UNIT 5 COLLECTION OF DATA

Structure

5.0 OBJECTIVES

After studying this unit, you should be able to :

- describe the factors affecting choice of data
- explain the problems of collecting primary data
- narrate the different methods of collecting primary data
- state the sources of secondary data
- explain the precautions to be taken while using secondary data, and
- describe the advantages and disadvantages of using secondary data.

5.1 INTRODUCTION

In Unit 2 we have briefly discussed the meaning of primary and secondary data, different methods of collecting primary data and the sources of obtaining secondary data. In this unit we will discuss in detail the various factors determining the choice of data, the problems in collecting primary data, and different methods of collecting primary data. You will also study the merits and limitations of secondary data, sources of secondary data, and the precautions one should observe in using secondary data.

5.2 FACTORS AFFECTING CHOICE OF DATA

As you know, statistical data can be categorised as primary and secondary data. Primary data refers to the data collected for the first time by the investigator as original data. Primary data are generally in the shape of raw material to which statistical methods have to be applied for analysis, If the investigator collects data which has been collected and p r o d by someone else also, such data is referred to as secondary data. Secondary data is generally in the shape of finished product since it has already been treated in some form or the other. When you plan a statistical survey, you have to indicate whether you should collect primary or secondary data. This is a basic question which must be settled in advance before you start data collection. The choice of data depends on several factors which are stated below :

1) Object of the enquiry
2) Scope of the enquiry
3) Financial resources

4) Time factor
5) Status of the investigating agency
6) Human resources
7) Availability of secondary data
8) Degree of accuracy desired

1) Object of the Enquiry : This is the most important factor affecting the choice of data. You should collect the data which can serve the object of the enquiry. The object will indicate the type of information required for the survey. If the objective of the study is served by the primary data, you have to go for primary data.

2) Scope of the Enquiry : Scope refers to the coverage of the survey with regard to the type of information, the subject matter, geographical area covered, etc. If the information to be obtained happens to be quite comprehensive and basic, the primary data would prove to be more suitable.

3) Financial Resources : Finance, in fact, is a big constraint in statistical survey and you have to act within this limitation. The availability of funds determines, to a large extent, the type of data to be used in an investigation. In most of the cases, collection of primary data 'requires more funds than the secondary data. When funds are very limited, we usually go for secondary data. If adequate finance is available, then we can plan for collecting primary data.

4) Time Factor : The required time for the collection of primary data and secondary data should also be considered in deciding whether to select primary data or to use secondary data. Collection of primary data requires relatively more time than the collection of secondary data. If there is sufficient time for the accomplishment of the investigation, we may use primary data. But if the time is a constraint, then we should consider secondary data. The time factor. thus, affects the type of data to be collected for the survey.

5) Status of the Investigating Agency : This is another important factor in making choice of data to be collected. Much depends upon whether the investigating agency is the Government or some public organisation/institution or an individual. In the first case one can think of collecting primary data on a large scale. However, for individuals, it is very difficult to collect primary data on a large scale. For an individual it is economical and practically feasible to use secondary data. Public organisations or institutions may also take-up field surveys for obtaining relevant information but that may not be the case with private organisations. The Government or the public institutions can afford to spend more money and employ adequate number of trained and competent staff for the collection of primary data. But individuals or private organisations have lots of constraints in this regard.

6) Human Resources : Availability of human resources also affects our choice concerning the data. As you know, for collecting primary data you require more persons. If you have competent and well trained staff you can easily organise field surveys and may collect primary data. If you do not have enough human resources, you can plan to use secondary data for your investigation.

7) Availability of Secondary Data : Secondary data can be made use of only if they are available. If the secondary data are not available or if they are not adequate or not suitable, there is no alternative except to collect primary data.

8) Degree of Accuracy Desired : The choice of data also depends upon the degree of accuracy desired. Before making the choice of data, we have to decide on the degree of accuracy desired. If the secondary data attain the same degree of accuracy as desired in the present investigation, you can use secondary data. Otherwise it is advisable to plan for collecting primary data that fulfil the requirement of the desired level of accuracy.

Among all the factors discussed above, no single factor can form the basis of our choice for data. Keeping all the factors in view, a decision has to be taken whether to use primary or secondary'data.

