

## UNIT 12 COLLECTION AND SCRUTINY OF DATA

Collection and  
Scrutiny of Data

### Structure

- 12.1 Introduction
  - Objectives
- 12.2 Primary Data
- 12.3 Secondary Data
- 12.4 Scrutiny of Primary Data
- 12.5 Preparation of Questionnaire
- 12.6 Summary
- 12.7 Solutions/Answers

### 12.1 INTRODUCTION

Recall the definition of statistics, given in Sec. 11.3 of previous unit “Statistics is a branch of science which deals with collection, classification, tabulation, analysis and interpretation of data”.

In the above definition out of the five successive steps (i.e. collection, classification, tabulation, analysis and interpretation) used in any statistical investigation, the first step is collection of data, and last step is interpretation of data which ultimately depends on the collection of data. So, if collection of data is not done carefully and sincerely then goal of the statistical investigation is not achieved or objective(s) of the statistical investigation is/are not fulfilled or final outcomes of the investigation will not be satisfactory.

Therefore, it becomes very important to focus on the collection of data in detail. In this unit two main types of collection of data namely primary and secondary will be discussed in detail. Also we shall discuss their different methods of collection with their merits and demerits. Scrutiny of data is also discussed in this unit. Finally we conclude this unit by throwing some light on preparing a questionnaire.

### Objectives

After completing this unit, you should be able to:

- define primary data and get familiar with the different methods of collection of primary data;
- define secondary data and get familiar with the different sources of collection of secondary data;
- get an idea about the scrutiny of data; and
- know some important points related to the preparation of a questionnaire.

### 12.2 PRIMARY DATA

In previous section of this unit, we have seen why collection of data is important in any statistical investigation. In fact there are mainly two types of data namely Primary Data and Secondary Data. Both have their own importance. In a statistical investigation which of the two is to be used, totally

depends upon many factors such as nature of problem, purpose of investigation, time period in which conclusion required and finally another important factor is availability of money/resources.

In this section we will discuss concept of primary data, different methods of collection of primary data with their merits and demerits.

Let us first formally define primary data and secondary data. Here we will just define secondary data with an example. Next section is devoted to discuss secondary data in detail.

### **Primary Data**

Data which are collected by an investigator or agency or institution for a specific purpose and these people are first to use these data, are called primary data. That is, these data are originally collected by these people and they are first to use these data.

For example, suppose a research scholar wants to know the mean age of students of M.Sc. Chemistry of a particular university. If he collects the data related to the age of each student of M.Sc. Chemistry of that particular university by contacting each student personally. The data so obtained by the research scholar is an example of primary data for the same research scholar.

### **Secondary Data**

The data obtained/gathered by an investigator or agency or institution from a source which already exists, are called secondary data. That is, these data were originally collected by an investigator or agency or institution and have been used by them at least once. And now, these data are going to be used at least second time.

For example, consider the same example as discussed in case of primary data. If the research scholar collects the ages of the students from the record of that particular university, then the data thus obtained is an example of secondary data. Note that, in both the cases data remain the same, only way of collecting these data differs. Our aim of just defining secondary data in this section is over, because we just want to click the idea how primary and secondary data differ from each other and are discussed into two different sections.

Now we move towards the aim of this section, which is to focus on primary data. We have already define primary data. There are a number of methods of collection of primary data depending upon many factors such as geographical area of the field, money available, time period, accuracy needed, literacy of the respondents/informants, etc.

Here we will discuss only following commonly used methods.

- (1) Direct Personal Investigation Method
- (2) Telephone Method
- (3) Indirect Oral Interviews Method
- (4) Local Correspondents Method
- (5) Mailed Questionnaires Method
- (6) Schedules Method

Let us discuss these methods one by one with some examples, merits and demerits.

