
26		JOBSTA	Same as in BEAFTER	Code as
27		JOBSTB	Same as in BEAFTER	in
				BEAFTER
28		JOBSTA	Same as in BEAFTER	Code as in
29		JOBSTA	Same as in BEAFTER	BEAFTER
			and so on	

To explain how to code, let us take each of the questions listed in the example in [Figure 15.2](#).

Question 1(a)

Your current age in completed years: _ _ _ _ _

This is an open-ended quantitative question. In questions like this it is important to determine the range of responses – the respondent with the lowest and the respondent with the highest age. To do this, go through a number of questionnaires/interview schedules. Once the range is established, divide it into a number of categories. The categories developed are dependent upon a number of considerations such as the purpose of analysis, the way you want to communicate the findings of your study and whether the findings are going to be compared with those of another study. Let us assume that the range in the study is 23 to 49 years and assume that you decide on the following categories to suit your purpose: 20–24, 25–29, 30–34, 35–39, 40–44 and 45–49. If your range is correct you should need no other categories. Let us assume that you decide to code 20–24 = 1, 25–29 = 2, 30–34 = 3, and

so on. To accommodate ‘no response’ you decide to assign a code of 9. Let us assume you decided to code the responses to this question in column 5 of the code sheet.

Question 1(c)

Your marital status: (Please tick)

*Currently married*_____

*Living in a de facto relationship*_____

*Separated*_____

*Divorced*_____

*Never married*_____

This is a closed categorical question. That is, the response pattern is already provided. In these situations you just need to assign a numerical value to each category. For example, you may decide to code ‘currently married’ = 1, ‘living in a de facto relationship’ = 2, ‘separated’ = 3, ‘divorced’= 4 and ‘never married’= 5. You may add ‘no response’ as another category and assign it a code of 9. The response to this question is coded in column 6 of the code sheet.

Question 2(b)

If tertiary/university, please specify the level achieved and the area of study. (Please specify all postgraduate qualifications.)

In this question a respondent is asked about tertiary qualifications. The question asks for two aspects: (1) level of achievement, which is categorical; and (2) area of study, which is open-ended. Also, a person may have more than one qualification, which makes it a multiple response question. Both aspects of the question are to be coded – the level of achievement (e.g. associate diploma, diploma) and the area of study (e.g. engineering, accounting). When coding multiple responses, decide on the maximum possible number of responses to be coded. Let us assume you code a maximum number of three levels of tertiary education. (This would depend upon the maximum number of levels of achievement identified by the study)

population.) Firstly, code the levels of achievement TEDU (= tertiary education) and then the area of Study STUDY. In the above example, let us assume that you decided to code three levels of achievement. To distinguish them from each other we call the first level TEDU1, the second TEDU2 and the third TEDU3, and decide to code them in columns, 7, 8 and 9, respectively. Similarly, the names given to the three areas of study are STUDY1, STUDY2 and STUDY3 and we decide to code them in columns 10–11, 12–13 and 14–15. The codes (01 to 22) assigned to different qualifications are listed in the code book. If a respondent has only one qualification, the question of second and third qualifications is not applicable and you need to decide on a code for ‘not applicable’. Assume you assigned a code of 88. ‘No response’ would then be assigned a code of 99 for this question.

Question 11

What, in your opinion, are the main differences between your jobs prior to and after redeployment?

This is an open-ended question. To code this you need to go through the process of content analysis as explained earlier. Within the scope of this chapter it is not possible to explain the details, but response categories that have been listed are based upon the responses given by 109 respondents to the survey on occupational redeployment. In coding questions like this, on the one hand you need to keep the variation in the respondents’ answers, and on the other you want to break them up into meaningful categories to identify the commonalities. Because this question is asking respondents to identify the differences between their jobs before and after redeployment, for easy identification let us assume this variable was named DIFWK (*difference in work*). Responses to this question are listed in [Figure 15.3](#). These responses have been selected at random from the questionnaires returned.

A close examination of these responses reveals that a number of themes are common, for example: ‘learning new skills in the new job’; ‘challenging tasks are missing from the new position’; ‘more secure in the present job’; ‘more interaction in the present job’; ‘less responsibility’; ‘more variety’; ‘no

difference'; 'more satisfying'. There are many similar themes that hold for both the before and after jobs. Therefore, we developed these themes for 'current job' and 'previous job'.

One of the main differences between qualitative and quantitative research is the way responses are used in the report. In qualitative research the responses are normally used either verbatim or are organised under certain themes and the actual responses are provided to substantiate them.

In quantitative research the responses are examined, common themes are identified, the themes are named (or categories are developed) and the responses given by respondents are classified under these themes. The data then can also be analysed to determine the frequency of the themes if so desired. It is also possible to analyse the themes in relation to some other characteristics such as age, education and income of the study population.

The code book lists the themes developed on the basis of responses given. As you can see, many categories may result. The author's advice is not to worry about this as categories can always be combined later if required. The reverse is impossible unless you go back to the raw data.

Let us assume you want to code up to five responses to this question and that you have decided to name these five variables as DIFWK1, DIFWK2, DIFWK3, DIFWK4 and DIFWK5. Let us also assume that you have coded them in columns 16–17, 18–19, 20–21, 22–23 and 24–25, respectively.

Respondent 3

Hours now FT: totally different skills required; deal with public; busier, more structured day and duties; now a part of the team instead of an independent worker.

Respondent 20

That I am happy and made to feel as though I am a valuable part of a team.

Respondent 41

This one is great, other one was lousy due to mismanagement, poor morale and feelings that dedication and hard work counted for nothing. Department of ... is well managed and morale is good and the graphic design work is fun and I am supported by my supervisors and subordinates.

Respondent 48

<i>Before</i>	<i>After</i>
<i>15 hours per week</i>	<i>24 hours per week</i>
<i>Monday to Friday</i>	<i>Very satisfactory</i>
<i>On the go the whole time</i>	<i>Dealing with the public</i>
<i>No sitting</i>	<i>Dealing with severe psychiatric</i>
<i>Dealings with the public</i>	<i>patients can be very stressful.</i>

Respondent 52

No difference.

Respondent 54

My substantive position has been the same before and after the redeployment, but before the redeployment I was acting as Project Manager (level 5) and after being redeployed I was assigned to programming duties (level e).

Respondent 63

This position has more day-to-day administration (it includes corporate services, finance, PR, IT, etc).

Respondent 69

Had to find my own job.

Respondent 72

The job I was doing I looked after the needs of a workshop where in this job I process files and deliver them.

Respondent 78

Challenging tasks are missing from the new position now that I am familiar with it. Many routine activities (that I have been glad to learn in terms of career development) but that I am becoming bored with. My previous job was much more difficult and interesting than this one.

Respondent 79

I am more secure in my present job and it is better paid as I work longer hours now. I only worked 6½ hours as a tea attendant.

Respondent 81

My previous job required me to be involved in the whole area of government/community and nongovernment requiring assistance in relation to settlement needs of migrants. My current position does not provide job satisfaction, particularly in relation to the offenders that I currently deal with.

Respondent 97

Less responsibility, more specific job, restricted job.

Respondent 105

More variety, more flexibility, more responsibility in the present job, but less confidence, more caution and some resentment.

Figure 15.3 Some selected responses to the open-ended question (no. 11) in [Figure 15.2](#)

Question 12

We would like to know your perception of the two jobs before and after redeployment with respect to the following aspects of your job. Please rate them on a five-point scale using the following guide:

5 = extremely satisfied, 4 = very satisfied, 3 = satisfied, 2 = dissatisfied, 1 = extremely dissatisfied

This is a highly structured question asking respondents to compare on a five-point ordinal scale their level of satisfaction with various areas of their job before and after redeployment. As we are gauging the level of satisfaction before and after redeployment, respondents are expected to give two responses to each area. In this example let us assume you have used the name JOBSTA for job status after redeployment (*job status after redeployment*) and JOBSTB for before redeployment (*job status before redeployment*). Similarly, for the second area, job satisfaction, you have decided that the variable name JOBSATA, will stand for the level of job satisfaction after redeployment (*job satisfaction after redeployment*) and JOBSATB for the level before redeployment (*job satisfaction before redeployment*). Other variable names have been similarly assigned. In this example the variable JOBSTA is entered in column 26, JOBSTB in column 27, and so on.

Stage II: Pre-testing the code book

Once a code book is designed, it is important to pre-test it for any problems before you code your data. A **pre-test** involves selecting a few questionnaires/interview schedules and actually coding the responses to ascertain any problems in coding. It is possible that you may not have provided for some responses and therefore will be unable to code them. Change your code book, if you need to, in light of the pre-test.

Stage III: Coding the data

Once your code book is finalised, the next stage is to code the raw data. There are three ways of doing this:

- 1. coding on the questionnaire/interview schedule itself, if space for coding was provided at the time of constructing the research instrument;**
- 2. coding on separate code sheets that are available for purchase;**
- 3. coding directly into the computer using a program such as SPSS or SAS.**

To explain the process of coding let us take the same questions that were used in developing the code book. We select three questionnaires at random from a total of 109 respondents ([Figures 15.4–15.6](#)). Using the code book as a guide, we code the information from these sheets onto the coding sheet ([Figure 15.7](#)). Let us examine the coding process by taking respondent 3 ([Figure 15.4](#)).

Respondent 3

The total number of respondents is more than 99 and this is the third questionnaire, so the identification number 003 was entered in columns 1–3 ([Figure 15.7](#)). Because it is the first record for this respondent, 1 was entered in column 4. This respondent is 49 years of age and falls in the category 45–49, which was coded as 6. As the information on age belongs in column 5, 6 was entered in this column of the code sheet. The marital status of this person is ‘divorced’, hence 4 was entered in column 6. This person has a bachelor’s degree in librarianship. The code chosen for a bachelor’s degree is 3, which was entered in column 7. Three tertiary qualifications have been provided for,

and as this person does not have any other qualifications, so TEDU2 and TEDU3 are “not applicable”, and therefore a code of 8 is entered in columns 8 and 9. This person’s bachelor’s degree is in librarianship, for which code 09 was assigned and entered in columns 10–11. Since there is only one qualification, STUDY2 and STUDY3 are “not applicable”; therefore, a code of 88 was entered in columns 12–13 and 14–15. This person has given a number of responses to question no. 11 (DIFWK), which asks respondents to list the main differences between their jobs before and after redeployment. In coding such questions much caution is required.

Examine the responses named DIFWK1, DIFWK2, DIFWK3, DIFWK4 AND DIFWK5, to identify the codes that can be assigned. A code of 22 (now deal with public) was assigned to one of the responses, which we enter in columns 16–17. The second difference, DIFWK2, was assigned a code of 69 (totally different skill required), which is coded in columns 18–19. DIFWK3 was assigned a code of 77 (current job more structure) and coded in columns 20–21. Similarly, the fourth (DIFWK4) and the fifth (DIFWK5) difference in the jobs before and after redeployment are coded as 78 (now part of the team instead of independent worker) and 38 (hours – now full-time), which are entered in columns 22–23 and 24–25, respectively. Question 12 is extremely simple to code. Each area of a job has two columns, one for before and the other for after. Job status (JOBST) is divided into two variables, JOBSTA for a respondent’s level of satisfaction after redeployment and JOBSTB for his/her level before redeployment. JOBSTA is entered in column 26 and JOBSTB in column 27. For JOBSTA the code 5 (as marked by the respondent) was entered in column 26 and for JOBSTB the code 4 is entered in column 27. Other areas of the job before and after redeployment are similarly coded.

The other two examples are coded in the same manner. The coded data is shown in [Figure 15.7](#). In the process of coding you might find some responses that do not fit your predetermined categories. If so, assign them a code and add these to your code book.

Stage IV: Verifying the coded data

Once the data is coded, select a few research instruments at random and record the responses to identify any discrepancies in coding. Continue to verify coding until you are sure that there are no discrepancies. If there are discrepancies, re-examine the coding.

Developing a frame of analysis

Although a frame of analysis needs to evolve continuously while writing your report, it is desirable to broadly develop it before analysing the data. A **frame of analysis** should specify:

- **which variables you are planning to analyse;**
- **what type of analysis they should be subjected to;**

RESPONDENT = 3

1 Please indicate:

(a) Your current age in completed years: 49

(c) Your marital status: (Please tick)

Currently married _____

Living in a de facto relationship _____

Separated _____

Divorced

Never married _____

2 (b) If tertiary/university, please specify the level achieved and area of study. (Please specify all postgraduate qualifications.)

