
UNIT 4 DATA COLLECTION

Structure

- 4.0 Introduction
- 4.1 Objectives
- 4.2 The Concept of Data
- 4.3 Methods of Data Collection
 - 4.3.1 Asking Questions
 - 4.3.2 Observation of Behaviour
 - 4.3.3 Utilisation of Existing Records or Data
- 4.4 Ensuring the Quality of Data
- 4.5 Key Points at a Glance
- 4.6 Let Us Sum Up
- 4.7 Glossary
- 4.8 Check Your Progress: The Key

4.0 INTRODUCTION

To carry out a research study, you have to collect the relevant data so that the hypotheses or generalisations you hold tentatively can be verified. This involves selection of samples from the population concerned. In units 1, 2 and 3, we discussed the concepts of population and sample, various sampling techniques along with the different data gathering tools and techniques varying in their complexity, design and ways of administration. Each tool or technique is appropriate for collecting a particular type of data or information, which lends itself to a particular type of analysis and interpretation for drawing meaningful conclusions and generalisations. In this unit we shall focus on various methods used for collection of data.

4.1 OBJECTIVES

On the completion of this Unit, you should be able to:

- define the concept of data;
- describe the various types of data;
- describe the methods of collecting data; and
- suggest precautions which are needed to ensure the quality of data.

4.2 CONCEPT OF DATA

The information collected from various sources through the use of different tools and techniques generally comprise numerical figures, ratings, descriptive narrations, responses to open-ended questions, field notes, etc. This information is called data. Usually, data are categorised as Quantitative and Qualitative data.

1) Quantitative Data

Quantitative data are obtained by applying various scales of measurement. The experiences of people are fit into standard responses to which numerical values are

attached. These data are close-ended and hardly provide any depth or details. Quantitative data are either parametric or non-parametric. Parametric data undergo interval or ratio scale measurement. For example, in measuring reaction time, we make use of ratio scale measurement. The score on a psychological test or inventory is an illustration of interval scale measurement. Non-parametric data are obtained by applying nominal or ordinal scales of measurement. These data are either counted or ranked.

2) **Qualitative Data**

Qualitative data are verbal or symbolic. The detailed descriptions of observed behaviours, people, situations and events, are some examples of qualitative data. For example, the responses to open-ended questions of a questionnaire or a schedule, first hand information from people about their experiences, ideas, beliefs, and selected content or excerpts from documents, case histories, personal diaries and letters are other examples of qualitative data.

Check Your Progress 1

List the various types of data.

Note: a) Write your answer in the space given below.

b) Compare your answer with the one given at the end of this unit.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

4.3 METHODS OF DATA COLLECTION

As discussed earlier, there are mainly three methods of obtaining data (i) one can ask questions; (ii) one can observe the behaviour of persons, groups or organisations, and their products or outcomes; or (iii) one can utilise existing records or data already gathered for purposes other than one’s research. In all the three methods the researcher needs to familiarise himself/herself with the procedure he/she is to adopt for collecting data from sample, groups or records.

4.3.1 Asking Questions

In the first method, the researcher may use psychological tests, inventories, questionnaires, or schedules. In Unit 2 you have learnt that tests are useful tools of research. They are devised to evaluate and measure behaviour in a standardized way for the purpose of providing data for most experimental and descriptive studies. Tests and inventories yield objective and standardized descriptions of behaviour,

quantified in numerical scores. Under ideal situations, intelligence, aptitude or achievement tests measure the best performance of which individuals are capable. The inventories attempt to measure typical behaviour. In experimental studies, a researcher may use test scores to equate the experimental and control groups, to describe relative skill at this task prior to the application of the teaching methods, to assess gains in achievement resulting from the application of the experimental and control teaching methods, and to evaluate the relative effectiveness of teaching methods. Tests and inventories are frequently used to describe prevailing conditions at a particular time in descriptive research studies. For example, achievement tests are used extensively in school surveys in the appraisal of instruction.

In selecting tests or inventories for collecting data in research situations, a researcher must evaluate their validity, reliability and usability. The information about these criteria are available in their accompanying manuals. The researcher should carefully examine the standardized data of the tests/inventories contained in their manuals and extensive analysis of published evaluations of the instruments. It makes the researcher aware about their usefulness and limitations in different test situations.