## (1) Direct Personal Investigation Method

In this method, the investigator personally contacts the informants and collects the related data through face to face interviews of the informants. Due to face to face meeting of investigator and informants data collected under this method has maximum degree of accuracy. But the degree of accuracy depends upon the sincerity, honesty, unbiasedness and expertness of the investigator, because it is the investigator, who ultimately gives the final shape to the information provided by the informants. This method of collection of primary data is recommended/suitable when field of enquiry is small, secrecy related to data is need, high accuracy is required and time as well as money is sufficient.

Following are some **merits** of this method:

- (i) It is simple to apply.
- (ii) It is convenient for both investigator as well as informants.
- (iii) Data have high degree of accuracy.
- (iv) Data are homogenous in nature as there is only one investigator.
- (v) Due to the presence of both investigator and informants, there is flexibility to clear any doubt or some other modification are possible according to physibility.
- (vi) Confidential information can also be obtained.

Having so many plus points, this method is not free from the demerits.

Following are some **demerits** of this method:

- (i) It is time consuming and costly.
- (ii) It is not suitable when area of investigation is large.
- (iii) It suffers from the biasness of the investigator.
- (iv) Data may be misleading if the investigator do not collect the data sincerely and honestly.
- (v) If the investigator does not have expertise, data again may be misleading.

## (2) Telephone Method

In the direct personal investigation method investigator has to personally contact with the informants, but now a day's telephone is very good communication tool. If the information of the interest is collected through telephone then data so obtained come under telephone method.

Some merits and demerits of this method are listed below:

### Merits

- (i) All merits of method 1 are also the merits of this method. Some additional merits are given below.
- (ii) It is easy to apply compare to method 1.
- (iii) It is time and cost saving method compare to method 1.
- (iv) It is suitable when area of investigation is large compare to method 1.

### Demerits

- (i) Information related to those informants who do not have telephone and those who actually have telephone but their number are not in the list will not be included.
- (ii) This method also suffers from the demerits (iii) to (v) listed in method 1.

### (3) Indirect Oral Interviews Method

In this method, investigator does not meet to the actual informants directly, but the related information is collected/obtained from other persons who are supposed to have the required type of information. The informants who provide the information about actual informants are known as 'witnesses'. The success of this method mainly depends upon the skill and experience of the investigator. Because it depends upon the way, sequence and trick of questions prepared and asked by the investigator from the informants. It also depends on the behavior and how much he/she is capable to create confidence in the informants. This method of collection of primary data is generally adopted to obtain the information related to the cases such as

- (i) murder
- (ii) theft
- (iii) in the cases where a person hesitates to provide correct information.

For example, if a researcher wants to collect the data on the smoking habit of students of a particular class of a college then it may happen that actual informants does not provide correct information. So, data can be collected with the help of class met or college met or from the neighbours.

Following are some **merits** of this method:

- (i) As far as time and money is concerned, it is economical compare to direct personal investigation.
- (ii) This method is easy to use, even if the area of investigation is large.
- (iii) Informants being a third person, so it is free from the biasness of both actual informants and investigator.

Following are some **demerits** of this method:

- (i) Here data totally depends upon the information provided by the third person. So, data suffer from the biasness of the third person.
- (ii) If the investigator is not experienced and well behaved, then data will not be reliable.
- (iii) It also depends on the honesty of the investigator.

### (4) Local Correspondents Method

In this method, first of all some local correspondents or agents are appointed by the investigator or agency or institute to collect data. These local agents directly meet to the informants and collect data related to the required purpose in hand. Data collected by these local agents have high degree of accuracy, because they are familiar with the local language, traditions, general behaviour of the people, etc. of that particular area.

This method of collection of primary data is suitable if

- Area of investigation is large, e.g. news channels have their reporters throughout India.
- Time period in which information is needed is very short. For example, news related to the happening of any special incident can be easily seen on the news channels in very short time period after its happening.
- Information is needed on regular basis. For example, news are provided daily using this method.

Below some **merits** of this method are listed.