Level of achievement	Area of study: e.g. engineering, accounting
Associate diploma	
Diploma	
Bachelor	Librarianship
Graduate diploma	
Masters	
PhD	

11 What, in your opinion, are the main differences between your jobs prior to and after redeployment?

Hours-now F.T, Totally different skills required, Now deal with public,

Busier, more structured day and duties, Now part of the team instead of independent worker

12 We would like to know your perception of the two jobs *before* and *after* redeployment with respect to the following aspects of your job. Please rate them on a five-point scale, using the following guide:

5 = extremely satisfied, 4 = very satisfied, 3 = satisfied, 2 = dissatisfied, 1 = extremely dissatisfied

Before redeployment					Areas	After redeployment				
1	2	3	4	5		1	2	3	4	5
			✓		Job status					✓
			✓		Job satisfaction					✓
			✓		Motivation to work					✓
			✓		Interest in the job					✓
			✓		Self-esteem					✓
			✓		Professional competence				✓	
			✓		Peer interaction					✓
			✓		Morale				✓	
			✓		Work environment				✓	
				✓	Social interaction			✓		

Figure 15.4 Some questions from a survey – respondent 3

- **what cross-tabulations you are planning;**
- **which variables you need to combine to construct your major concepts or to develop indices (in formulating a research problem concepts are changed to variables – at this stage change them back to concepts);**
- **which variables are to be subjected to which statistical procedures.**

To illustrate, let us take the example from the survey used in this chapter.

RESPONDENT = 59

1 Please indicate:

(a) Your current age in completed years: 45

(c) Your marital status: (Please tick)

Currently married

Living in a de facto relationship _____

Separated _____

Divorced _____

Never married _____

2 (b) If tertiary/university, please specify the level achieved and area of study. (Please specify all postgraduate qualifications.)

Level of achievement	Area of study: e.g. engineering, accounting
Associate diploma	
Diploma	
Bachelor	<i>Behavioural Sciences</i>
Graduate diploma	
Masters	
PhD	

11 What, in your opinion, are the main differences between your jobs prior to and after redeployment?

Less responsibility, More specific jobs, Restricted scope.

12 We would like to know your perception of the two jobs *before* and *after* redeployment with respect to the following aspects of your job. Please rate them on a five-point scale using the following guide:

5 = extremely satisfied, 4 = very satisfied, 3 = satisfied, 2 = dissatisfied, 1 = extremely dissatisfied

Before redeployment					Areas	After redeployment				
1	2	3	4	5		1	2	3	4	5
			✓		Job status		✓			
			✓✓		Job satisfaction			✓		
			✓✓		Motivation to work		✓			
				✓	Interest in the job			✓		
				✓	Self-esteem			✓		
				✓	Professional competence			✓		
				✓	peer interaction				✓	
				✓	Morale		✓			
				✓	Work environment			✓		
			✓		Social interaction				✓	

Figure 15.5 Some questions from a survey – respondent 59

RESPONDENT = 81

1 Please indicate:

(a) Your current age in completed years: 42

(c) Your marital status: (Please tick)

Currently married

Living in a de facto relationship _____

Separated _____

Divorced _____

Never married

2 (b) If tertiary/university, please specify the level achieved and area of study.
(Please specify all postgraduate qualifications.)

Level of achievement	Area of study: e.g. engineering, accounting
Associate diploma	
Diploma	
Bachelor	Social work
Graduate diploma	
Masters	
PhD	

11 What, in your opinion, are the main differences between your jobs prior to and after redeployment?

My previous job required me to be involved in the whole area of government/community and nongovernment requiring assistance in relation to settlement needs of migrants. My current position does not provide job satisfaction particularly in relation to the offenders that I currently deal with.

12 We would like to know your perception of the two jobs *before* and *after* redeployment with respect to the following aspects of your job. Please rate them on a five-point scale using the following guide:

5 = extremely satisfied, 4 = very satisfied, 3 = satisfied, 2 = dissatisfied, 1 = extremely dissatisfied

Before redeployment					Areas	After redeployment				
1	2	3	4	5		1	2	3	4	5
		✓			Job status		✓			
			✓		Job satisfaction		✓			
				✓	Motivation to work			✓		
			✓		Interest in the job		✓			
			✓		Self-esteem		✓			
		✓			Professional competence		✓			
		✓			Peer interaction					
			✓		Morale		✓			
			✓		Work environment		✓			
		✓			Social interaction		✓			

Figure 15.6 Some questions from a survey – respondent 81

Frequency distributions

A **frequency distribution** groups respondents into the subcategories into which a variable can be divided. Unless you are not planning to use answers to some of the questions, you should have a frequency distribution for all the

variables. Each variable can be specified either separately or collectively in the frame of analysis. To illustrate, they are identified here separately by the names used in the code book. For example, the frame of analysis should include frequency distributions for the following variables:

Programmer										Date																														
Identification number																																								
Statement Number	Marital status	Study area	Difference in work					Level of satisfaction					Fortrain S																											
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	
0	0	3	1	6	4	3	8	8	0	9	8	8	8	2	2	6	9	7	7	7	7	8	3	8	5	4	5	4	5	4	5	4	5	4	5	4	5	4	5	
0	5	9	1	6	1	3	8	8	0	1	8	8	8	8	0	3	1	9	7	7	8	8	8	8	2	4	4	3	4	2	4	3	5	3	5	3	5	3	5	
0	8	1	1	5	1	2	8	8	1	3	8	8	8	8	1	7	0	1	8	8	8	8	8	8	2	3	2	4	3	5	2	4	2	4	2	4	2	4	2	4

Figure 15.7 An example of coded data on a code sheet

- AGE;

- MS;
- TEDU (TEDU1, TEDU2, TEDU3 – multiple responses, to be collectively analysed);
- STUDY (STUDY1, STUDY2, STUDY 3 – multiple responses, to be collectively analysed);
- DIFWK (DIFWK1, DIFWK2, DIFWK3, DIFWK4, DIFWK5 – multiple responses, to be collectively analysed);
- JOBSTA, JOBSTB;
- JOBSATA, JOBSATB;
- MOTIVA, MOTIVB.
- etc.

Cross-tabulations

Cross-tabulations analyse two variables, usually independent and dependent or attribute and dependent, to determine if there is a relationship between them. The subcategories of both the variables are cross-tabulated to ascertain if a relationship exists between them. Usually, the absolute number of respondents in each cell, and the row and column percentages, give you a reasonably good idea as to the possible association.

In the study we cited as an example in this chapter, one of the main variables to be explained is the level of satisfaction with jobs before and after redeployment. We developed two indices of satisfaction:

- 1. satisfaction with the job before redeployment (JOBSATB);**
- 2. satisfaction with the job after redeployment (JOBSATA).**

Differences in the level of satisfaction can be affected by a number of personal attributes such as the age, education, training and marital status of the respondents. Cross-tabulations help to identify which attributes affect the levels of satisfaction. Theoretically, it is possible to correlate any variables, but it is advisable to be selective or an enormous number of tables will result. Normally only those variables that you think have an effect on the dependent variable should be correlated. The following cross-tabulations are an example of the basis of a frame of analysis: JOBSATA and JOBSATB by AGE, MS, TEDU, STUDY, and DIFWK (these determine whether job satisfaction before and after redeployment is affected by age, marital status, education, and so on); and JOBSATA by JOBSATB (this will reveal whether there is a

relationship between job satisfaction before and after redeployment). You can specify as many variables as you want.

Reconstructing the main concepts

There may be places in a research instrument where you look for answers through a number of questions about different aspects of the same issue, for example the level of satisfaction with jobs before and after redeployment (JOBSATB and JOBSATA). In the questionnaire there were 10 aspects of a job about which respondents were asked to identify their level of satisfaction before and after redeployment. The level of satisfaction may vary from aspect to aspect. Though it is important to know respondents' reactions to each aspect, it is equally important to gauge an overall index of their satisfaction. You must therefore decide, before you actually analyse data, how you will combine responses to different questions.

In this example the respondents indicated their level of satisfaction by selecting one of the five response categories. A satisfaction index was developed by assigning a numerical value to the response given by a respondent. The numerical values corresponding to the categories ticked were added to determine the satisfaction index. The satisfaction index score for a respondent varies between 10 and 50. The interpretation of the score is dependent upon the way the numerical values are assigned. In this example the higher the score, the higher the level of satisfaction.

Statistical procedures

In this section you should list the statistical procedures that you want to subject your data to. You should identify the procedures followed by the list of variables that will be subjected to those procedures – for example, regression analysis (JOBSATA on JOBSATB); multiple regression analysis (JOBSATA on AGE, EDUCATION and MS; JOBSATB on AGE, EDUCATION and MS); and analysis of variance. Similarly, it may be necessary to think about and specify the different variables to be subjected to the various statistical procedures. There are a number of user-friendly programs such as SPSS and SAS that you can easily learn.

Analysing quantitative data manually

Coded data can be analysed manually or with the help of a computer. If the

Figure 15.8 Manual analysis using graph paper

If you want to analyse data using a computer, you should be familiar with the appropriate program. You should know how to create a data file, how to use the procedures involved, what statistical tests to apply and how to interpret them. Obviously in this area knowledge of computers and statistics plays an important role.

Data processing in qualitative studies

How you process and analyse data in a qualitative study depends upon how you plan to communicate the findings. Broadly, there are three ways in which you can write about your findings in qualitative research:



CODING IN QUALITATIVE RESEARCH

- (1) developing a narrative to describe a situation, episode, event or instance;**
- (2) identifying the main themes that emerge from your field notes or transcription of your in-depth interviews and writing about them, quoting extensively verbatim; and**
- (3) in addition to (2) above, also quantifying, by indicating their frequency of occurrence, the main themes in order to provide their prevalence.**

Editing, as understood for quantitative studies, is inappropriate for qualitative research. However, it is possible that you may be able to go through your notes to identify if something does not make sense. In such an event, you may be able to recall the context and correct the contents, but be careful in doing so as inability to recall precisely may introduce inaccuracies (due to **recall error**) in your description. Another way of ensuring whether you are truly reflecting the situation is to transcribe the interviews or observational notes and share them with the respondents or research participants for confirmation and approval. Validation of the information by a respondent is an important aspect of ensuring the accuracy of data collected through unstructured interviews.

For writing in narrative form there is no analysis per se; however, you need

to think through the sequence in which you need or want to narrate. For the other two ways of writing about the findings, you need to go through content analysis, as mentioned earlier. Content analysis means analysing the contents of interviews or observational field notes in order to identify the main themes that emerge from the responses given by your respondents or the observation notes made by you. This process involves a number of steps:

Step 1 Identify the main themes. You need to carefully go through descriptive responses given by your respondents to each question in order to understand the meaning they communicate. From these responses you develop broad themes that reflect these meanings. You will notice that people use different words and language to express themselves. It is important for you to select the wording of your themes in a way that accurately represents the meaning of the responses categorised under a theme. These themes become the basis for analysing the text of unstructured interviews. Similarly, you need to go through your field notes to identify the main themes.

Step 2 Assign codes to the main themes. Whether or not you assign a code to a main theme is dependent upon whether or not you want to count the number of times a theme has occurred in an interview. If you decide to count these themes you should, at random, select a few responses to an open-ended question or from your observational or discussion notes and identify the main themes. You continue to identify these themes from the same question till you have reached saturation point. Write these themes and assign a code to each of them, using numbers or keywords, otherwise just identify the main themes.

Step 3 Classify responses under the main themes. Having identified the themes, the next step is to go through the transcripts of all your interviews or your notes and classify the responses or contents of the notes under the different themes. You can also use a computer program such as NVivo, ATLAS.ti or MAXQDA for undertaking this thematic analysis. You will benefit from learning one of these programs if your data is suitable for such analysis.

Step 4 Integrate themes and responses into the text of your report. Having identified responses that fall within different themes, the next step is to integrate them into the text of your report. How you do so is mainly your choice. Some people, while discussing the main

themes that emerged from their study, use verbatim responses to keep the ‘feel’ of the responses. Others count how frequently a theme has occurred, and then provide a sample of the responses. It entirely depends upon the way you want to communicate the findings to your readers.