Ease of administration, scoring, and interpretation, are important factors in selecting a test or inventory. The tests or inventories which are easily and effectively administered, scored, and interpreted should generally be used.

The procedure given in the manual for administering a test or an inventory should be strictly followed to collect dependable data. The cooperation of the subjects must be ensured at each stage of data collection. The subjects should be encouraged to provide objective information. The responses of the subjects should be independent of the personal judgement/view of the researcher who is using the test or inventory. The testing conditions should be made interesting and fascinating so as to gain the cooperation of the subjects.

Questionnaires and interview schedules are the other tools of research through which information is sought. The reliability and validity of the data gathered through questionnaires or interview schedules depends not only on their design but also on the manner of administering the questionnaire or the technique of interviewing. The questionnaire is generally sent through mail to the subjects for answering without any further assistance from the researcher. The schedule, on the other hand, is generally filled out by the researcher who can interpret the questions whenever necessary. In certain situations when the researcher administers the questionnaire personally it creates an opportunity to establish rapport with the subjects, explain them the purpose of the study, and the meaning of items that may not be clear. The availability of a number of subjects in one place helps in exercising economy of time and expense and provides a high proportion of usable responses. However, subjects who have the desired information cannot always be contacted personally without spending of a great deal of time and money in travel. It is in such situations that mailed questionnaires are useful.

The researcher should choose the respondents carefully before administering the questionnaire. It is important that questionnaires be sent only to those who possess the desired information and are interested to respond conscientiously and objectively. It is advisable to send a preliminary letter to respondents individually asking whether the individual would be willing to participate in the proposed study. This is not only a courteous approach but also a practical way of identifying those who will cooperate in furnishing the desired information. The researcher should also consider the possibility of providing for anonymous responses if the desired information is delicate or of confidential in nature. This approach is helpful in producing objective and honest responses.

4.3.2 Observation of Behaviour

Direct observation of the behaviour of persons, groups or organisations provides reliable and conceptually meaningful data in field studies as well as in laboratory experimentation. You have already learnt that observation is the technique in which one or more persons observe what is occurring in some real-life situation. This technique, like other research tools and techniques needs proper planning, expert execution, and adequate recording and interpretation. Observation is always directed towards a specific goal. It is neither haphazard nor unplanned. The planning for observation includes definition of specific activities or units of behaviour to be observed, the nature of the groups of subjects to be observed, determination of the length of each observation and recording. Effective execution of observation ensures proper arrangement of specific conditions for the subject or subjects to be observed, objective and effective use of recording tools, and interpretation of observation data. Observation may be either participant or non-participant in structured or unstructured situations. Structured observations are executed in controlled situations like classroom or laboratory settings. Unstructured observations is mainly associated with participant observation and it is often an exploratory technique. The recording of the observation data may either be simultaneous or soon after the observation. In the former case, the observer goes on recording his/her observation along with the occurrence of the phenomena observed. In the latter case, the observer undertakes to record his/her observations not simultaneously with the actual observation process, but immediately after he/she has observed for a unit of time while the details are still fresh in the mind. As discussed earlier in Unit 3, an observation schedule is developed and used for recording and taking notes about the observed observed behaviours. The specific behaviours to be observed and recorded are listed in this schedule.

Fig. 4.1: Observing behaviour

Observation is done either directly, as when the observer plays a passive role and observes without intervening in any way; or in an interview, where the observer plays a more active role, by asking a series of questions or administering a test, and where he/she observes the behaviour of the subject (interviewee) as well as records his/her responses.