- (i) It is very economical in terms of money, time and man power.
- (ii) Time period in which information are provided by this method is very short.
- (iii) Large area and heterogeneity of the informants can easily handled by this method.
- (iv) This method can provide regular basis information.

Having so many merits, this method also has some **demerits**, listed below.

- (i) Data suffer from the biasness of the local agents.
- (ii) If local agents do not perform their duties with honesty and sincerity, then data will not be reliable.

### (5) Mailed Questionnaires Method

In this method, first of all a list of questions related to the information required by the investigator is prepared. At the time of preparing the list of questions following points should keep in mind.

- Number of questions should not be too many.
- Each question should be related directly or indirectly to the objective(s) of the investigation.
- Each question should be clear.
- Generally objective type questions should be used, but if necessary multiple choice and open-ended questions can also be used.
- Language should be simple and effective.

**Open ended questions** are defined on page number 87 of this unit.

After preparing the 'final list' of questions known as questionnaire, it is sent through mail to the informants. With this questionnaire a covering letter in which it is requested to the informants that please sent it back after completing and a brief introduction about the objective of the investigation is also attached.

This method is suitable when the informants are literate and area of investigation is large.

Following are some **merits** of this method:

- (i) This method is very useful if area of investigation is large.
- (ii) This method is very economical as far as time, money and labour is concerned.
- (iii) This method provides very good results when informants are literate.
- (iv) Informants have enough time to think and give correct information. Thus data obtained by this method have high degree of accuracy.
- (v) Biasness of the investigator is not involved.

Having being very economical and suitable for large area, following are some **demerits** of this method.

- (i) This method fails if informants are illiterate.
- (ii) Generally percentage of responses are very less because people take less interest in filling up the questionnaires.
- (iii) If informants do not fill up the questionnaire sincerely or honestly then biasness of the informants may mislead the investigator.

## (6) Schedules Method

In this method first of all a list of questions based on the information to be required is prepared like mailed questionnaires method, known as schedule. After doing this whole area of investigation is divided into sub areas. Then a number of people are appointed to collect the information directly from the informants. The appointed people are known as enumerators. The exact figure of the enumerator to be appointed depends upon the area of investigation and the amount of information to be required. These enumerators meet with the informants face to face and after giving a brief introduction about the objectives of the investigation they ask the answer of each question listed in the schedule. The answers provided by the informants are filled up in the schedule by enumerator themselves. This is one of the main differences of the two methods namely mailed questionnaires method and schedules method. In mailed questionnaires method answers are filled up by the informants themselves while in schedules method this job is done by enumerators. This characteristic of the schedules method make it superior to mailed questionnaire method in the case when informants are illiterate or semi-illiterate. Here enumerators have to meet directly with the informants, therefore it becomes important that enumerators should be well behaved, honest sincere and unbiased in nature.

This method is suitable when area of investigation is large and informants are illiterate or semi-illiterate.

Following are some **merits** of this method:

- (i) This method is suitable when informants are illiterate or semi-illiterate.
- (ii) It is application whatever large the area is.
- (iii) After every ten years census data is collected by using this method in India.
- (iv) Data are least affected by the bias of the enumerators and investigators.
- (v) Since the information is directly obtained from the informants, so data collected by this method are more reliable and have a high degree of accuracy.
- (vi) Because enumerator is present in front of the informants so if informants have any doubt, he/she can easily clear it from the enumerator.

There are so many merits of this method, even though it is not free from the demerits. Following are some **demerits** of this method:

- (i) It is very time-consuming and large amount of money is needed.
- (ii) Because a large number of enumerators have to be appointed, so it becomes too difficult to get all well trained and experienced persons.
- (iii) Training is also needed to the enumerators.
- (iv) Even after providing the training, some enumerators may not do their responsibilities sincerely, honestly and efficiently.
- (v) Accuracy of the data will suffer if enumerator is bias or not devoted.