Content analysis in qualitative research – an example

The above four-step process was applied to a study recently carried out by the author to develop an operational service model, based upon the principle of family engagement. The information was predominantly gathered through in-depth and focus group discussions with clients, service providers and service managers. After informal talks with a number of stakeholders, a list of possible issues was drawn up to form the basis of discussions in these in-depth interviews and group discussions. The list was merely a guiding framework and was open to inclusion of any new issue that emerged during the discussions. Out of the several issues that were identified to examine various aspects of the model, here the author has taken only one to show the process of identifying themes that emerged. Note that these themes have not been quantified. They are substantiated by verbatims, which is one of the main differences between qualitative and quantitative research. The following example shows perceived strengths of the family engagement model (FEM) as identified by the stakeholders during in-depth interviews and focus groups. Information provided in [Figure 15.8](#) serves as an example of the outcome of this process.

Example

Developing themes through content analysis

Perceived strengths of the model

The framework developed for the perceived strengths of the model is based upon the analysis of the information gathered, which suggested that the various themes that emerged during the data collection stage reflecting strengths of the model can be classified under four perspectives. The following diagram shows the framework that emerged from the analysis.

Perceived strengths from the perspective of the family

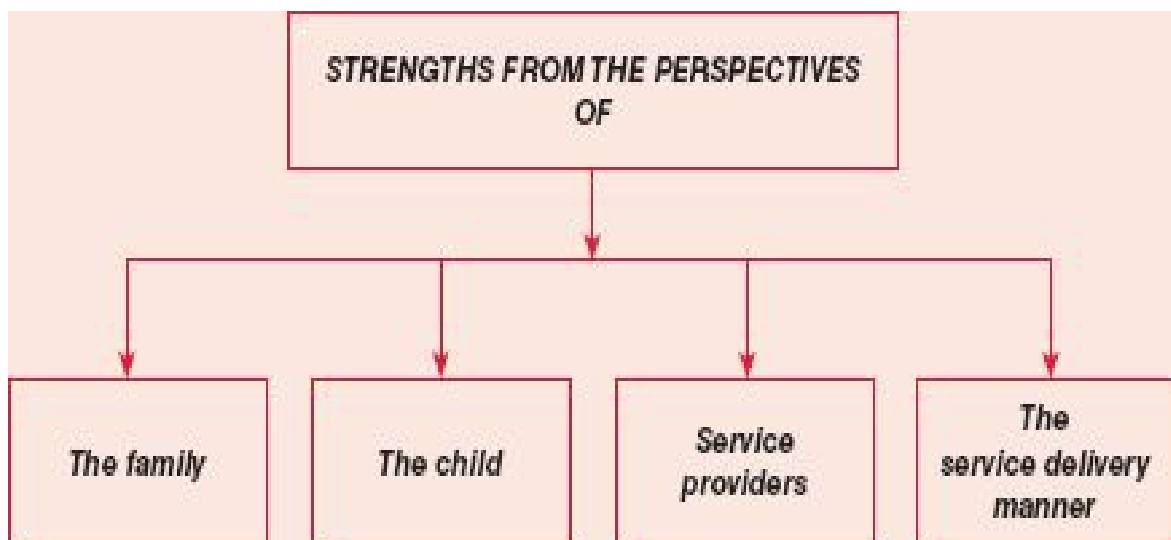
This section details the perceived strengths of the model from the perspective of the family. Keep in mind that the sequential order of the perceived strengths is random and does not reflect any order or preference. Also, the naming of these themes is that of the author, which to the best of his knowledge captured the ‘meanings’ of the intentions of the research participants.

Empowerment of families

Almost everyone expressed the opinion that one of the main strengths of the model is that it empowers families and clients to deal with their own problems. The model provides an opportunity to families to express their feelings about issues of concern to them and, to some extent, to take control of their situations themselves. It seems that in ‘preparing a plan for a child under this model, the family of the child will play an extremely important role in deciding about the future of the child, which is the greatest strength of the model’. One of the respondents expressed his/her opinion as follows:

Oh, the family engagement model actually gives the power back to

the family but with the bottom line in place, like the Department's bottom lines, they have to meet them. Oh ... the old model would have been black and white; kids remain in Mum's care, he [the father] would have supervised contacts with kids and it all would have been set up ... the family engagement model was about pulling them in the whole family then coming up with the solutions as long as they reach the Department's bottom line. They actually have to come up and nominate what they were willing to do ... He [the father] returned home, which was much better ... If they have relapse we bring them back in and we talk about it, get them back on track, make sure they were engaging with the services ... In the old method, kids just would have been removed and kids would have gone into the Department's care ... It is more empowering to the family, and it is much easier to work with the family at that level than you are standing over and telling them that you have to do this and this, and holding it against them that if you do not do, well, the kids are out. It is much, much better for the families. You've got more opportunity to work with the family at that level, rather than being on the outside dictating.



Different perspectives classifying perceived strengths of the model.

Another participant said:

I think this model empowers the family a lot more ... you are having

meetings all the time. You give them the bottom line, and they develop their own strategy ... I think it empowers the family when they come back ... because they are developing their plan, they are using their own network and resources ... I think it is empowering.

Yet, according to another respondent: 'it allows them to feel that they can make some decisions ... They are able to work with the Department and that their voices or views are as valid as the Department's.'

Building of capacity of families

Another advantage that came out of the discussions is that the process adopted as a part of the FEM makes clients aware of community resources which, in turn, help them to build their capacity to deal with a situation effectively and independently. As one participant pointed out:

They know that, ok, if something goes wrong in this aspect of their life, they know they can go there for support, they do not need to be calling us ... they may have resolved their own issues ... that is really empowering.

Another participant said: 'under this model, a family has taken a much stronger role in bringing about change as compared to the case conference approach.'

Acknowledgement of positives in families

One of the strengths of the approach is that it acknowledges the strengths of families. The model is primarily based upon designing interventions based on the strengths and positives of a situation rather than on the negatives. 'In the old model the strengths were not acknowledged to a large degree and certainly not of the parents.' During one of the focus groups, a participant expressed the view that:

The family engagement model starts with the strengths of the

family, so bringing the family in at an earlier stage, and trying to get them to help make decisions about what is going to happen to children who are in crisis ... so ... it is involving more people, their extended family, and getting them to come up with a plan.

In another focus group, a respondent said:

It is only because you actually do work with that strength-based approach and you acknowledge it. It is a huge part of what happens. You can actually say to somebody that you are doing so well, it is great to see the change in you, and even though you personally have nothing to do with those changes, you say, well done ... it is so good to see you looking so well ... You get to a point in a process where you are no longer seen as an outsider, you are no longer seen as a prescriptive organisation, but you are seen as a supportive organisation which is actually assisting that person in the process ... This Department has not been good in acknowledging change, we have not been.

Collaborative decision making and solutions

Another strength pointed out by many is that solutions pertaining to a child are now developed in close collaboration with the family, extended family and other appropriate stakeholders, which makes them (decisions and solutions) more acceptable and workable. In one of the focus group interviews a participant expressed this strength of the model as follows:

They come up with answers, they got it. You are up front with the family. It gives the family a very clear idea what exactly is expected of them ... this is what they have done, what are the concerns of the Department ... So by having family support meetings, you are telling the family this is what we feel is happening with the child and these are the things that cannot happen to your children, what we intend to do about it in the future to be able to have them back or to improve their environment ... it is straight in front of them, not behind their back ... Previously I know of a case where a family was not involved in any of the discussions and they did not understand

why their children were removed from them.

Many participants felt very positive about this collaborative approach. They felt that 'having a family support meeting clearly tells them what has happened, what are the intentions of the Department, and how the Department is going to work with them'. Another participant in one of the focus group discussions added:

There are differences between how the meetings are held but, I guess, oh, sometimes to get the family to develop, and remind them of the bottom line, rather than us saying, 'This is what I want to happen'. I mean, obviously in the discussion of the general situation, you make things clear, but you let the family take the responsibility to develop their own plan.

A respondent in an in-depth interview expressed the opinion that, now, 'We are identifying the members of the extended family. Once upon-a-time we just had parents; now you have to go around and search and get them all together to make a decision.'

Yet another respondent, talking about the strengths of the model, said 'now we approach very differently.' According to him/her:

We inform you, we advise you, that the children are at risk ... whatever with the children, we want to sit down and work with you. Throughout this process we want to work with you, and also, plans have been set up. We want you to be a part of that.

It appears that, under this model, decisions are made not by a single individual, but by all those involved. According to one respondent, 'you are sharing responsibility with other agencies and family members; it is not only your decision, it is the decision of everybody.'

Keeps families intact

Some respondents also felt, that in certain cases, 'The children may not even be taken from the family so quickly'. It seems there is a greater attempt to keep the children in the family. One participant

said: 'I am actually working with quite a few kids where they are trying to keep the family together.' Another participant added that 'The apprehension rate has come down substantially.'

Perceived strengths from the perspective of the child

A greater focus on children

Many family workers as well as team leaders felt that the whole approach is a lot more focused on children. The approach is child centred and, at every step, concerns for children form the core of an intervention: 'it is a lot more child-focused as well. Rather than focusing on the parents, it is focusing on how we are going to make this child safe, and how we are going to achieve that.'

Returns children to their parents quicker

Some respondents felt that the new approach helped children to get back to their parents quicker. In one of the focus group discussions, one participant said:

I think it gets the children back to their parents quicker because at the meeting it identifies strong people in the family that can support the parents to keep their kids. So what I found in the office is that some of the kids get back to their parents quicker than through the Case Conference. The Case Conference is every year ... what the families have to jump through by the end of the twelve months at the next meeting ... here it is none of that. It is a strong person, how you are going to support the Mum to keep getting the kids back ... what you do need ... sometimes the kids go back, just like that.

Prevention of removal of children

Some respondents felt that the model actually prevented kids from being removed from their families. According to one participant, 'that is the big advantage of this model; to prevent kids from being removed'.

Perceived strengths from the perspective of service

providers

Greater job satisfaction

Almost every service provider said that their work after the introduction of the model had become ‘much more satisfying’ because ‘it is enabling workers and clients’.

Easier for the workers to work under the model

Many respondents felt that ‘initially it is more work for a worker but, in the long run, it is less work because of the shared responsibility’. According to one participant:

As a case worker, I remember the days when I had to really work so hard to meet so many people and do so many things individually and all the responsibility was on my shoulders ... but now there is a shared responsibility ... you have to do the ground work but when it is done the long term engagement is easier because there are more people involved and they in part make the decisions.

Decline in hostility towards the Department

Another advantage that some respondents pointed out was a decline in hostility among families towards the Department. It was pointed out that though it depended upon the circumstances, there was a feeling that, on the whole, hostility among clients towards the Department had declined. They also felt that though, in the beginning, there might have been hostility, working under the new model, in most situations, made that hostility disappear. In some situations, an increase in hostility was possibly attributable to a situation such as the apprehension of a child. Most respondents were of the opinion that, as compared to the Case Management model, they had experienced far less hostility towards them while working under the new approach.

Increased trust in the staff by clients

Another advantage some workers saw in the new approach was that they felt that clients had started trusting them a lot more. A participant in one of the focus groups described his/her feelings as follows:

They call us now, I do not have to go and look for them. They are calling me now and asking what is this? ... which means they are taking an interest. They are not sitting back and saying, oh well, they are going to tell us or not tell us.

Better rapport with families

Because of the increased number of contacts, it appears that staff were able to build more congenial and trusting relationships with families. In a focus group, one of the respondents said, 'I think the relationship is more respectful and trustful'. Another respondent said:

Family relationships are a little bit better, and a family also understands that a DCD worker is not someone who goes to homes and removes their children ... how horrible people we are, but, by interacting with them they actually understand that we are people at work, and that we are not going to do these things, that is the old way of doing. We are not going to remove a child without saying anything. We have communication with them.

Develops better understanding by workers of the family dynamics

'One of the strengths that I have seen is that it allows the social worker to feel the family dynamics, to think about the dynamics and it allows the families to participate in whatever they want to', said a participant of a focus group.

Fewer aggressive clients

Another obvious difference between the two models, according to some respondents, was fewer aggressive clients. One respondent

said:

There are far less aggressive clients here as compared to other places. I think it is partly because of the approach. Here you very rarely see people who get agitated, it is much more controlled, and it is a calmer atmosphere.