5.3 PROBLEMS IN COLLECTING PRIMARY DATA

We have discussed various factors which influence the decision relating to the choice of data. Now, let us discuss the problems involved in the collection of primary data. In fact, the problems in collecting primary data are different from the problems we face when we collect secondary data. When you go for primary data, you generally confront the following problems:

1) When you decide to collect primary data, you have to plan the field work quite comprehensively. This is necessary because the quality of result of an enquiry depends, to a large extent, on the preparations made before starting the data collection. Various steps in planning the enquiry were already discussed in Unit 2.

2) To collect the primary data, the unit in terms of which the data have to be gathered must be very clearly and unambiguously stated. The unit must possess the entire characteristics of a good statistical unit. Properly defined unit is a prerequisite for smooth collection of primary data.

3) The problem concerning the technique of data collection is also to be looked into. As you know, there are two techniques of data collection : (i) Census method, and (ii) sample method. The former method requires the collection of information from all the units in the population whereas the latter method obtains information from only a part of the universe. The investigator must decide which technique he will use. The choice would depend upon the availability of resources, the time factor, nature and scope of enquiry and similar other factors. In case of sampling technique, the sampling design has to be carefully specified.

4) Developing the frame is another problem which has to be set prior to the collection of data. 'Frame' refers to a list, map or other specification of the units which constitute the available information relating to the population designated for a particular enquiry. Data can be collected easily if a suitable frame already exists. Otherwise a suitable frame has to be developed before embarking on the collection of data.

5) Before collecting primary data it is necessary to decide on the degree of accuracy desired. As you know, a reasonable level of accuracy is desired in all statistical enquiries. The desired level of accuracy is to be determined keeping in view the object and purpose of enquiry.

6) Designing the forms for the collection of data is another problem in the context of collecting primary data. Careful attention should be given to the designing of various forms (viz., questionnaire, schedule, etc.) that will be used. They should be designed in such a way that required information can be collected through them. Before the commencement of data collection, the forms should be pre-tested, in order to examine their effectiveness. If some shortcomings are identified in the pre-testing, the same should be eradicated before finalising the instrument i.e., the form.

7) The selection, training and supervision of the field staff is more important for the collection of primary data than is necessary in the case of secondary data. Since the success of survey depends upon the field staff, it is essential that they are properly selected, thoroughly trained and their work closely supervised. The enumerators selected should be honest, intelligent and hard working. They should be able to elicit the needed information from the respondents.

8) Then comes the problem of exercising control over the quality of field work. Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently. A careful watch should be kept for unanticipated factors in orders to keep the data collection work as much realistic as possible. In other words, steps should be taken to ensure that the data collection work is under control so that the information collected is in accordance with the pre-defined standard of accuracy. The collected data should be checked for omissions, inconsistencies, and other errors before they are passed on for further processing.

9) In spite of the best effart, the problem of non-response may remain. Some suitable method(s) should be designed to tackle this problem. One method of tackling the problem of non-response is to make a list of non-respondents and take a small sample of them and efforts can be made for securing response with the help of experts. But in any case enumerators should not be allowed to substitute for anyone who is not considered to be a good respondent. Otherwise the bias would creep in the collected information. Proper organisation has to be set up so that the data collection work proceeds smoothly. This problem is more serious particularly in big enquiries or investigations. You must judiciously select the method(s) of data collection (like observation, questionnaire, schedule, interview, etc.) which may prove appropriate for your study. Of course, while making a choice of the method to be used, you should pay attention to the nature and object of enquiry, availability of funds, the time factor and precision required,

Check Your **Progress A**

1) Differentiate between primary and secondary data.

....

....

....

....

2) List the factors that affect the choice of data to be used in an investigation.

....

....

....

....

3) Mention five problems with which the investigator is generally contronted while collecting primary data.

....

....

....

....

4) State whether the following statements are True or False.

i) It takes more time to collect secondary data compared to primary data.

ii) National income data collected by the Government is a secondary data in the hands of a researcher who uses it for his study.

iii) Any single factor, considered in isolation, should not form the basis of our choice of data.

iv) There is no difference in the problems confronted by the investigator whether he collects primary data for himself or uses secondary data.

v) Secondary data should be used after it has been ensured that it is reliable, adequate and suitable.

vi) The problem of developing a suitable frame arises while using secondary data.

vii) The problem of selection, training and supervision of the field staff arises whenever we collect primary data irrespective of the method we use for ,the purpose.

viii) Problem of non-response is a major problem in case of secondary data.