---

## 12.3 SECONDARY DATA

---

Discussion in the previous section shows that collection of primary data requires lot of time, money, manpower, etc. But sometimes some or all these resources are not sufficient to go for the collection of primary data. Also, in some situations it may not be feasible to collect primary data easily. To

overcome these types of difficulties, there is another way of collecting data known as secondary data. The data obtained/gathered by an investigator or agency or institution from a source which already exists, are called secondary data. That is, these data were originally collected by an investigator or agency or institution and has been used by them at least once and now, these are going to be used at least second time. Already existed data in different sources may be in published or unpublished form. So sources of secondary data can broadly be classified under the following two heads.

- (1) Published Sources, and
- (2) Unpublished Sources.

Let us discuss these two main types of sources one by one:

### **(1) Published Sources**

When an institution or organisation publishes its own collected data (primary data) in public domain either in printed form or in electronic form then these data are said to be secondary data in published form and the source where these data are available is known as published source of the secondary data of the corresponding institution or organisation. Some of the published sources of secondary data are given below:

#### **(a) International Publications**

There are many international organisations including the governments' organisations of different countries which collect data regarding the characteristics under their objectives and publish these data. Some of these organisations or publications are:

- (i) Annual Abstract of Statistics (United Kingdom)
- (ii) Annual Reports of International Labour Organisation (ILO)
- (iii) World Health Organisation (WHO)
- (iv) World Bank
- (v) UNESCO Institute for Statistics, etc.

#### **(b) Government Publications in India**

There are number of government organisations or bodies at national level or state level which collect and publish data on different characteristics of interest. Data published by these bodies play a significant role in the sources of secondary data. Some of these organisations or bodies or publications are:

- (i) Office of Registrar General of India
- (ii) Central Statistical Organisation (CSO)
- (iii) National Sample Survey Organisation (NSSO)
- (iv) Reserve Bank of India Bulletin
- (v) Directorate of Economics and Statistics (DES), etc.

#### **(c) Published Reports of Commissions and Committees**

From time to time, Central government and State governments constitute or appoint some commissions and committees to get a road map on some issues of interest. Some of these (with time of appointment in brackets) are listed below:

- (i) Sarkaria Commission (1983)
- (ii) Sixth Pay Commission (2006)

## **Matrices, Determinants and Collection of Data**

- (iii) Shunglu Committee (2010)
- (iv) Kalodkar Committee (2010), etc.

### **(d) Research Publications**

Published research work of research scholars in different journals throughout the world is another major source of published data. Most of these research works are carried out in universities or other research institutes. Some of the journals related to the subject Statistics are:

- (i) International Journal of Probability and Statistics
- (ii) American Journal of Mathematics and Statistics
- (iii) Sankhya
- (iv) Journal of Statistics & Management Systems
- (v) Journal of Statistical Research, etc.

### **(e) Reports of Trade and Industry Associations**

In India, there are number of trade and industry associations whose published reports are also secondary sources of published data. Some of these associations are given below:

- (i) All India Association of Industries
- (ii) The Indian Cotton Mill Association
- (iii) All India Resort Development Association
- (iv) All India Biotech Association, etc.

### **(f) Published Printed Sources**

Books, directories, newspapers, magazines, etc. are published printed sources of secondary data. Some of these are:

- (i) Statistical Year Book, India
- (ii) Business Line
- (iii) The Economic Times
- (iv) Business Standard
- (v) Business Today, etc.

### **(g) Published Electronic Sources**

Today, internet is a huge source of published data because most of the published material is available on internet. Information of interest which is available on internet can easily be obtained in very short time. Some of the electronic sources of secondary data are listed below:

- (i) Online databases
- (ii) e-books
- (iii) e-journals
- (iv) Websites of different institutes or organisations or agencies, etc.