Develops a sense of ownership of a case

One of the advantages of the FEM as pointed out by various respondents was that under the model, 'you feel the case belongs to you; you "own" a child'. Because of this, according to some participants, there was a greater affiliation between the family worker and the child.

Greater community interaction

Another advantage pointed out by family workers was that the model resulted in their having a greater interaction with community agencies and, consequently, had more knowledge of their community and the services available in it. This was primarily because of restructuring under the family engagement model whereby family workers were allocated particular geographical areas, called, 'patches': 'you also develop really nice working relations with those people. You are working together collaboratively towards the goals, and I think, that is really a great benefit', said one participant. There was 'a lot more linking with other agencies' under the model. Not only was the interaction between community agencies increased but, it appears, clients had also started making more use of community agencies.

Greater knowledge about community members

Another benefit of working under the FEM and within a 'patch' was that family workers got to know a lot more community members. According to one respondent, 'The relationship with people in your community is much stronger and widespread.'

Another respondent said:

After a certain time you get to know who lives on what street, family links between people, especially, when you are working with Aboriginal families. Family links are so important, and knowing who is dealing with whom ... knowing who is in the area, what resources you have, makes your job a lot more effective.

Greater control over personal values by workers

Another advantage identified by some respondents was that, with the new model, 'case workers own values and morals cannot be imposed'.

Perceived strengths from the perspective of the service delivery manner

An open, honest and transparent process

The whole process is open to all stakeholders. 'All the cards are on the table', said one participant, and another expressed the opinion that: 'The case worker may be honest, but, I guess, the process, how it was done, was not'. One respondent said that one of the good things about this model was that 'everyone knows what is going on'. According to a respondent:

Another good thing about these family meetings is that there is the parents, there is the family. The parents might have been telling us one thing or a part of the story and Jaime, another part of the story, not telling Uncle Jimmy ... so it is good in a way that everyone knows what is going on. The whole information is there for everyone that is there. So they cannot push it to us and push it to the family. That is another good thing: they all get the same information, and we get and give the same information to them ... and it is amazing the plan they want to come up with ... it is based upon the information given to them.

Another respondent in a focus group said:

And you are actually fighting the parents about the guardianship of the child: at the end of the day that is what you are doing, and, I think, just to have the transparent working relationship within the family engagement model actually makes that process a lot easier because everything is out in the open and when it comes up in the court, they are not going to be surprised.

Greater informality in meetings

Family meetings under the FEM are far less formal: ‘The family members and others are encouraged to say whatever they want to. They can interrupt and stop the chairperson any time, if they disagree. They can even come back later.’ ‘What is important is that the minutes are written up, and the family gets a copy of the minutes so that they can go back home and read the minutes. They can come back to us.’

More frequent review of cases

Many people felt that the model provided an opportunity to review cases more frequently which helped them to achieve goals more quickly and, if an intervention was not working, it helped them to change the intervention. As one participant pointed out, ‘Changes in the plan to reflect the changes in the family dynamics are undertaken frequently.’ Hence, under the model, ‘The plan for a child is continually being reviewed.’ There seemed to be a lot more flexibility in terms of changing a plan under the model.

Increased honesty, transparency and accountability

Some respondents also felt that, because of the transparency and accountability of the process, simply working within the parameters of the model had helped to keep workers honest and accountable. According to one participant, ‘From a practice viewpoint, it allows the social worker to be honest, accountable and to be transparent.’

A fairer approach

Many respondents felt that the FEM was fairer, as it was open, participatory and empowered a family.

Goals set for clients are more attainable and workable

According to one of the participants of a focus group: 'I think the plan of family engagement meetings is more attainable and more workable ... what they are actually capable of doing. We are not setting what they are not going to achieve, so they are not going to fail.' In addition, it seems, because families were involved in developing a plan, they had a feeling of ownership, and hence they attained the tasks set out in it. Another participant was of the opinion that: 'if you are a part of the solution, then you actually have an investment in making the change.'

Equality in relation to expression of opinion

Some respondents felt that the model provided freedom of expression to parties. All involved were free to express their opinions, and they were encouraged to share their views. As long as the bottom line was met, their opinions were taken into consideration in developing a strategy.

A less chaotic process

As one participant observed: 'it is far less chaotic, just the perception of what was going on. They [referring to workers in the CMM] felt a bit chaotic because work was coming in all the time and they were holding on to cases. Here it is more organised', one participant observed. With the old structure, 'case workers were very stressed; they were not operating particularly well'.

A less stressful approach

Many participants felt that the new approach was less stressful because of its many benefits. It was less stressful for them, and for families, as well as children. In one of the focus groups, a participant

expressed his/her feeling in the following words:

You do not feel that I hate to go to this home ... how are they going to react, what are they going to say to me, or how should I leave or how should I protect myself? You do not have to have those stresses now; it is a calmer situation, it is a happier situation and that is good for the kids, not only for us, but for the kids ... it is actually the kids who also benefit from the approach.

Fewer conflicts with families

Many respondents felt that ongoing conflicts with families were far fewer after the introduction of the new model.

Equality regarding choice of a facilitator for meetings

Another strength was that some participants thought that under the new model facilitation work was not only confined to case managers. Under the model, anyone could become a facilitator.

Increased reflection on practice

Some people also felt that the model provided an opportunity to reflect on practice, thus helping them to improve it.

Total responsibility for cases

Some also pointed out that workers have the total responsibility for cases which seemed to be much better from a number of viewpoints. As pointed out by one person, ‘under the model, a field worker is responsible for the total intervention, from A to Z. You do everything in a patch.’

Compliance with government’s child placement policy

One of the participants pointed out that the model actually complied with the government’s legislative obligation to place Aboriginal and Torres Strait Islander children with their families. According to this

participant:

The model actually meets, for Aboriginal and Torres Strait Islanders child placement principle which is now enshrined in our legislation, where it actually states that children will be placed with family, extended family, immediate community and extended community and a non-Aboriginal person is a last resort ... So this model actually meets that.

Data analysis in mixed methods studies

In mixed methods studies you need to combine a number of data analysis procedures depending upon the way data was collected and the way you want to communicate your findings. Suppose you are interested in finding out the changes in the number of tourists coming to an area or country and how do they feel about their visit. To do so you might collect data either from secondary sources, that is, from the records kept by say the Department of Immigration or through cross-sectional surveys over a period of time to provide information on the number of tourists coming in a year. By analysing these numbers you can estimate the change in the number of tourists over a period of time. This type of analysis that helps you to study the trend in the number of students is quantitative in nature. Further, suppose you found out about their attitudes towards the facilities and stay in the area from some respondents through in-depth interviewing. The data collected through in-depth interviewing can be better subjected to content analysis which is qualitative in nature. In mixed/multiple methods studies, as you use two or more methods either belonging to the same or different research paradigms, mostly the methods used in data collection and your intention about the way you want to communicate your findings determine the way the data should be analysed. If the data is collected by two methods which belong to the two paradigms, you need to use different methods for its analysis appropriate to the way the data was collected.

The role of statistics and computers in research

The role of statistics in research is sometimes exaggerated as you do not have

to have a sound statistical knowledge in the designing of a research study (though it may help indirectly). Knowledge of statistics plays a significant role only when you have collected the required information, adhering to the requirements of each operational step of the research process. Up to data collection the role of statistics is limited; it acquires its importance thereafter. This does not mean that without statistics you cannot conduct a research study. You can conduct a perfectly valid study without using any statistics. The use of statistics depends entirely upon the type of study you are conducting, how you want to communicate your findings, who are going to be your readership, the convention in your academic area (some departments place a lot of emphasis on the use of statistical procedures and others do not), and the requirements of your department. Once data is collected you encounter two questions:

- 1. How do I organise this data to understand it?**
- 2. How to understand what the data means?**

In a way, the answer to the first question forms the basis for the answer to the second. Statistics can play a very important role in answering your research questions in such a manner that you are able to quantify, measure, place a level of confidence on the findings, assess the contribution each variable has made in bringing out change, measure the association and relationship between various variables, and help predict what is likely to happen in the light of current trends.

From individual responses, particularly if there are many, it becomes extremely difficult to understand the patterns in the data, so it is important for the data to be aggregated and summarised. It is in this process of aggregation and summarisation that the statistics play a very important role. Simple aggregations such as frequency distribution and cross-tabulation tables and some simple statistical measures such as percentages, means, standard deviations and coefficients of correlation can reduce the volume of data, making it easier to understand. In computing summary measures, it is possible that some information is lost, making it possible to misinterpret the findings. Hence, caution is required when interpreting data.

Statistics also play a vital role in understanding the relationship between variables, particularly when there are more than two. With experience, it is easy to 'read' the relationship between two variables from a table, but not to

quantify this relationship. Statistics help you to ascertain the strength of a relationship. They confirm or contradict what you read from a piece of information, and provide an indication of the strength of the relationship and the level of confidence that can be placed in findings. When there are more than two variables, statistics are also helpful in understanding the interdependence between them and their contribution to a phenomenon or event.

Indirectly, knowledge of statistics helps you at each step of the research process. Knowledge of the problems associated with data analysis, the types of statistical test that can be applied to certain types of variable, and the calculation of summary statistics in relation to the measurement scale used play an important role in a research endeavour. However, you can also carry out a perfectly valid study without using any statistical procedures. This depends upon the objectives of your study.

Computers in research can play a very significant role in many ways. In the modern age you can hardly imagine anyone who is computer illiterate. Almost everyone would use a computer for word-processing. All your research activities, from writing a research proposal to completion of dissertation, can be typed and saved on a computer. You will often need to go through a number of drafts to incorporate the development of new thoughts and improvement of your written communication, and computers can help you make additional changes without redoing everything.



VIDEO TUTORIALS ON SPSS

There are a number of statistical packages, such as IBM SPSS Statistics , SAS and NVivo, that will make it very easy for you to carry out simple or complex data analyses.

Another advantage of computers is in displaying your data the way you want and changing the presentation to better suit your written communication style. Most programmes can help you to present you data the way you want it with extreme ease.

It will pay you to learn the program(s) you are likely to use. In the long run it will save you a great deal of time and money.

Summary

In this chapter you have learnt about data processing. Irrespective of the method of data collection, qualitative or quantitative, the information is called 'raw data' or simply 'data'. The processing of data includes all operations undertaken from collection to analysis either manually or by computer. Data processing in quantitative studies starts with editing, which is basically 'cleaning' your data. This is followed by coding, which entails developing a code book, pre-testing it, the actual coding, and verifying the coded data. In this chapter we have provided a prototype for developing a code book, detailing descriptions of how to develop codes for open-ended and closed questions, and a step-by-step guide to coding data, taking an example from a survey. The chapter also includes detailed information about content analysis. An extended example of content analysis is provided. The chapter also deals with how to treat data for narrative and thematic styles of writing.

Though the development of a frame of analysis continues until you have finished the report, it helps immensely in data analysis to develop this before you begin analysing data. In the frame of analysis the type of analysis to be undertaken (e.g. frequency distribution, cross-tabulation, content analysis), and the statistical procedures to be applied, should be specified.

Computers primarily help by saving labour associated with analysing data manually. Their application in handling complicated statistical and mathematical procedures, word processing, displaying and graphic presentation of the data analysed saves time and increases speed. Statistics are desirable but not essential for a study. The extent of their application depends upon the purpose of the study. Statistics primarily help you to make sense of data, 'read' the data, explore relationships and the interdependence between variables, ascertain the magnitude of an existing relationship or interdependence and place confidence in your findings.

FOR YOU TO THINK ABOUT

- Refamiliarise yourself with the keywords listed at the beginning of this chapter, and if you are uncertain about the meaning or application of any of them revisit them in the chapter before moving on.
- What procedures can you set in place to ensure the accuracy of the information obtained in both quantitative and qualitative studies?
- Thinking of examples from your own area of study, consider the advantages and disadvantages of having used open-ended or closed questions when you come to process your data.
- Assess the role of statistics for a study in your area of interest.

Now, as you have gone through the chapter, try answering the following questions:

- In terms of data processing, what are the advantages and disadvantages of closed and open-ended questions?
- In undertaking content analysis for descriptive information, what considerations would you keep in mind when developing themes?
- Critically examine the applications of statistics with particular reference to quantitative and qualitative research.