5.4 METHODS OF COLLECTING PRIMARY DATA

There are many methods for the collection of primary data and any one of them can be employed depending upon the nature of the survey. In Unit **2** you studied in brief about these methods. Now let us **study** about them in more detail.

5.4.1 Observation

Observation is a systematic viewing, coupled with consideration of the seen phenomena. The Concise Oxford Dictionary defines observation as accurate watching **and** noting of phenomena as they occur in nature with regard to cause and effect or mutual relations. The required information is obtained directly through observation rather than through the reports of others. In the case of behaviour one finds out what the individual does, rather than what the individual says he does. If the informants are unable to provide the information or can give only very in exact answer, questioning is not useful and observation is the only way to proceed. For instance, when you are studying the behaviour of small children who cannot **speak,** you can collect the information by abserving the children under different circumstances. You should remember that all phenomena are not open to observation. Even if a phenomenon is open to observation, it may not find a ready observer at hand.

Observation may be participant observation or non-participant observation. In Participant Observation Method the observer joins in the daily life of the group or organisation he is studying. He watches what happens to the members of the community and how they behave. He also engages in conversation with them to find out their reactions to, and interpretations of, the events that have occurred. In the **Non-Participant** Observation Method in order to collect information the observer will not join the group or organisation he is studying but will watch it from outside.

Merits : This method has the following advantages :

1) This is the best suitable method when the informants are unable to provide information or can give in exact information.
2) This method provides first hand information and provides deeper insights into the problem. Therefore, this is useful for intensive studies.

Limitations : This method suffers from the following limitations :

1) In many cases, you cannot predict when the events occur. So a phenomenon which is open to observation may not find a ready observer at hand.
2) The observer should be very objective in interpreting the events he has observed. Otherwise, the bias of the observer may creep into the results.
3) This is not suitable for large scale extensive studies.
4) The presence of the observer may influence the behaviour he is observing. In such cases he may not get the actual information.

5.4.2 Personal Interviewing

Under this method data are collected by the investigator himself through interviews. Therefore, the enquiry is intensive rather than extensive. Under this method the investigator meets the informants personally, asks them questions pertaining to enquiry and collects the desired information. Thus, if a person wants to collect data on the wages of workers of the National Ball Bearing Company, he would go to the factory site of this company, contact the workers and collect the relevant information. Thus, this method is generally used in small size surveys confined to a small locality.

Interviews can be formal or informal. In **Formal Interviewing,** set questions are asked and the answers are recorded in a standardised form. This is the practice in large scale interviews where a number of investigators are assigned to the job of interviewing. In a formal interview, the interviewer's bias is minimised. This type of interview is most suitable when you know very clearly what type of information you require for your survey. In the case of Informal Interviewing, the investigator may not have a set of questions but have only a number of key points around which to build the interview. The interviewer is at liberty to vary the sequence of questions, to explain their meaning, to add additional ones and even to change the wording. Informal interviews are preferred in the case of an explorative survey where you are not sure about the type of data you collect.

Merits : The major advantages of this method are as follows :

1) The response of the persons interviewed is more encouraging as most people are willing to supply information when approached personally.
2) The information obtained is likely to be more accurate because the doubt of any of the informants can be cleared by the investigator himself.
3) Additional information about the personal characteristics of informants which are helpful in interpreting the results later on may as well be collected.

Limitations : The limitations of this method are as follows :

1) Major limitations of this method are the subjective factors or the biases of the investigator coming in either consciously or unconsciously.
2) It is a costly and time consuming method especially when the number of persons to be interviewed is large and they are spread over a wide area. So, this is not suitable for big surveys.

5.4.3 'Through Local Reports and Correspondents

Under this method the investigator appoints local agents or correspondents in different places of the field of enquiry and the relevant information is obtained through them. These correspondents collect and transmit the information to the office of the investigator. Newspaper agencies generally adopt this method. This method is also used by various departments of the Government in cases where regular information is to be collected from a relatively wide area. In case of making crop estimates or for obtaining regular information regarding prices of different commodities for the preparation of price index this method is generally used.

Merits : The chief merit of this method is that it is comparatively cheap and also gives approximately good results. The method is equally appropriate for extensive enquiries.