4.3.3 Utilisation of Existing Records or Data

In Unit 3, you read that when the researcher uses the method of observation, either participant or non-participant, he/she focuses on those aspects of the behaviour of an individual or individuals which are of interest to him. If he/she uses tests or questionnaires, he/she chooses or frames the tool to suit his/her research needs. He/She uses interviews if he/she needs information on matters of confidential or personal nature. In using these tools, the researcher controls or manipulates the situation according to his/her research objectives. In contrast, the existing data or records/documents bring to the researcher's notice, certain data over which he/she has relatively little control. These come to the researcher readymade. Some other person, either a participant in a social situation or process, the originator of a system of recording, or the creator of an index, has already determined the form of the data. The data obtained through observation, through tests and questionnaires, and through interviews are gathered for a specific purpose. Documents and records, on the other hand, may bring together data for scientific analysis from remote periods and places. These data provide unique access to historical and social situations too, which are otherwise difficult or expensive to observe. Personal letters, life histories, diaries, autobiographies, court records, proceedings of commissions, seminars and conferences, newspaper stories, registration and census records/information pertaining to births, deaths, marriages, divorces, school attendance, drop-out rate, performance on psychological tests, crimes known to police, arrests, court actions, prison records, voting pattern in the parliament and assembly election, automobile registrations, enrolment of distance learners in different study or regional centers etc. are examples of documents and records.

Check Your Progress 2

Describe briefly the methods of collecting data.

Note: a) Write your answer in the space given below.

b) Compare your answer with the one given at the end of this unit.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

4.4 ENSURING THE QUALITY OF DATA

The adequacy of tool or technique for collection of data is ordinarily judged in terms of the criteria of reliability (consistency), validity and usability. Reliability requires that repeated measurements yield results which are identical or fall within narrow and predictable limits of variability. The criterion of validity demands that

measurement be meaningfully related to the research objectives; that is, it should measure what it purports to measure. The requirements of usability ensure objectivity in the use of a tool or technique and economy of time and cost in field situations. A good tool and its objective use in the collection of data ensures quality.

Psychological tests or inventories are likely to gain the cooperation of subjects and conserve the time of all those involved in their administration provided they are administered in a short period of time. The active cooperation of the subjects is

likely to enhance the quality of the data. Hence, the researcher should take utmost care in selecting a reliable and valid test from the available standardized tests. These tests are easily and effectively administered, scored and interpreted by the researcher. They should also be interesting and enjoyable for the subjects so as to ensure objective data. Boring tests that discourage or antagonise the subjects should not be used for collecting quality data. The testing conditions should be favourable; otherwise, the test is not likely to yield useful and quality data.

The collection of data by means of questionnaires or interview schedules is a highly complicated and technical job which demands considerable effort on the part of the researcher. Much of the quality of the data obtained depends on the skill with which the tools are administered.

Now, we shall discuss some guiding principles which a researcher should consider while using questionnaires or schedules.

The researcher should choose the subjects carefully. It is important to know that only those subjects should be selected who possess the desired information and are likely to be keen to respond conscientiously and objectively. A questionnaire or a schedule, unlike a psychological test or inventory, has a very limited purpose. Hence, the question of validity and reliability of these tools is not dealt with as seriously as in the case of psychological tests or inventories. However, the validity of data gathered through questionnaires or schedules is improved considerably by making the language of the questions less ambiguous. The meaning of all the terms used in the questionnaire/schedule must be clearly defined so that they carry the same meaning for all respondents. The predictive validity of some specific types of questions can also be estimated by follow-up observations of respondent behaviour. The reliability of the responses to the questions can be inferred by a second administration of the tool and comparing the responses with those of the first.

One major problem in interviewing could be due to the inability or unwillingness of the respondent to communicate. His/her involvement in the data in terms of his/her responses and the likelihood of any bias should be looked into carefully by the researcher. It needs to be ensured that the respondent does not withhold or distort facts while communicating them to the researcher. Hence, it is advisable to use other means of data collection to cross-validate the information obtained through a questionnaire or schedule.

Memory bias is another factor that affects the quality of data obtained through interviews. To overcome this problem it is suggested that the research design be carried out over a period of time, applying appropriate tools at reasonable intervals as indicated by the research objectives.

We have already discussed that an interview is a highly flexible tool, provided it is conducted by a skilful researcher. It allows a more liberal atmosphere than in the case of other tools of research. Questions not readily grasped by interviewees should be rephrased, or repeated with proper emphasis and explanations wherever necessary. This is quite useful for handling contradictory statements made by the respondents. It also ensures the quality and consistency of responses. However, questions which are generally vague or obscure should be avoided so as to permit precise answers from the respondents. Sometimes, misunderstandings occur when questions involve usage of technical terms or unfamiliar expressions. Inadequate responses are secured if too many choices are offered for too long during the interview.