## **(2) Unpublished Sources**

Collected information in term of data or data observed through own experience by an individual or by an organisation which is in unpublished form is known as unpublished source of secondary data. Some of the sources of unpublished secondary data are given on the next page:



- (i) Records and statistics maintained by different institutions or organisations whether they are government or non-government
- (ii) Unpublished projects works, field works or some other research related works submitted by students in their corresponding institutes
- (iii) Records of Central Bureau of Investigation
- (iv) Personal diaries, etc.

After discussing sources of secondary data, natural questions which may arise in your mind are:

- (1) What are the precautions one should use before using secondary data?
- (2) What are the advantages of secondary data?
- (3) What are the limitations of secondary data?

Let us address these questions one by one.

### **(1) Precautions to be taken before using Secondary Data**

Every investigation in hand has some specific objectives and data are collected keeping these objectives in view. So, secondary data which we are planning to use in our investigation may be collected for some different objectives. Therefore some precautions which are necessary before using the secondary data in our investigation are given below.

#### **(i) Reliability of Data**

Reliability of data is judged by:

- Reliability and experience of the investigator or institution for collecting data.
- Reliability of the source(s) from where data were collected.
- Whether the proper methods of collecting data were used. Whether the sample size was proper if sample technique was used in data collection?
- Whether data collected in normal times? That is, whether data were free from periods such as floods, famines, earthquakes, etc?
- Whether data were free from the biasness of the collecting investigator or institution?

If above criteria are met, we assume, generally, that reliability of data is all right.

#### **(ii) Suitability of Data**

Suitability of data is judged by:

- Comparing the nature, scope and objectives of the investigation at hand to the original one.
- Comparing definitions of different terms and units used in original investigation to the one at hand. For example, if word “large” is used as a measurement unit then what figure it represents such as 100-200 or 500-1000 or 10000 and above, etc.
- Checking uniformity of different terms or units, i.e. we have to check whether the definition of different terms and units is maintained throughout or not.

#### **(iii) Adequacy of Data**

Adequacy of data is judged by:

- Comparing geographical area covered and to be covered in original and one at hand investigations respectively. If variation between the areas of two investigations is large then data will not be adequate. For example, data collected for the purpose of estimating per person income of a state say Delhi cannot be used to estimate the per person income in India.
- Similar argument applies on time factor also. For example, if price of a commodity are available for a particular month of a year then on the basis of prices of one particular month one cannot accurately estimate the price of that commodity for whole year.

## **(2) Advantages of Secondary Data over Primary Data**

Use of secondary data in an investigation has following advantages:

- It saves lot of time and money.
- It is easy to use.
- In some investigations primary data cannot be collected.
- The only source in case of historical documents.
- Longitudinal study can be possible.
- Secondary data complements primary data in many ways such as better understanding of the problem(s) in terms of what are gaps and deficiencies in the earlier investigation(s) which need to improve.

## **(3) Limitations of Secondary Data**

Some of the limitations of the secondary data are given below:

- It is very difficult to get secondary data which is appropriate for all objectives of our investigation at hand.
- It is very difficult to get secondary data which meet all the requirements like reliability, suitability, adequacy and accuracy.
- Secondary data are generally not available for all types of investigations.
- Data may be beyond our reach.
- Available data may be out dated.

**Example 1:** Giving reason(s) in each of the following cases, specify whether data are primary or secondary?

- (i) A Television channel telecasts the published survey report of an agency XYZ (say) based on the data collected by the agency before the general election in India to know the opinion of the people about casting their votes.
- (ii) Data source in part (i) used by agency XYZ.
- (iii) An Industrial Statistics student estimate the average life of electric bulbs of a company in which he/she works by observing the lives of a random sample of 100 bulbs.
- (iv) A Bio-Statistics student collected data from the records of 10 hospitals of a state in order to conduct his/her study, whether smoking and T.B. are associated?

**Solution:**

- (i) Secondary because TV channel used the data which already existed in a published form.

- (ii) Primary source because agency itself collected data from the field.
- (iii) Data used by the student were primary as data observed by him/her were original in character.
- (iv) Student collected data from already existed sources (i.e. records of the hospitals) so data used by the student were secondary.