Limitations : The danger of entering personal bias of the correspondents in the reports submitted by them at more or less regular intervals is, however, great in this method. Thus, it is necessary that the correspondents or the agents appointed for the purpose must be selected very carefully and trained properly.

5.4.4 Questionnaire

Collection of data through questionnaires is the most popular method for collecting primary data. A questionnaire is a list of questions pertaining to the enquiry. Under this method a questionnaire is sent to various informants with a request to answer the questions and return the questionnaire. The questionnaire is mailed to the respondents who are expected to read the questions and record their response in the space meant for the purpose on the questionnaire itself. The respondents have to answer the questions on their own. This method is extensively employed in various economic and business surveys.

Merits : The merits of this method are as under :

1) This method is very economical particularly when the universe is large and spread geographically on a vast area.
2) Since the answers happen to be in the respondent's own words, he/she is free from the bias of the interviewer.
3) Respondents can take their own time to answer the questions. So they give well thought out answers.
4) Respondents that are at remote places and are not easily approachable can also be reached conveniently.
5) Large samples can be covered and thus the results can be more dependable and reliable.

Limitations : This method also suffers from the following limitations :

1) Sometimes the respondents do not bother to return the questionnaires. So there is the problem of low rate of return of the duly filled in questionnaires. And also bias due to non-response cannot often be determined.
2) Questionnaires can be circulated only among the respondents who are educated and cooperative.
3) Once the questionnaires are sent to the respondents, the investigator cannot change or modify the questions for individual respondents.
4) There is no flexibility because of the difficulty of amending the approach once the questionnaires have been despatched.
5) There is also the possibility of ambiguous replies or omission of replies to certain questions. Interpretation of omissions is difficult.
6) It is difficult to know whether willing respondents' are truly representative.
7) This method is likely to be the slowest of all, because the respondents take their own time to return the filled in questionnaires.

Before sending them to the respondents, it is advisable to conduct a 'Pilot **Survey'** for pre-testing it. Pilot Survey is in fact the replica and rehearsal of the main survey. From the experience gained in this sort of survey, changes can be made in the questionnaire for the final collection of data. The pre-testing is necessary particularly in case of a big enquiry.

Features of a Good Questionnaire : In order to make the questionnaire more effective, it must be very carefully drafted. The' form and tone of the questionnaire must be designed so as to bring in the personal element which is lost in the mailed questionnaire. The following are the qualities of a good questionnaire :

1) It should be short and simple.

2) Questions should proceed in logical sequence starting with easy questions and then moving on to more difficult ones. Personal questions should generally be avoided or may be left to the end.

3) Questions may be dichotomous (yes or no type), or multiple choice. Open ended questions are difficult to analyse and should be avoided to the extent possible.

4) In order to ensure the reliability of respondent there should be some control questions. They introduce a cross-check to see whether the information collected is correct or not.

5) Adequate space for answers should be provided in the questionnaire itself. There should always be provision for indications of uncertainty e.g., "do not know", "no preference". and so on.

6) Layout and design of the questionnaire should also be attractive so that it may attract the attention of the respondents.

5.4.5 Schedule

This method of data collection is similar to that of the questionnaire. The schedule is also a proforma containing a set of questions. The'difference between the questionnaire and the schedule is that the schedule is being filled in by the enumerators who are specially appointed for the purpose. These enumerators go to respondents with the schedules and ask them the questions from the schedule in the order they are listed. The enumerator records the replies in the space meant for the same in the schedule itself. In certain situations, schedules are handed over to respondents and the enumerators help the respondents in recording the answers. Enumerators explain the objectives of the investigation and also remove the difficulties which the respondent may feel in understanding the implications of a particular question(s) or the definition or concept of difficult terms. Thus, the essential difference between the questionnaire and schedule is that the former (i.e., questionnaire) is sent to the informants by post and in the latter case the enumerators carry the schedule personally to informants and fill them in their own handwriting. This method is usually adopted in investigations conducted by governmental agencies or by some big organisations. For instance, population census all over the world is conducted through this method.

Data collection through schedules requires enumerators for filling up schedules and as such they should be very carefully selected. They should be trained to perform their job well. They should be intelligent and must possess the capacity of cross-examination in order to find out the fact. Above all, they should be honest, sincere, hard working and should have the patience and perseverance. In drafting the schedules, all points stated for a good questionnaire, must as well be observed.