Want to learn more? Visit
[http://www.uk.sagepub.com/
kumar4e](http://www.uk.sagepub.com/kumar4e)

or scan this QR code to gain access to a range of online resources to support your study including practice quizzes, videos, weblinks, flashcards, and journal articles.

DISPLAYING DATA

In this chapter you will learn about

- Methods of communicating and displaying analysed data in quantitative and qualitative research
- How to present your data in tables
- Different types of graphs and how to use them to represent your data

Keywords

area chart,
bar diagram,
bivariate,
cumulative frequency polygon,
data display,
frequency graph,
line diagram,
pie chart,
polygon,
polyvariate,
scattergram,
table,
univariate.

At the end of this chapter, you should have an understanding of

- **Different ways of communicating data**
- **The anatomy of a table and types of tables**
- **Different graphics presentations**

Methods of communicating and displaying analysed data

Having analysed the data that you collected through either quantitative, qualitative or mixed/multiple method(s), the next task is to present your findings to your readers. The main purpose of using data display techniques is to make the findings easy and clear to understand, and to provide extensive and comprehensive information in a succinct and effective way. Broadly,

there are two styles of writing in research: firstly, a descriptive and narrative style; and secondly, an analytical style with factual information incorporated in the text. The descriptive and narrative style is more prevalent in qualitative studies and the analytical in quantitative studies. However, in practice you may mix them if the data permits. The choice of a particular style should be determined primarily by the type of writing style the data gathered lends itself to, your impressions/knowledge of your likely readership's preference, your familiarity with research methodology and statistical procedures, your own writing style, and your impression as to which style of writing would communicate the findings more effectively to the readership.

Because of the nature and purpose of investigation in qualitative research, text becomes the dominant and usually the sole mode of communication. In quantitative studies the text is very commonly combined with other forms of data display, the extent of which depends upon your familiarity with them, the purpose of the study and what you think would make it easier for your readership to understand the content and sustain their interest in it. Hence as a researcher it is entirely up to you to decide the best way of communicating your findings to your readers. Within these two approaches there are many ways of presenting information. If your readers are likely to be familiar with 'reading' data, you can use complicated methods of data display; if not, it is wise to keep to simple techniques.

Although there are many ways of displaying data, this chapter is limited to those more commonly used. There are many computer programs that can help you with this task. In this chapter we confine the communication and display of analysed data to the following four methods: text; tables; graphs; and statistical measures.

Text

Text is by far the most common method of communication in both quantitative and qualitative research studies, and perhaps the only method in the latter. It is, therefore, essential that you know how to communicate effectively, keeping in mind the level of understanding, interest in the topic and need for academic and scientific rigour of those for whom you are writing. Your style should strike a balance between academic and scientific rigour on the one hand, and attracting and sustaining the interest of your readers on the other. Of course, it goes without saying that a reasonable command of the language and clarity of thought are imperative for good

communication.

Your writing should be thematic: that is, written around the various themes of your report; findings should be integrated into the literature citing references using an acceptable system of citation; your writing should follow a logical progression of thought; and the layout should be attractive and pleasing to the eye. Language, in terms of clarity and flow, plays an important role in communication. According to the *Commonwealth of Australia Style Manual* (2002: 49):

The language of well-written documents helps to communicate information effectively. Language is also the means by which writers create the tone or register of a publication and establish relationships with their readers. For these relationships to be productive, the language the writer uses must take full account of the diversity of knowledge, interests and sensitivities within the audience.

Tables

Structure

Other than text, **tables** are the most common method of presenting analysed data in quantitative studies. According to the *Chicago Manual of Style* (1993: 21), 'Tables offer a useful means of presenting large amounts of detailed information in a small space'. According to the *Commonwealth of Australia Style Manual* (2002: 46), 'tables can be a boon for readers. They can dramatically clarify text, provide visual relief, and serve as quick point of reference.' It is, therefore, essential for beginners to know about their structure and types. [Figure 16.1](#) shows the structure of a table.

A table has five parts:

- 1. Title** – This normally indicates the table number and describes the type of data the table contains. It is important to give each table its own number as you will need to refer to the tables when interpreting and discussing the data in the text of the report. The tables should be numbered sequentially as they appear in the text. The procedure for numbering tables is a personal choice. If you are writing an article, simply identifying tables by number is sufficient. In the case of a

dissertation or a report, one way to identify a table is by the chapter number followed by the number of the table within the chapter – the procedure adopted in this book. The main advantage of this procedure is that if it becomes necessary to add or delete a table when revising the report, you only need to change the table numbers for that chapter rather than for the whole report.

The diagram shows a table with the following structure:

- Title:** Table X. YY: Attitudes towards uranium mining by age
- Column heading (x-axis):** Age of respondent
- Stub (y-axis):** Attitude towards uranium mining
- Body:** The main data area of the table.
- Supplementary notes:** Source: Hypothetical data

Attitude towards uranium mining	Age of respondent					Total
	<25	25–34	35–44	45–54	55+	
Strongly in favour						
In favour						
Uncertain						
Against						
Strongly Against						
Total						

Figure 16.1 The structure of a table

The description accompanying the table number must clearly specify the contents of that table. In the description identify the variables about which information is contained in the table, for example ‘Respondents by age’ or ‘Attitudes towards uranium mining’. If a table contains information about two variables, the dependent variable should be identified first in the title, for example ‘Attitudes towards uranium mining [dependent variable] by gender [independent variable]’.

- 2. Stub** – The subcategories of a variable, listed at the head of each row (the left-hand column of the table). According to the *McGraw-Hill Style Manual* (Longyear 1983: 97), ‘the stub, usually the first column

on the left, lists the items about which information is provided in the horizontal rows to the right.’ The Chicago Manual of Style (1993: 331) describes the **stub** as ‘a vertical listing of categories or individuals about which information is given in the columns of the table’.

3. **Column headings** – The subcategories of a variable, listed along the top of the table. In univariate tables (tables displaying information about one variable) the column heading is usually the ‘number of respondents’ and/or the ‘percentage of respondents’ ([Tables 16.1](#) and [16.2](#)). In bivariate tables (tables displaying information about two variables) it is the subcategories of one of the variables displayed in the column headings ([Table 16.3](#)).
4. **Body** – The cells housing the analysed data.
5. **Supplementary notes or footnotes** – There are four types of footnote: source notes; other general notes; notes on specific parts of the table; and notes on the level of probability (*Chicago Manual of Style* 1993: 333). If the data is taken from another source, you have an obligation to acknowledge this. The source should be identified at the bottom of the table, and labelled by the word ‘source’, as in [Figure 16.1](#). Similarly, other explanatory notes should be added at the bottom of a table.

Types of table

Depending upon the number of variables about which information is displayed, tables can be categorised as:

- univariate (also known as frequency tables) – containing information about one variable, for example [Tables 16.1](#) and [16.2](#);
- bivariate (also known as cross-tabulations) – containing information about two variables, for example [Table 16.3](#); and
- polyvariate or multivariate – containing information about more than two variables, for example [Table 16.4](#).

Table 16.1 Respondents by age (frequency table for one population – hypothetical data)

Age	No. of respondents
<20 years	2 (2.0)
20–24	12 (12.0)
25–29	22 (22.0)
30–34	14 (14.0)
35–39	17 (17.0)
40–44	10 (10.0)
45–49	11 (11.0)
50–54	9 (9.0)
55+	3 (3.0)
Total	100 (100.0)

Note: Figures in parentheses are percentages.

Table 16.2 Respondents by age (frequency table comparing two populations – hypothetical data)

Age	Population A	Population B
<20	2 (2.0)	1 (0.6)
20–24	12 (12.0)	17 (10.9)
25–29	22 (22.0)	23 (14.7)
30–34	14 (14.0)	18 (11.5)
35–39	17 (17.0)	26 (16.7)
40–44	10 (10.0)	16 (10.3)
45–49	11 (11.0)	18 (11.5)
50–54	9 (9.0)	27 (17.3)
55+	3 (3.0)	10 (6.4)
No response	0 (0.0)	0 (0.0)
Total	100 (100.0)	156 (100.0*)

Note: Figures in parentheses are percentages (*these sum to 99.9 due to rounding).

Table 16.3 Respondents by attitude towards uranium mining and age (cross-tabulation – hypothetical data)

Attitude towards uranium mining	Age					Total
	<25	25-34	35-44	45-54	55+	
Strongly in favour	(0.0)* 0 (0.0) [†] (0.0)*	(5.5) 2 (12.5) [†] (8.3)	(14.8) 4 (25.0) [†] (18.5)	(35.0) 7 (43.6) [†] (20.0)	(100.0) 3 (18.6) [†] (0.0)	16 (100.0) [†]
In favour	(0.0) 0 (0.0) [†]	(25.0) 3 (0.0)	(41.7) 5 (7.4)	(33.3) 4 (20.0)	(0.0) 0 (0.0)	12 (100.0)
Uncertain	(0.0) 0 (14.3) [†]	(0.0) 0 (19.4)	(33.3) 2 (3.7)	(66.7) 4 (0.0)	(66.7) 0 (0.0)	6 (100.0)
Against	(20.0) 2 (85.7) [†]	(70.0) 7 (66.7)	(10.0) 1 (55.6)	(0.0) 0 (25.0)	(0.0) 0 (0.0)	10 (100.0)
Strongly against	(21.4) 12 (100.0) [†]	(42.9) 24 (100.0)	(26.8) 15 (100.0)	(8.9) 5 (100.0)	(0.0) 0 (100.0)	56 (100.0)
Total	14	36	27	20	3	100

* =column percentage; † = row percentage.

Table 16.4 Attitude towards uranium mining by age and gender (hypothetical data)

Attitude towards uranium mining	Number of respondents												
	<25		25-34		35-44		45-54		55+		Total		
	F	M	F	M	F	M	F	M	F	M	F	M	T
Strongly in favour	0	0	1	1	3	1	5	2	3	0	12	4	16
In favour	0	0	1	2	3	2	3	1	0	0	7	5	12
Uncertain	0	0	0	0	1	1	2	2	0	0	3	3	6
Against	1	1	4	3	1	0	0	0	0	0	6	4	10
Strongly against	4	8	17	7	8	7	2	3	0	0	31	25	56
Total	5	9	23	13	16	11	12	8	3	0	59	41	100

Types of percentage

The abilities to interpret data accurately and to communicate findings effectively are important skills for a researcher. For accurate and effective interpretation of data, you may need to calculate measures such as percentages, cumulative percentages or ratios. It is also sometimes important

to apply other statistical procedures to data. The use of percentages is a common procedure in the interpretation of data. There are three types of percentage: 'row', 'column' and 'total'. It is important to understand the relevance, interpretation and significance of each. Let us take some examples.

[Tables 16.1](#) and [16.2](#) are univariate or frequency tables. In any univariate table, percentages calculate the magnitude of each subcategory of the variable out of a constant number (100). Such a table shows what would have been the expected number of respondents in each subcategory had there been 100 respondents. Percentages in a univariate table play a more important role when two or more samples or populations are being compared ([Table 16.2](#)). As the total number of respondents in each sample or population group normally varies, percentages enable you to standardise them against a fixed number (100). This standardisation against 100 enables you to compare the magnitude of the two populations within the different subcategories of a variable.

In a cross-tabulation such as in [Table 16.3](#), the subcategories of both variables are examined in relation to each other. To make this table less congested, we have collapsed the age categories shown in [Table 16.1](#). For such tables you can calculate three different types of percentage – row, column and total – as follows:

- **Row percentage** – This is calculated from the total of *all* the subcategories of one variable that are displayed along a row in different columns, in relation to only one subcategory of the other variable. For example, in [Table 16.3](#) figures in parentheses marked † are the **row percentages** calculated out of the total (16) of all age subcategories of the variable age in relation to only one subcategory of the second variable (i.e. those who are strongly in favour of uranium mining) – in other words, one subcategory of a variable displayed on the stub by all the subcategories of the variable displayed on the column heading of a table. Out of those who are strongly against uranium mining, 21.4 per cent are under the age of 25 years, none is above the age of 55 and the majority (42.9 per cent) are between 25 and 34 years of age ([Table 16.3](#)). This row percentage has thus given you the variation in terms of age among those who are strongly against uranium mining. It has shown how the 56 respondents who are strongly against uranium mining differ in age from one another.