For enhancing satisfactory face-to-face relationship between the researcher (interviewer) and the respondent (interviewee), the former should be properly introduced to the latter. General letter of introduction is of little value. The introduction should be personal. The interviewer also needs to choose a 'suitable' time and place for conducting the interview. The most suitable place would be one where it is believed that the interviewee will be most at ease. Politeness on the part of the interviewer is essential for gaining the confidence of the interviewee. It is helpful in gathering reliable and quality responses from the interviewees. In the initial meeting, after friendly greetings are exchanged, the interviewer should explain the purpose of interview to the interviewee. It should be stated in terms that can be understood by the interviewees easily. There will be no difficulty in getting frank and sincere responses from the interviewees if they are confident that the interviewer has no ulterior motives but seeks information only for scientific/research purpose.

Patience and perseverance are two important traits of a good interviewer. Listening to responses of an interviewee, especially in unstructured situations, is hard work. It requires self-restraint, self-discipline, patience and humility. Ability to listen with understanding, respect, and curiosity is the gateway to meaningful communication. An interviewee is likely to provide truthful data if he/she feels that he/she will not meet with interruption, denial, contradictions, and other harassments from the interviewer. He/She is motivated to communicate when the atmosphere is congenial and permissive. However, mere listening is not sufficient. A quiet listener (interviewer) must at the same time be an analytical researcher. Hence, the questions must not only be precise and wisely formulated, but must also be phrased in such a way that they display concern for the interviewee's problems. Some questions are necessary and often unavoidable in a long interview. At times, the interview 'runs dry' and needs restimulation. The description of some incidents lacks clarity or completeness. Hence, it is essential for an interviewer to clarify the doubts, if any, to an interviewee so that the accurate information is made available.

Blunt questions must be avoided as they cause antagonism and withdrawal. Indirect questions are helpful in seeking co-operation of the interviewees. Direct questions

satisfy only the interviewer and the data or information collected through such questions gives rise to unrelated facts or incidents. These data are useless for scientific purposes.

As a data-gathering device, observation also makes an important contribution to descriptive research. A number of devices like check-lists, schedules, rating scales, and score cards are used for collecting and recording observations. Quality of these data mostly depends on the application and use of these devices. The observer should constantly keep in mind that it is easy to become attracted by conspicuous, dramatic and interesting events/situations. However, the researcher should safeguard himself/herself against merely observing unique and striking events which have hardly any relationship with actual reality. The validity and the reliability of measurements are improved when observations are not hurried and are made at frequent intervals by the same observer, or when several observers record their observations independently. An observer must try to minimise the error of 'halo-effect' when using a rating scale or a score card. The tendency to rate someone with a pleasing personality high on other traits such as intelligence or professional interest should be curbed. The halo effect is likely to increase when the observer is asked to rate too many factors or traits; he/she is not trained to judge. Hence, it is advisable to allow only a small number of traits for rating while making observations.

The quality and veracity of data depends greatly on the selection of the tools and their judicious use by researchers. It requires careful calibration of the individuals involved in data collection and also their orientation for the development of skills in the use of various research tools and techniques.

Check Your Progress 3

Describe how you can ensure the quality of data collected?

- 1) Using an interview schedule.
- 2) Using observation?

Note: a) Space is given below for your answer.

b) Compare your answer with the one given at the end of this unit.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

For the identification of data from various documents and records, the researcher must learn to read them with understanding and insight as a basis for being able to interpret the past, which in turn may help in interpreting present trends and possibly in predicting future events. For this the researcher needs to subject the documents/

records to rigorous evaluation. It will involve the dual processes of establishing the authenticity of the source and validity of its contents. This evaluation is called ‘criticism’ of information/data provided by the document/records. The process of establishing authenticity of the data is termed as ‘external criticism’ and that of establishing the validity of their content is termed as internal criticism. External criticism checks the genuineness and authenticity of the source material. For this the researcher has to determine whether it is what it appears or claims to be and whether it matches with the original so as to save himself/herself from being the victim of a fraud. Through internal criticism the researcher established the validity, credibility and merit of the contents of the document.