Now you can try the following exercises.

---

**E 1)** Give reasons in each case whether the data are primary or secondary?

- i) 2011 census data published by Office of Registrar General of India and to be used by itself.
- ii) 2011 census data to be used by a demographer in his study.

---

**E 2)** What are the differences between primary and secondary data?

---

## 12.4 SCRUTINY OF PRIMARY DATA

---

At the time of collecting primary data, information provided by the informants is recorded either in the form of questionnaire or schedule or by other means. Before tabulating primary data, it is necessary to scrutinise the questionnaires or schedules or other means used while collecting primary data to maintain

- (i) Completeness of data
- (ii) Consistency of data
- (iii) Accuracy of data, and
- (iv) Homogeneity of data

### (i) Completeness of Data

We may find some questionnaires or schedules which are incomplete. Incomplete questionnaires or schedules, if possible, should be completed by revisiting to the informants otherwise we should reject them. Incomplete questionnaires or schedules should not be entertained.

### (ii) Consistency of Data

Sometimes, the information provided by the informants either does not make any sense or may contradicts some other information. For example,

- If an informant puts his age 40 years while age of his son as 32 years. This information does not make any sense.
- If total age and date of birth provided by the informant do not match then these two types of information contradict each other. Again either these information should be corrected by revisiting to the informants or we should reject them.

### (iii) Accuracy of Data

To maintain 100% accuracy is not an easy task. We can only do correction(s) related to certain figures by checking their sum, subtraction, etc. But we have no control if some informants provide wrong information. For example, some informants may provide wrong figure about their annual income.

### (iv) Homogeneity of Data

Keeping the homogeneity of data is also an important part of editing/scrutiny of data. Here we have to check that the units of measurements used by the informants are same or different. For example,

- Some informants may put their heights in centimeters while others may put in inches.
- Some informants may put their monthly income while others may put yearly income, etc.

To maintain homogeneity of data, we have to convert all the informations provided by all the informants in the same unit(s).

---

## **12.5 PREPARATION OF QUESTIONNAIRE**

---

We have already become familiar with the words questionnaire and schedule during our discussion of different methods of collecting primary data. It has also been mentioned there that the difference between questionnaire and schedule is that questionnaires are filled by informants themselves while schedules are filled by the enumerators/investigators. Both are formalised set of the questions for obtaining information from the informants. Questionnaire plays the role of an instrument for the investigator to get the information from the informants. So, it becomes necessary that questionnaire or schedule should be well drafted.

This whole section is devoted to discuss various aspects of the questionnaire. There are no hard and fast rules for preparing a questionnaire or schedule. Actually designing of a good questionnaire is an art not a science. Development of this art requires skill, experience, dedication and hard work. Following are some guidelines to improve the quality of the questionnaire.

### **(1) Identification of Objectives and Target Population**

First of all, we identify objectives of the investigation at hand. Once our objectives are clear we can accordingly set the type of information needed and target population. Target population is very important because questions appropriate for one group may not be appropriate for other, e.g. questions appropriate for university students may not be appropriate for farmers.

### **(2) Selection of Method**

After setting the information needed and the target group, we have to decide one of the methods of primary data collection from amongst those studied in Sec 12.2. The method used for the data collection also affects the questionnaire. For example, in direct personal interviews method, investigator personally contacts the informants and hence lengthy, complex and varied questions can be asked whereas in the telephone method we have to ask simple and short questions because on telephone lengthy questions may confuse the informants.

### **(3) Content of Individual Questions**

Keep only those questions in the questionnaire which directly or indirectly contribute to the information needed or serve some specific purpose. Unnecessary questions should not be included in the questionnaire.

### **(4) Wording of the Questions**

Whatever method of data collection is used, one should maintain the following features in the questionnaire:

- Language should be simple, i.e., use of technical terms should be avoided unless informants are technically trained.
- Questions should be short, simple and easy to understand.