Similarly, you can select any other subcategory of the variable (attitude towards uranium mining) to examine its variation in relation to the other variable, age.

- **Column percentage** – In the same way, you can hold age at a constant level and examine variations in attitude using **column percentages**. For example, suppose you want to find out differences in attitude among 25–34-year-olds towards uranium mining. The age category 25–34 (column) shows that of the 36 respondents, 24 (66.7 per cent) are strongly against while only 2 (5.5 per cent) are strongly in favour of uranium mining. You can do the same by taking any subcategory of the variable age, to examine differences with respect to the different subcategories of the other variable (attitudes towards uranium mining).
- **Total percentage** – This standardises the magnitude of each cell; that is, it gives the percent age of respondents who are classified in the subcategories of one variable in relation to the subcategories of the other variable. For example, what percentage of the total population is constituted by those who are under the age of 25 years and are strongly against uranium mining?

It is possible to sort data for three variables. [Table 16.4](#) (percentages not shown) examines respondents' attitudes in relation to their age and gender. As you add more variables to a table it becomes more complicated to read and more difficult to interpret, but the procedure for interpreting it is the same.

The introduction of the third variable, gender, helps you to find out how the observed association between the two subcategories of the two variables, age and attitude, is distributed in relation to gender. In other words, it helps you to find out how many males and females constitute a particular cell showing the association between the other two variables (attitude towards uranium mining and age of the respondents). For example, [Table 16.4](#) shows that of those who are strongly against uranium mining, 24 (42.9 per cent) are 25–34 years of age. This group comprises 17 (70.8 per cent) females and 7 (29.2 per cent) males. Hence, the table shows that a greater proportion of female than male respondents between the ages of 25 and 34 are strongly against uranium mining. Similarly, you can take any two subcategories of age and attitude and relate these to either subcategory (male/female) of the third

variable, gender.



GRAPHS IN SPSS

Graphs

Graphic presentations constitute the third way of communicating analysed data. Graphic presentations can make analysed data easier to understand and effectively communicate what it is supposed to show. One of the choices you need to make is whether a set of information is best presented as a table, a graph or text. The main objective of a graph is to present data in a way that is easy to interpret and interesting to look at. Your decision to use a graph should be based mainly on this consideration: 'A graph is based entirely on the tabled data and therefore can tell no story that cannot be learnt by inspecting a table. However, graphic representation often makes it easier to see the pertinent features of a set of data' (Minium 1978: 45).

Graphs can be constructed for every type of data – quantitative and qualitative – and for any type of variable (measured on a nominal, ordinal, interval or ratio scale). There are different types of graph, and your decision to use a particular type should be made on the basis of the measurement scale used in the measurement of a variable. It is equally important to keep in mind the measurement scale used in the measurement of a variable when it comes to interpretation. It is not uncommon to find people misinterpreting a graph and drawing wrong conclusions simply because they have overlooked the measurement scale used in the measurement of a variable. The type of graph you choose depends upon the type of data you are displaying. For categorical variables you can construct only bar charts, histograms or pie charts, whereas for continuous variables, in addition to the above, line or trend graphs can also be constructed. The number of variables shown in a graph is also important in determining the type of graph you can construct.

When constructing a graph of any type it is important to bear the following points in mind:

- **A graphic presentation is constructed on two axes: horizontal and vertical. The horizontal axis is called the 'abscissa' or, more commonly, the x-axis, and the vertical axis is called the 'ordinate' or, more commonly, the y-axis (Minium 1978: 45).**

- If a graph is designed to display only one variable, it is customary, but not essential, to represent the subcategories of the variable along the x-axis and the frequency or count of each subcategory along the y-axis. The point where the axes intersect is considered as the zero point for the y-axis. When a graph presents two variables, one is displayed on each axis and the point where they intersect is considered as the starting or zero point.
- A graph, like a table, should have a title that describes its contents. The axes should also be labelled.
- A graph should be drawn to an appropriate scale. It is important to choose a scale that enables your graph to be neither too small nor too large, and your choice of scale for each axis should result in the spread of axes being roughly proportionate to one another. Sometimes, to fit the spread of the scale (when it is too spread out) on one or both axes, it is necessary to break the scale and alert readers by introducing a break (usually two slanting parallel lines) in the axes.

The histogram

A **histogram** consists of a series of rectangles drawn next to each other without any space between them, each representing the frequency of a category or subcategory ([Figure 16.2](#)). Their height is in proportion to the frequency they represent. The height of the rectangles may represent the absolute or proportional frequency or the percentage of the total. As mentioned, a histogram can be drawn for both categorical and continuous variables. When a histogram is based upon a continuous variable you can interpret the trend exhibited by it, but the same cannot be done if it is based on a categorical variable as you can arrange categorical variables in any order. Therefore when interpreting a histogram you need to take into account whether it is representing categorical or continuous variables. [Figure 16.2](#) provides three examples of histograms using data from [Tables 16.1](#) and [16.4](#). The histogram in [Figure 16.2b](#) is effectively the same as that in [Figure 16.2a](#) but is presented in a three-dimensional style.

The bar chart

The **bar chart** or bar diagram is used for displaying categorical data ([Figure 16.3](#)). A bar chart is identical to a histogram, except that in a bar chart the rectangles representing the various frequencies are spaced, thus indicating

that the data is categorical. The bar chart is used for variables measured on nominal or ordinal scales. The discrete categories are usually displayed along the x-axis and the number or percentage of respondents on the y-axis. However, as illustrated, it is possible to display the discrete categories along the y-axis. The bar chart is an effective way of visually displaying the magnitude of each subcategory of a variable.

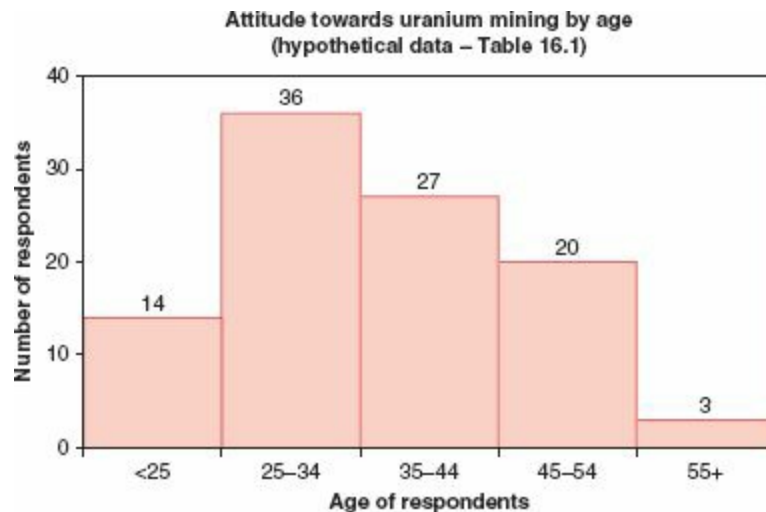


Figure 16.2a Two-dimensional histogram

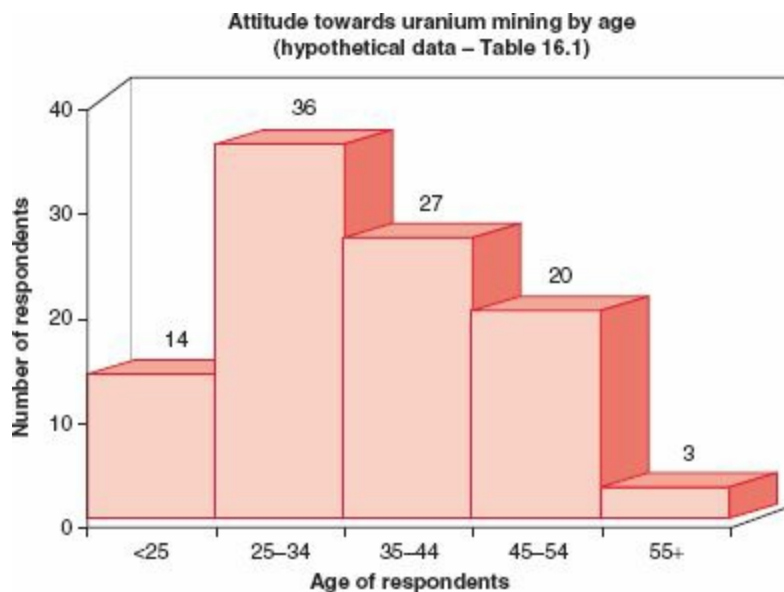


Figure 16.2b Three-dimensional histogram

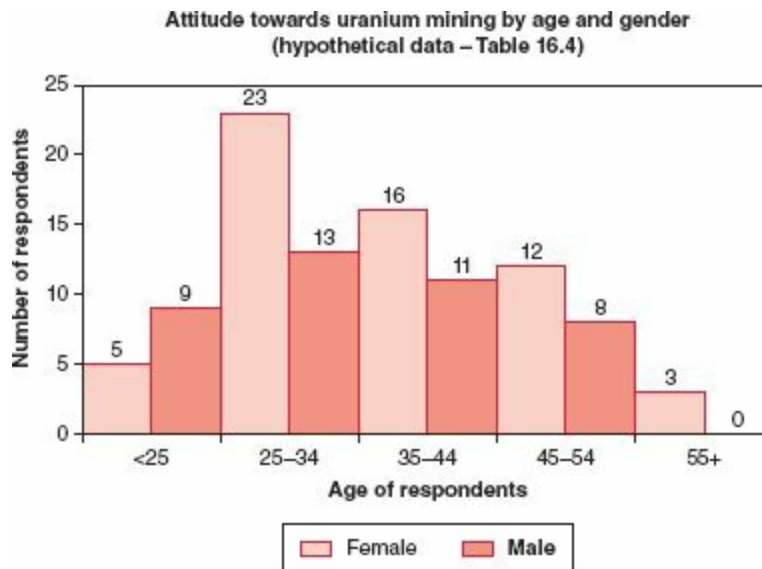


Figure 16.2c Two-dimensional histogram with two variables



STACKED BARS GRAPH IN SPSS

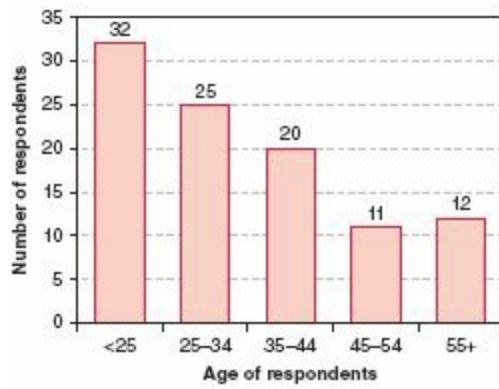
The stacked bar chart

A **stacked bar chart** is similar to a bar chart except that in the former each bar shows information about two or more variables stacked onto each other vertically ([Figure 16.4](#)). The sections of a bar show the proportion of the variables they represent in relation to one another. The stacked bars can be drawn only for categorical data.

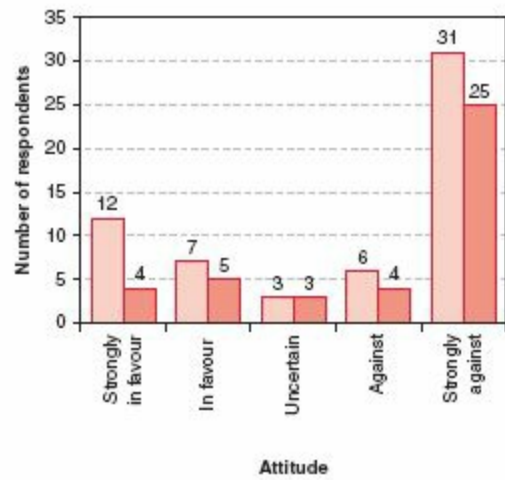
The 100 per cent bar chart

The **100 per cent bar chart** ([Figure 16.5](#)) is very similar to the stacked bar chart. In this case, the subcategories of a variable are converted into percentages of the total population of the subcategory and drawn as a stacked bar chart. Each bar, which totals 100, is sliced into portions relative to the percentage of each subcategory of the variable.