Check Your Progress 4

Briefly, state the importance of ‘external and internal’ criticism of documents/ records.

Note: a) Write your answer in the space given below.

b) Compare your answer with the one given at the end of this Unit.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

4.5 KEY POINTS AT A GLANCE

- 1) Data are information collected from various sources by using different tools and techniques.
- 2) Data are either quantitative or qualitative
- 3) Parametric data are measured on interval or ratio scales whereas non-parametric data are obtained by applying nominal or ordinal scales of measurement.
- 4) Qualitative data are verbal and symbolic.
- 5) The three methods of obtaining data in research are (i) one can ask people questions, (ii) one can observe the behaviour of persons, groups, or organisations, and their products or outcomes and, (iii) one can utilise existing records or data already gathered for purposes other than one’s research.
- 6) In asking people questions, the researcher may use psychological tests, inventories, questionnaires or schedules.

- 7) Interview schedules provide an opportunity to the researcher to establish rapport with the subjects (interviewees), explain them the purpose of the study, and the meaning of the items that may not be clear to them.
- 8) Observation schedules, rating scales, score cards and check-lists are used for recording data collected through observations.
- 9) Existing records and documents include personal letters, life histories, diaries, autobiographies, court records, proceedings of commissions, seminars and conferences, newspaper stories, registration and census records etc.
- 10) Past records and documents provide data for scientific analysis of remote period and place. These data are in a form over which the researcher has little control.
- 11) A valid, reliable and usable tool ensures quality data.
- 12) The quality of the obtained data depends on the skill with which the tools are used and administered.
- 13) The validity of the data gathered through questionnaires or schedules is improved considerably by making the language of the questions unambiguous.
- 14) The reliability of the responses to the questions can be inferred by a second administration of the tool and then comparing the questions with those of the first.
- 15) The quality of data obtained through interviews mostly depends on the skillful handling of situations by the interviewer. It requires a lot of patience and perseverance on the part of the interviewer.
- 16) A good interview requires self-restraint, self-discipline, patience and humility.
- 17) Criticism of the documents helps in establishing the authenticity and veracity of data/documents/records.

4.6 LET US SUM UP

In this unit we elaborated on the concept of data, methods of data collection and the precautions which are needed for ensuring quality data.

- 1) The information collected from various sources with the help of different tools and techniques generally comprises figures, ratings, descriptive narrations, responses to open-ended questions, field notes, life-histories, proceedings of seminars, conferences, etc. All these are called data.
- 2) Data are either quantitative or qualitative. Quantitative data are parametric or non-parametric.
- 3) Parametric data are measured by interval or ratio scales. Non-parametric data are obtained by applying nominal or ordinal scales of measurement. These data are either counted or ranked.
- 4) Qualitative data are verbal or symbolic materials. The detailed descriptions of observed behaviours, people, situations and events, are some examples of qualitative data.
- 5) There are three methods of obtaining data in research, (i) one can ask questions by using psychological tests, inventories, questionnaires or schedules, and interviews, (ii) one can observe the behaviour of persons, groups or organizations, and their products or outcomes by using participant or non-participant observation,

and (iii) one can utilise existing records or documents like personal letters, life histories, autobiographies, school records, performances or psychological or academic tests, etc.

- 6) The quality of the data obtained with the help of various tools depends upon their reliability, validity and objectivity in using them.
- 7) The validity and reliability of the data gathered through questionnaires and schedules are ensured by making the language of questions unambiguous, and also by selecting respondents who possess desired information and are likely to be keen to respond conscientiously and objectively.
- 8) A good quality data through interviews can be obtained through the willing co-operation of the interviewees.
- 9) Patience and perseverance, self-restraint and self-discipline, and listening with understanding are important traits of a good interviewer. These traits motivate an interviewee to disclose all the required information with ease and confidence.
- 10) Checklists, rating scales, score cards and observation schedules are the tools that are used for collecting and recording observations. The quality of these data mostly depends on the application and the objective use of tools. A good observation is not hurried. It is made at frequent intervals by the same observer or by several observers independently at a given time.