Merits : The main advantages of this method are as follows :

1) It can be adopted in those cases where informants are illiterate.

2) The problem of non-response is avoided as the enumerators go personally to obtain the information.

3) The method is very useful in extensive enquiries and can lead to fairly reliable results.

4) The identity of the respondent is known which is not always clear in case of a questionnaire.

Limitations : This method has the following limitations :

1) This method is very expensive as enumerators are generally paid persons. Money also has to be spent in training them.

2) Another limitation is that if the investigator is not good in interviewing, most of the information collected by him may be unreliable.

3) Since the investigator is present when the respondent is giving the answers, the respondent may not give answers to some personal questions freely.

5.4.6 Choice of Method

As discussed above, there are several methods for the collection of primary data. You have to choose the best suitable method for your study. You must make your choice of method very judiciously so that the method chosen will be quite appropriate and effective. For this purpose, you should consider (i) nature, scope and object of enquiry, (ii) availability of funds, (iii) the time factor, (iv) the precision required, (v) other relevant factors as stated in Section 5.2 above.

But you should always remember that each method of data collection has its own advantages and limitations and none of them is suitable in all situations. For instance, the observation method is suitable for intensive field surveys to be conducted when the incident is really happening. The interview method is considered suitable in cases where indirect sources of information are required to be tapped because direct observation is not possible. Information through local reports and correspondents is considered a suitable method when information is to be obtained at regular intervals from a relatively wide area. The questionnaire method is appropriate in extensive enquiries where informants are spread over a wide area. This method, however, can be adopted only when the respondents are educated and capable of filling in their responses themselves. Data collection through schedules is suitable in case of extensive enquiries spread over a wide area wherein informants may not always be literate. This method, however, requires lot of funds, relatively more time and a team of dedicated enumerators. Due to its high rate of response this method is usually adopted by the Government in extensive enquiries such as the population census.

In case funds are available and more information is desired, personal interview method can easily be adopted provided the enquiry is confined to a limited area. In case there is sufficient time and limited funds the questionnaire method will be more suitable. Where a wide geographic area is to be covered, the use of questionnaires supplemented by personal interview will yield more reliable results. In short, the most desirable approach while making a choice of method depends on the nature of the particular problem and on the time and resources (financial as well as human) available along with the desired degree of accuracy. Above all those, much depends upon the ability and experience of the investigator. In this context A.L. Bowley's remark "in collection of statistical data commonsense is the chief requisite and experience is the chief teacher", is quite appropriate.

Check Your Progress B

1) Differentiate between an interview and a schedule.

....

....

....

....

....

2) Distinguish between a questionnaire and a schedule.

....

....

....

....

....

3) Distinguish between an observation and a interview.

....

....

....

....

....

4) Name the appropriate method of primary data collection under each one of the following situations :

	Situation	Appropriate Method
i)	Intensive survey is to be conducted when the incident is really happening.	....
ii)	When the information is to be obtained at regular intervals over a wide area.	....
iii)	In extensive enquiries where literate informants are spread over a wide area and you have enough time.	....
iv).	When time is ample, funds are limited, much information is to be gathered, area to be covered is wide consisting of literate persons.	....
v)	When time is ample, funds are available area is wide, all persons may not be literate, team of sincere and honest enumerators is available.	....

5.5 SOURCES OF SECONDARY DATA

We have discussed various problems and methods for the collection of primary data. Now let us discuss the secondary data. In case you decide to collect secondary data, you have to look into various sources from where you can obtain it. The source from which you actually collect the data depends upon the nature of the problem. The sources of secondary data can broadly be classified into two categories : i) pubished, and ii) unpublished sources. Let us discuss these two sources in detail.

5.5.1 Published Sources

Data is published and made available to all the interested parties. Usual sources of pubiished data are the following :

1) Reports and official publications of the central and state governments.
2) Various publications of foreign governments or of international bodies and their subsidiary organisations such as the World Bank, International Monetary Fund, United Nations Organisation, etc.
3) Semi-official publications of various local bodies such as municipal corporations, district boards, etc.
4) Private publications, such as :
   i) Technical and trade journals such as Commerce, Capital, etc.
   ii) Publications of professional bodies like the Institute of Chartered Accountants of India, Institute of Company Secretaries, Institute of Bankers, etc.
   iii) Publications of trade and industry organisations like the Federation of Indian Chambers of Commerce, Stock Exchanges, etc.
   iv) Annual reports of banks and joint stock companies.
   v) Reports prepared by research scholars, universities, economists, etc.
   vi) Public records and statistics, historical documents and other sources of published information such as books, magazines, newspapers, etc.