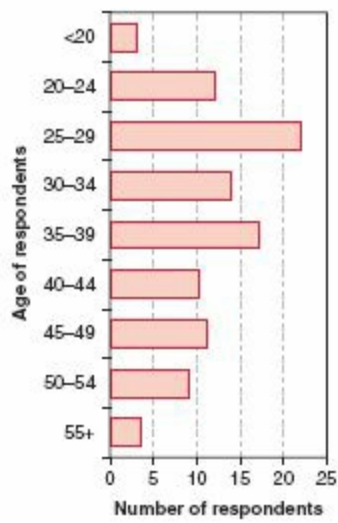
(a) one variable (2D)



(b) two variables (2D)



(c) one variable (vertical box)



(d) two variables (3D)

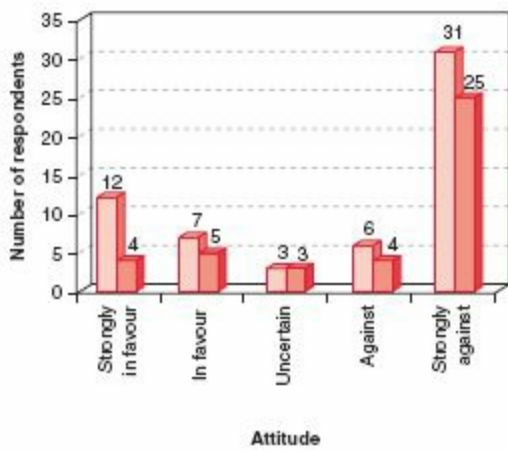


Figure 16.3 Different types of bar chart

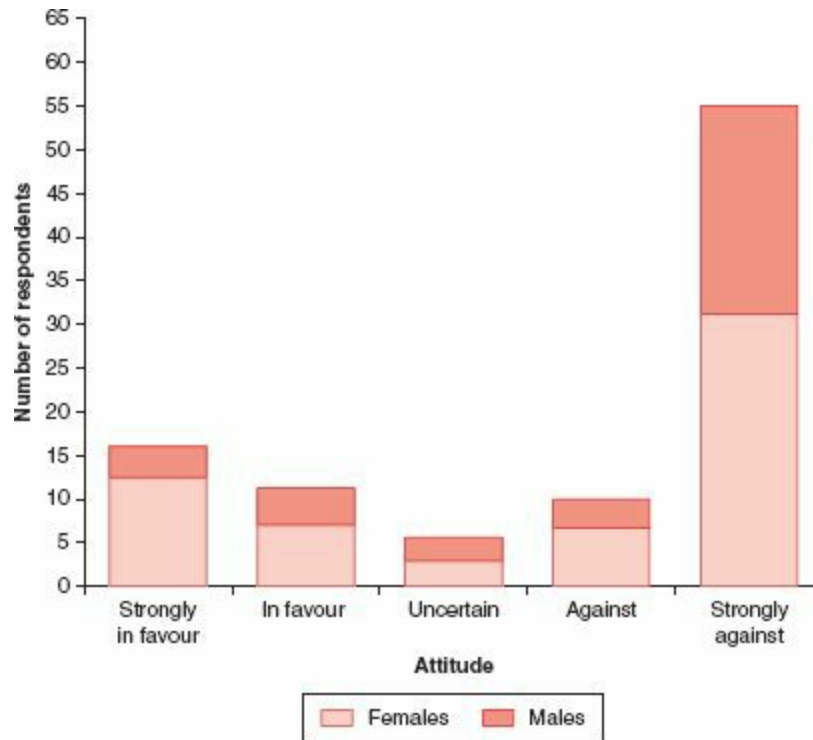


Figure 16.4 The stacked bar chart

The frequency polygon

The **frequency polygon** is very similar to a histogram. A frequency polygon is drawn by joining the midpoint of each rectangle at a height commensurate with the frequency of that interval ([Figure 16.6](#)). However, there is one problem in constructing a frequency polygon, that is, what to do with the two categories at either extreme if they are open-ended (e.g. if one is less than 15 and the other over 60). How do we draw a trend line for these extremes? To overcome this you imagine that the two extreme categories have an interval similar to the rest and assume the frequency in these categories to be zero. From the midpoint of these intervals, you extend the polygon line to meet the x-axis at both ends. A frequency polygon can be drawn using either absolute or proportionate frequencies.

The cumulative frequency polygon

The **cumulative frequency polygon** or cumulative frequency curve ([Figure 16.7](#)) is drawn on the basis of cumulative frequencies. The main difference between a frequency polygon and a cumulative frequency polygon is that the former is drawn by joining the midpoints of the intervals, whereas the latter is

drawn by joining the end points of the intervals because cumulative frequencies interpret data in relation to the upper limit of an interval. As a cumulative frequency distribution tells you the number of observations less than a given value and is usually based upon grouped data, to interpret a frequency distribution the upper limit needs to be taken.

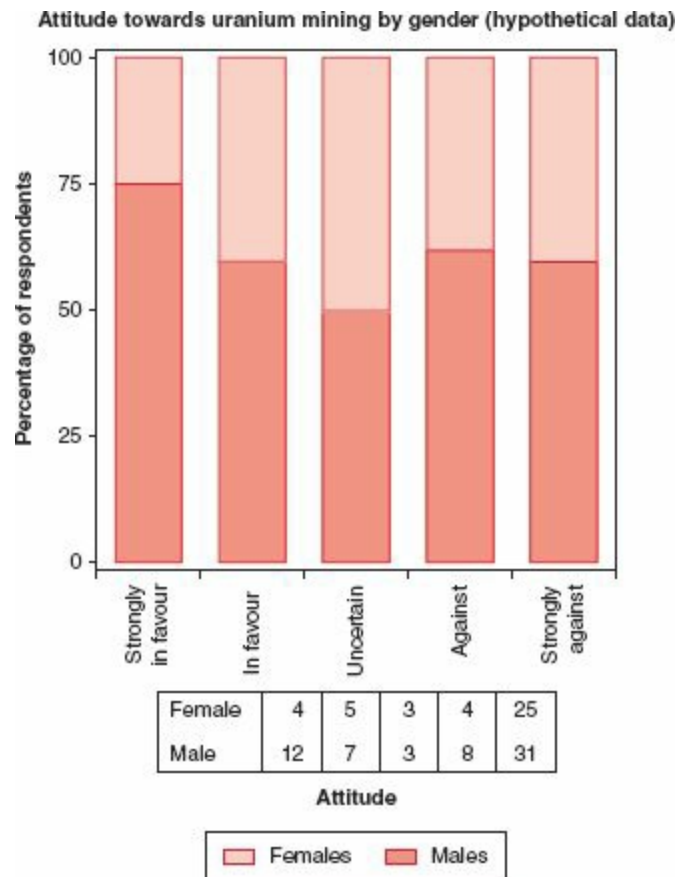


Figure 16.5 The 100 per cent bar chart

The stem-and-leaf display

The **stem-and-leaf display** is an effective, quick and simple way of displaying a frequency distribution ([Figure 16.8](#)). The stem-and-leaf diagram for a frequency distribution running into two digits is plotted by displaying digits 0 to 9 on the left of the y-axis, representing the tens of a frequency. The figures representing the units of a frequency (i.e. the right-hand figure of a two-digit frequency display) are displayed on the right of the y-axis. Note that the stem-and-leaf display does not use grouped data but absolute frequencies. If the display is rotated 90 degrees in an anti-clockwise direction, it

effectively becomes a histogram. With this technique some of the descriptive statistics relating to the frequency distribution, such as the mean, the mode and the median, can easily be ascertained; however, the procedure for their calculation is beyond the scope of this book. Stem-and-leaf displays are also possible for frequencies running into three and four digits (hundreds and thousands).

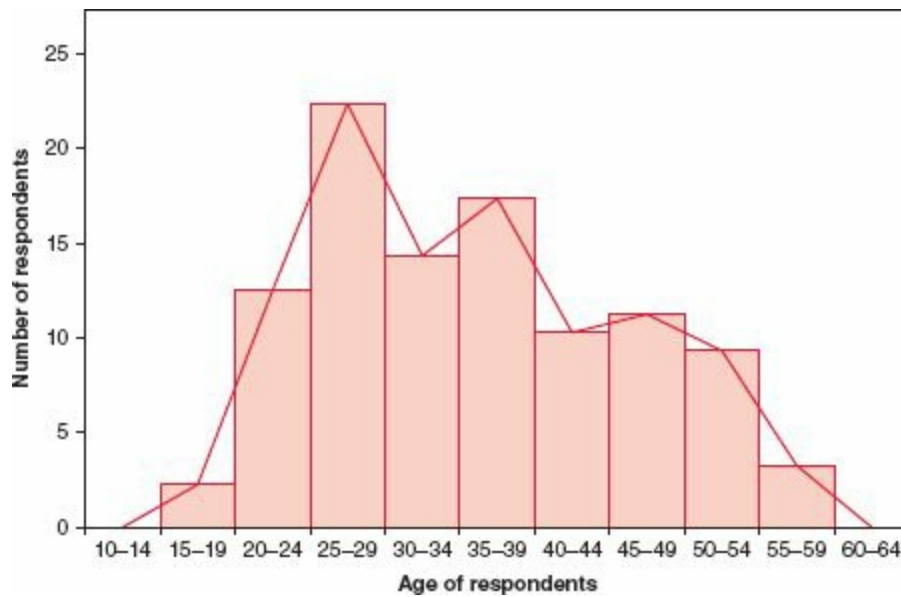


Figure 16.6 The frequency polygon

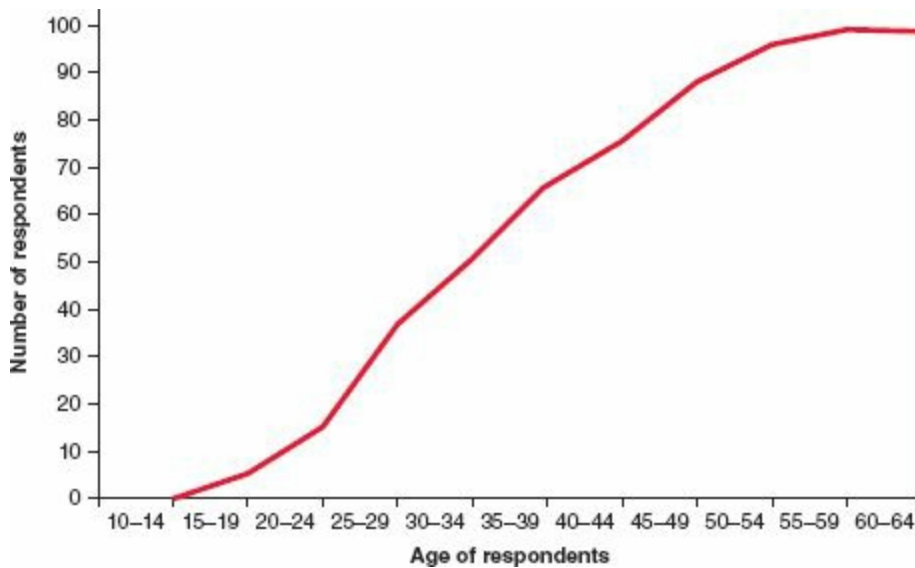


Figure 16.7 The cumulative frequency polygon

1	89
2	000111222222334455555666777788899
3	001112223333334455556667777889
4	112223333444555566779
5	000113355667

Figure 16.8 The stem-and-leaf display

The pie chart

The **pie chart** is another way of representing data graphically ([Figure 16.9](#)), this time as a circle. There are 360 degrees in a circle, and so the full circle can be used to represent 100 per cent, or the total population. The circle or pie is divided into sections in accordance with the magnitude of each subcategory, and so each slice is in proportion to the size of each subcategory of a frequency distribution. The proportions may be shown either as absolute numbers or as percentages. Manually, pie charts are more difficult to draw than other types of graph because of the difficulty in measuring the degrees of the pie/circle. They can be drawn for both qualitative data and variables measured on a continuous scale but grouped into categories.

The line diagram or trend curve

A set of data measured on a continuous interval or a ratio scale can be displayed using a line diagram or **trend curve**. A trend line can be drawn for data pertaining to both a specific time (e.g. 1995, 1996, 1997) or a period (e.g. 1985–1989, 1990–1994, 1995–). If it relates to a period, the midpoint of each interval at a height commensurate with each frequency – as in the case of a frequency polygon – is marked as a dot. These dots are then connected with straight lines to examine trends in a phenomenon. If the data pertains to exact time, a point is plotted at a height commensurate with the frequency. These points are then connected with straight lines. A line diagram is a useful way of visually conveying the changes when long-term trends in a phenomenon or situation need to be studied. For example, [Figure 16.10](#) clearly shows the trends in the number of current uses of methods for contraception over a period of five years among four study populations receiving four treatment modalities comprising different combinations of maternal and child health, nutritional supplements and family planning services. The same information could not have been depicted with such

clarity by a table.