4.7 GLOSSARY

Data	: Consists of all relevant materials, past and present, serving as bases for study and analysis. They are quantitative and qualitative. Quantitative data are numerical figures or ratings whereas descriptive narration, responses, to open-ended questions, quotations, field notes etc. are qualitative data.
Test Inventory	: (i) Any tool by which the presence, quality or genuineness of anything is determined; (ii) a device to evaluate the performance and capabilities of an individual or group, (iii) procedure for eliciting responses upon which an appraisal of the individual concerned can be made.
Rating	: A term applied to the expression of opinion or judgement regarding some situation, object or character.
Rating Scale	: A scale with a set of points which describe varying degrees of an attitude under study.
Score Card	: It is an elaborate form of a rating scale.
Questionnaire	: A series of questions dealing with psychological, social, educational or any other topic sent to an individual or a group, with the object of obtaining data with regard to the topic under study.

- Schedule** : A device consisting of a set of questions which are asked and filled in by an interviewer in a face-to-face conversation with an interviewee. It is used for recording responses of an interview or data in an observation.
- Checklist** : A device consisting of a prepared list of items which the researcher feels are relevant for his/her study. The researcher checks the presence or absence of the items by making 'yes' or 'no'.
- Interview** : An interview is in a sense an oral questionnaire. Instead of writing the response, the subject gives the needed information orally and face-to-face.
- Observation** : A technique of collecting data by observing activities of individuals in different settings, by talking to them, or studying their constructive or creative products.
- Structured Observation** : The process of observing individuals in controlled situations.
- Document** : A record of the events of the past. Personal letters, life histories or accounts of small group processes are the examples of documents.
- Reliability and Validity of Data** : Reliability and validity of the data means the consistency and the truthfulness of the data. These are ensured by using reliable, valid and objectively usable tools.
- Halo-effect** : An error which distorts an observer's ratings of the cluster of traits in an individual. The rater forms a general opinion about the subject's merits and his/her ratings on specific traits are generally influenced by this general impression. The result is that most ratings get distorted.

4.8 CHECK YOUR PROGRESS: THE KEY

- 1) Data are of two types: (i) quantitative data and (ii) qualitative data. Quantitative data are numerical figures obtained by applying interval, ordinal or nominal scales of measurement. Qualitative data are verbal or other symbolic materials.
- 2) There are three methods of obtaining/collecting data in research:
 - i) one can ask questions by using psychological tests/inventories, questionnaires or interview schedules.
 - ii) one can observe the behaviour of persons, groups or organisations, and their products or outcomes through participant or non-participant observations. The data are recorded using an observation schedule, a rating scale or a score card.

- iii) one can utilise existing records or documents available in the form of personal letters, life histories, diaries, registration/consus data, etc.
- 3) Using an Interview Schedule Assign as
- i) Establish rapport with the interviewee(s).
 - ii) Explain to the interviewee(s) the purpose of the study.
 - iii) Chose interviewees carefully by selecting those who are willing to provide the desired information and are likely to be keen to respond conscientiously and objectively.
 - iv) Provide enough flexibility to the interviewees by exhibiting patience, perseverance, self-restraint and self-discipline.
 - v) Motivate the interviewees by asking questions not only precisely and wisely, but which show concern for their problems.

Using observation

- i) Make proper planning with regard to (a) the selection of subject(s) to be observed; (b) selection and arrangement of special condition(s), if any, for the group; (c) physical position of the observer and possible effect on the subject(s) to be observed; (d) the definition of specific activities/ characteristics of behaviour to be observed; and (c) training of the observer in recording the behaviour, etc.
 - ii) Effective execution of the task of observing the specific conditions of the subjects; proper handling of the recording instruments/tools, and proper precautions in case of participant observaion.
 - iii) Minimal halo-effect error while rating/observing the behaviour.
 - iv) Ensuring the validity and reliability of the observational measurements through observations which are not hurried and are made at frequent intervals by the same observer, or when several observers record their observations independently.
- 4) External and Internal Criticism of Documents

A rigorous evaluation of the document is called the criticism of documents. It involves the dual processes of establishing the authenticity of the source and of establishing the validity of its contents. External criticism is helpful in establishing the autenticity of the document(s) and internal criticism establishes their truthfulness, credibility and merit.