5.5.2 Unpublished Sources

All statistical material is not necessarily available in published form. There are various sources of unpublished data which can also be used wherever necessary. Unpublished data may generally be found in diaries, letters, unpublished biographies and autobiographies. Unpublished data may also be available with scholars and research workers, trade associations, labour bureau and other public/private organisations and individuals.

Thus, there is a vast amount of information available both in published and unpublished sources which constitute the basis for several statistical studies. The investigator may use one or more sources suitable for his project.

5.6 PRECAUTIONS IN USING SECONDARY DATA

As you know, the secondary data has been collected and analysed by someone else. Therefore, while using it you should be very careful. You have to study the data carefully because it may be unsuitable or may be inadequate in the context of your study. It is never safe to take published statistics at their face value without knowing their meaning and limitations. You should always keep in mind the following precautions before using secondary data :

1) Reliability of Data : Secondary'data should only be utilised if they are found reliable. The reliability can be tested by examining the following aspects :

   i) Who collected the data?
   ii) What were the sources of data?
   iii) Were they collected in a proper manner?
   iv) At what time were they collected?
   v) Was the compiled biased?
   vi) What level of accuracy was desired? Was it achieved?

   If the collecting agency happens to be some government institution or international organisation or other competent authority, the secondary data can be taken as more reliable compared to the data collected by individuals or by some private organisation that is not well reputed. Secondary data collected from published sources of government departments and corporations established under the Act of Parliament, and international institutions are reliable.

2) Suitability of Data : The data which may be suitable in one enquiry may not necessarily be suitable in another enquiry. So you should examine whether the data is suitable for your study or not. If the available data is found to be unsuitable, it should not be used. So, you must carefully *scrutinise* the definition of various terms and units of data collected. Similarly, the object, scope and nature of the original enquiry must also be studied. If these aspects are not found sound, the data will not be suitable for the relevant enquiry and should not be used. For example, you are conducting a survey on wage levels including allowances of workers. If the secondary data is available only on basic wages, such data is not suitable for the present enquiry.

3) Adequacy of Data : Adequacy of the data has to be judged in the light of the requirements of the survey and the geographical area covered by the available secondary data. For example, if our object is to study the wage rate of workers in cotton textile industry of India and the published reports provide the data on wage rates of workers in all industries together, then the data would not serve the purpose. The question of adequacy may also be considered in the light of the time period for which the data are available. For example, for studying price trends we may require data for the last 20 years but the secondary data is available for the last 4 years only. Here the available data would be inadequate and would not serve our object. Similarly, if the level of accuracy achieved in a given data is found inadequate for the purpose of a relevant enquiry, such data should not be used by the investigator.

Thus, we should use given secondary data if it is reliable, suitable and adequate. If secondary data is available from authentic sources and also suitable and adequate for the particular study, it will not be economical to spend time, energy and money in organising field survey for collecting primary data. Thus, if the suitable secondary data is available, as discussed above, it should be utilised with due precaution.

5.7 ADVANTAGES AND DISADVANTAGES OF SECONDARY DATA

There are certain advantages and disadvantages in using the secondary data. Let us now discuss them.

Advantages

1) It is much more economical to use secondary data as we do not need to spend money on printing data collection forms, and hiring large numbers of enumerators.

2) Secondary data, if available, can be obtained more quickly compared to primary data. Secondary data can be collected in a few days whereas it may take months to complete field work for obtaining primary data. As such the investigation may be accomplished in lesser time with the help of secondary data.

3) Secondary data facilitates the work of individual investigator or research organisation when they find it impossible to collect primary data with regard to several subjects. Census data, national income data, etc., cannot be collected by an individual but they can be easily obtained from government publications.

4) The worldwide data concerning diverse phenomena like world trade, industry, population, health, etc., are usually obtainable through secondary sources published by international agencies like United Nations Organisation, World Bank, International Monetary Fund. etc.