- Questions should have clear meanings, i.e. words having multiple meaning should be avoided. For example, words like poor, rich, etc should be avoided because meaning of these words vary from person to person.
- Questionnaire should be self explanatory. That is, if a particular question needs some clarification, it should be provided with the help of footnote.
- Questions which are sensitive and personal in nature should be avoided.

### (5) Sequence of the Questions

Arrangement of the questions in the questionnaire should be logical and all questions which are related to a particular topic should be asked before beginning of a new topic.

### (6) Types of Questions

Main types of the questions which are generally used in a questionnaire are given below:

- **Dichotomous Questions:** Questions having only two alternatives like yes or no, agree or disagree, etc. are known as dichotomous questions. For example, Have you visited USA?
- **Multiple Choice Questions:** Questions having more than two alternatives are known as multiple choice questions. For example, which soap do you use for bath?
  - (i) Lux
  - (ii) Dettol
  - (iii) Dove
  - (iv) Kanti
  - (v) Other
- **Specific Information Questions:** Questions which are used to know some specific information from the informants are known as specific information questions. For example, which is your favourite soft drink?
- **Open Ended Questions:** Those questions which provide freedom to the informants to give their opinion are known as open ended questions. For example, what are your hobbies?
- **Scale Questions:** Consider following question:  
Are you satisfied with the service provided by your mobile company?  
Options for this type of questions are based on Likert scale as given below (Refer Unit 11):

Highly satisfied	Satisfied	Neutral	Dissatisfied	Highly dissatisfied
5	4	3	2	1
OR				
2	1	0	-1	-2

Such types of questions based on Likert Scale or some other scale are known as scale questions.

Any of the above types of questions can be used in a questionnaire depending on the suitability of the information required.

### (7) Size of the Questionnaire

Generally people do not like to answer lengthy questionnaire, so we should try to make the number of questions in the questionnaire the smallest possible. Usually, the number of questions in a questionnaire lies between 15-25. But less than 15 and more than 25 questions may be used depending on the type of information needed. For example, the number of questions in the questionnaire of 2011 census data collected by Office of Registrar General of India was 21.

### (8) Cross Questions

Some cross questions can also be put in a questionnaire to check the accuracy of the information provided by the informants. For example, we can put two questions:

- What is your date of birth?
- What is your age as on date?

These two questions put a cross check on each other.

### (9) Format, Layout and Quality of paper of a Questionnaire

To make the questionnaire eye catching and to give it professional appearance, one has to focus on:

- Proper line spacing
- Size and colour of headings and sub-headings
- Quality of paper used.

### (10) Pretesting of Questionnaire

This is the last step to give a final shape to a questionnaire. Pretesting means “testing the questionnaire on a small group of informants to get some valuable suggestion(s) from the informants to improve the questionnaire”. After incorporating valuable suggestions of the informants, questionnaire becomes ready to use.

This final questionnaire is sent to informants with a covering letter, stating briefly the objectives of the enquiry and with the request that please fill up the questionnaire and return it to the said address:

Let us prepare a questionnaire for the following example:

**Example 2:** A market research scholar wants to study the behaviour of Indian consumers regarding food and grocery items. Prepare a questionnaire for this purpose.

**Solution:** Questionnaire for this purpose is given on the next page:

Dear Informants,

I am conducting a market research study to explore the store choice behaviour of Indian consumers regarding food and grocery items. Information provided by you in the following questionnaire will be used only for academic purpose. So, you are requested to provide the feedback to the best of your knowledge and experience. Questionnaire is divided into three parts.

**PART A**

Put tick (✓) mark in the appropriate box.