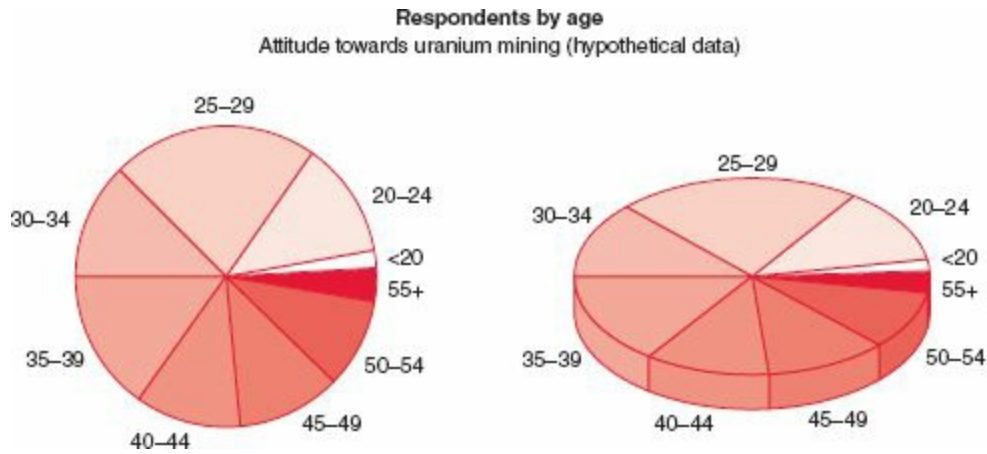


Figure 16.9 Two- and three-dimensional pie charts

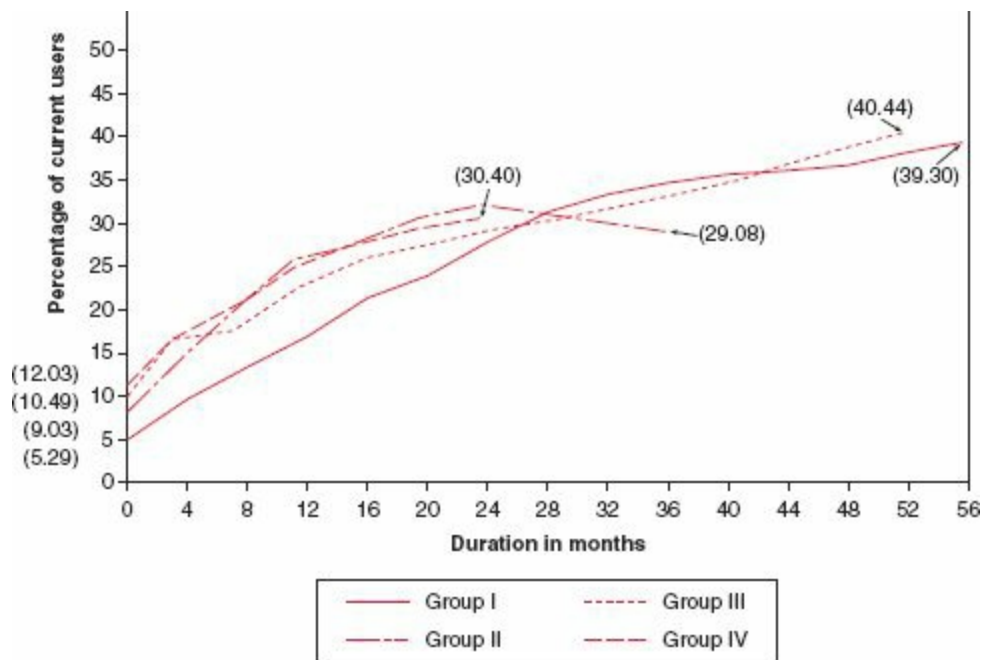


Figure 16.10 The line diagram or trend curve

The area chart

For variables measured on an interval or a ratio scale, information about the subcategories of a variable can also be presented in the form of an **area chart**. This is plotted in the same way as a line diagram but with the area under each line shaded to highlight the total magnitude of the subcategory in relation to other subcategories. For example, [Figure 16.11](#) shows the number

of male and female respondents by age.



SCATTERPLOTS IN SPSS

The scattergram

When you want to show visually how one variable changes in relation to a change in the other variable, a **scattergram** is extremely effective. Both the variables must be measured either on interval or ratio scales and the data on both the variables needs to be available in absolute values for each observation – you cannot plot a scattergram for categorical variables. Data for both variables is taken in pairs and displayed as dots in relation to their values on both axes. Let us take the data on age and income for 10 respondents of a hypothetical study in [Table 16.5](#). The relationship between age and income based upon hypothetical data is shown in [Figure 16.12](#).

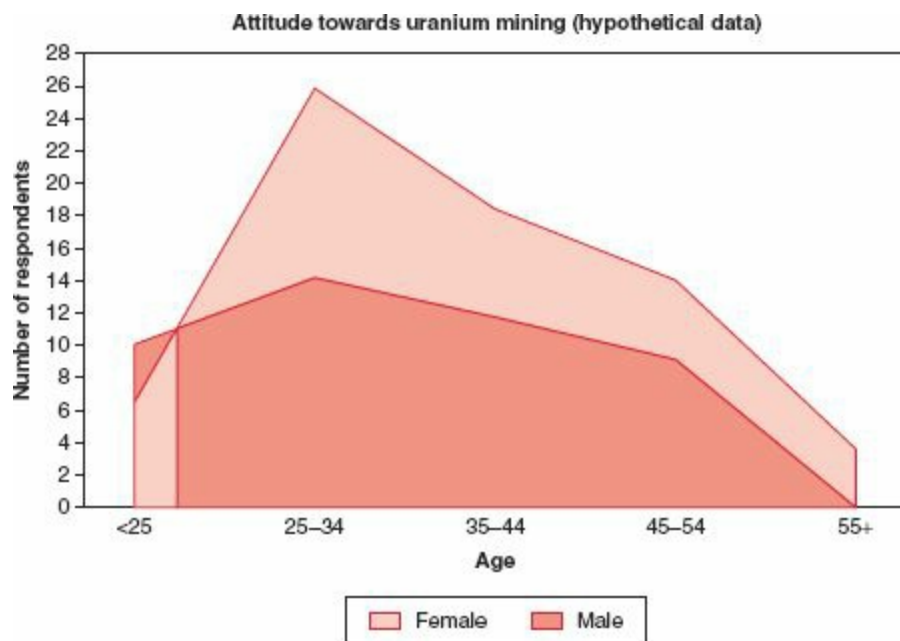


Figure 16.11 The area chart

Table 16.5 Age and income data

Respondent	Income	Age
A	25,500	24
B	46,000	50
C	30,500	36
D	55,000	45
E	27,000	29
F	35,000	38
G	40,000	37
H	52,000	48
I	47,000	41
J	38,000	47

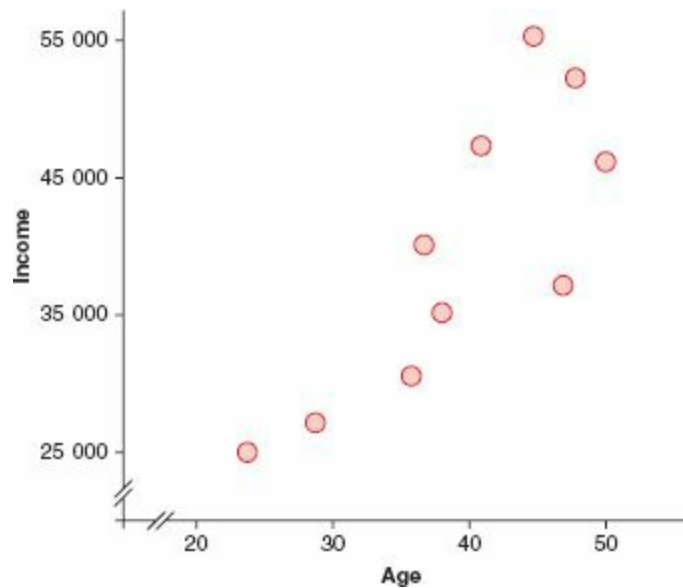


Figure 16.12 The scattergram

Statistical measures

The importance of statistics in research, particularly in quantitative studies, is unquestionably immense. The role of statistics lies in first describing data in aggregated and summary form to enhance its understanding and meaning, to establish whether or not there is a relationship between two variables and, if so, to quantify the magnitude of the relationship. Placing a degree of confidence in the findings is another function of statistics. Many academic disciplines are extremely particular about the integration of statistics into research reports and place immense importance on the application of statistical measures to research findings.



The exclusion of statistics from this book is not an indication of the author's lack of appreciation for statistics, but of the difficulty in accommodating such a vast area in one text. Because of its vastness, the author thought it better for the reader to refer to texts specifically written on statistics.

Though statistical measures are extremely effective in communicating findings in a precise and succinct manner, you can conduct a perfectly valid study without using any statistical measure. However, their use in certain situations is desirable and in some it is essential.

There are many statistical measures, ranging from very simple to extremely complicated. At one end of the spectrum you have simple descriptive measures such as the mean, mode, median, and at the other there are inferential statistical measures such as analysis of variance, factorial analysis and multiple regressions.

Because of its vastness, statistics is considered a separate academic discipline, and before you use these measures you need to learn about them.

Use of statistical measures is dependent upon the type of data collected, your knowledge of statistics, the purpose of communicating the findings, and the knowledge base in statistics of your readership.

Summary

Research findings in both quantitative and qualitative research are usually conveyed to readers through text. In qualitative research this is more or less the sole method of communication. However, in quantitative studies, though text is still the dominant method of communicating research findings, it is often combined with other forms such as tables, graphs and statistical measures, which can make the research easier to understand. What you use should be determined by what you feel comfortable with and what you think will be easiest for readers to understand. Tables have the advantage of containing a great deal of information in a small space, while graphs make it easy for readers to absorb information at a glance.

Usually, a table will have five parts: title, stub, column headings, body and supplementary notes or footnotes. Depending upon the number of variables

about which information in a table is stored, there are three types of table: univariate (frequency), bivariate (cross-tabulation) and multivariate.

To interpret a table, simple arithmetic procedures such as percentages, cumulative frequencies or ratios can be used. You can also calculate simple descriptive statistical procedures such as the mean, the mode, the median, the chi-square test, the *t*-test and the coefficient of correlation. If you have statistical knowledge, advanced statistics can be applied.

Statistics are not included in this book for two reasons: to maintain a sharp focus on research study design; and because the vastness of the subject of statistics would have made this book difficult for beginners in research.

While there are many types of graphs, the common ones are: the histogram, the bar diagram, the stacked bar chart, the 100 per cent bar chart, the frequency polygon, the stem-and-leaf display, the pie chart, the line or trend diagram, the area chart and the scattergram. The choice among these depends upon your purpose and the measurement scale used to measure the variable(s) being displayed. Some graphs are difficult to draw by hand but can be readily done by computer.

FOR YOU TO THINK ABOUT

- Refamiliarise yourself with the keywords listed at the beginning of this chapter, and if you are uncertain about the meaning or application of any of them revisit them in the chapter before moving on.
- Identify two specific examples where you could use a table rather than just text to communicate findings and two examples where graphs would be better.
- Construct a hypothetical bivariate table, within the context of an area of interest. Calculate different types of percentages and interpret the data.

Now, as you have gone through the chapter, try answering the following questions:

- What are the different ways you can communicate your research findings?
- Ideally, a table has a number of parts. Explain them.
- What are the different types of percentage? What functions do they perform in data interpretation?
- What are the different types of graph? Discuss their application in relation to measurement scales.



Want to learn more? Visit
[http://www.uk.sagepub.com/
kumar4e](http://www.uk.sagepub.com/kumar4e)

or scan this QR code to gain access to a range of online resources to support your study including practice quizzes, videos, weblinks, flashcards, and journal articles.