5) At time, there may be lots of usable information in the already available data which can well be utilised by the investigator and he can even have new insights concerning the problems he is studying.

6) Most statistical analysis in practice rest upon secondary data since they are readily available in many cases in diverse fields. We use primary data only when secondary data do not provide an adequate basis for analysis.

Disadvantages

1) Secondary data is very risky and is to be used only when their reliability, suitability and adequacy have been ensured. If this is not done, the results of the investigation may not be fully correct.

2) It is difficult to find secondary data which exactly fit the needs of your investigation.

3) There is also the problem of finding secondary data which is sufficiently accurate. Due to bias, inadequate size of sample, errors of definition, etc., the secondary data may be erroneous.

4) Many times, secondary data are not available and in such situations we have to compulsorily collect primary data.

Check Your Progress C

1) Write names of five published sources of secondary data.

....

....

....

....

2) What are the basic factors to be kept in mind while using the secondary data.

....

....

....

....

3) State whether each of the following statements is True or False.

i) Using secondary data is always economical than using primary data.

ii) Secondary data is always available in published sources.

iii) Secondary data when used without verifying the suitability, reliability and adequacy may not yield correct results.

iv) Secondary data which is not published is not reliable.

5.8 LET US SUM UP

Statistical data can be either primary data or secondary data. Primary data means the data which is collected for the first time as original data. Secondary data refers to the data collected and processed by someone else earlier and is being used now in the present enquiry. Important factors which affect the choice of data are : 1) object and scope of enquiry, 2) financial resources, 3) the time factor, 4) status of the investigator, 5) human resources, 6) precision required, and 7) availability of secondary data.

There are several problems in the collection of primary data. Important among them are : The problem of defining the unit of data collection, the problem concerning the technique of collection, the problem of developing the frame, the problem of deciding the degree of accuracy, the problem of designing questionnaire/schedule, selection and training of enumerators, the problem of tackling non-response, the problem of control of the field work and other administrative aspects.

There are several methods of collecting the primary data viz., observation, interview, information from correspondents and local reports, questionnaire and schedules. Each of these methods has its own merits and limitations. As such no one method is appropriate in all situations. The investigator must make a choice of one or the other method as per the needs of the situation after taking into consideration the object and nature of enquiry, availability of funds, the time factor and the precision required.

Secondary data may be obtained from both published and unpublished sources. Secondary data should always be used with precaution. It should be used after ensuring reliability, suitability and adequacy, otherwise it may result in misleading conclusions. Secondary data have certain advantages and disadvantages.

5.9 KEY WORDS

Enumerators : Persons who go to respondents and gather information from them through pre-designed schedules.

Interview : A method of collecting primary data by meeting the informants personally and asking them questions.

Non-response : Problem of not getting the questionnaire duly filled in by the informants,

Observation : A method of collecting the information by observation when the incident is actually happening.

Pilot Survey : Replica and rehearsal of the main survey. The experience gained through the pilot survey is made use of in finalising the questionnaire.

Primary Data : Data that is collected for the first time as original data.

Published Sources : Sources which contain published statistical information.

Questionnaire : An instrument for collection of primary data containing a list of questions pertaining to enquiry, generally sent by post to informants and the respondent himself writes the answers.

Schedule : An instrument for the collection of primary data which contains a set of questions to be filled in by the enumerators who are specially appointed for the purpose.

Secondary Data : Data which were collected and processed by someone else but are being used in the present enquiry.

5.10 ANSWERS TO CHECK YOUR PROGRESS

A) 4) i) Fasle ii) True iii) True iv) False v) True vi) False vii) False viii) False.

B) 4) i) Observation
ii) Through correspondents and local reports
iii) Questionnaire
iv) Questionnaire
v) Schedule

C) 3) i) False ii) False iii) True iv) False

5.11 TERMINAL QUESTIONS

1) Differentiate between primary and secondary data. Describe the factors that affect the choice of data to be used in an investigation.
2) Describe the problems that are usually confronted by the investigator when he decides to collect primary data.
3) Explain various methods of collecting primary data and also narrate their merits and demerits.
4) "It is never safe to use secondary data without proper scrutinisation" explain.
5) What are the sources of secondary data? Explain the advantages and disadvantages of using secondary data.

Note : These questions will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the University. These are for your practice only.