1. **Gender:** ☐ Male ☐ Female
2. **Marital Status:** ☐ Single ☐ Married
3. **In which age (in years) group you belong?**  
☐ 0 – 20 ☐ 20 – 30 ☐ 30 – 40 ☐ 40 – 50 ☐ above 50
4. **Education Level**  
☐ Below 10<sup>th</sup> ☐ 10<sup>th</sup> ☐ 10 + 2  
☐ Graduation ☐ PG and above
5. **Monthly Income in Rupees**  
☐ Below 10000 ☐ 10000 – 20000 ☐ 20000 – 40000  
☐ 40000 – 60000 ☐ Above 60000
6. **Employment Status**  
☐ Employed ☐ Unemployed ☐ Business person  
☐ House Wife ☐ Other (Please specify)
7. **Store where I often buy my most of food and grocery items**  
☐ Kirana Store ☐ Kirana Store with self service  
☐ Chain Store – Individual Outlet (Please specify)  
☐ Chain Store – Located in Mall (Please specify)

**PART B**

8. **Please put a tick (✓) mark according to your importance level for each of the following factors while selecting a food and grocery retail store.**

Factor	Not at all Important	Not Important	Sometimes Important	Important	Extremely Important
Quality of product					
Variety of products					
Proper product display					
Price					
Spaciousness of the store					
Distance from residence					
Knowledge of sales person					
Behaviours of the workers of the store					
Location of store					
Music					
Air conditioning facility					
Cleanliness of the store					
Store lighting					

**Matrices, Determinants  
and Collection of Data**

Parking facility					
Payment options (debit card, credit card, cash)					
Checkout time/billing time					
Home delivery facility					
Easy return and exchange					
Complaint handling					
Long opening hours					
Low crowd size					
Clear indication of prices					
Advance communication of discounts and offers					
Facility for membership holders					
Various offers and schemes					

9. Give your response relating to the following statements by putting a tick (✓) mark in the appropriate box. Some of the statements may or not be applicable to you.

Statement	Strongly disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
I wait for special offers to buy food and grocery products					
I like to compare prices of different stores before buying the items					
I like to try the new grocery outlets					
I frequently look for new products					
I make unplanned visits to stores					
I usually shop from a nearest grocery store					
I tend to buy from a particular grocery store					
I like to shop with my family					
I like to shop alone					

Now you can try the following exercise.

- E 3)** Suppose you are a member of the team involved in preparing units of this block. Being a team member you want to get the feedback of the learners related to this material so that at the time of revising the material valuable suggestions of the learners can be incorporated. Prepare a questionnaire to mail your learners.



## 12.6 SUMMARY

In this unit, we have discussed the following topics:

- 1) Primary data and different methods of collection of primary data.
- 2) Secondary data and its major sources.
- 3) Scrutiny of primary data.
- 4) Preparation of questionnaires.

## 12.7 SOLUTIONS/ANSWERS

**E 1)** (i) 2011 census data will be considered as primary data for Office of Registrar General of India because census data were collected by this office after every 10 years.

(ii) 2011 census data will be secondary data for the demographer because for him it will be an already existed source.

**E 2)** Difference between primary and secondary data is explained in the following table:

Factor of Difference	Primary Data	Secondary Data
Time	Long time is required for collection	Less time is required for collection
Money	Needs more money	Needs less money
Reliability	More reliable	Less reliable
Suitability	More suitability	Less suitable
Adequacy	More adequacy	Less adequate
Hand	First hand data	Second hand data
Precaution	Needs no extra precautions	Needs many precautions like reliability, suitability, adequacy and accuracy
Manpower	Needs more manpower	Needs less manpower

**E 3)** Answer of this exercise is not unique. One of the possible answers is given below:

Dear Learner,

During the study of the units of this block you may have found certain portions of the material where you faced some difficulty to grasp.

We believe that there is always a scope for improvement. So we wish to know your difficulties and valuable suggestions to improve the material. Therefore, we request you to fill out and send us the following questionnaire. If you find that space provided is insufficient, kindly use a separate sheet.

## QUESTIONNAIRE

1. Enrolment No

2. Mathematical Background

☐ Up to 10<sup>th</sup>

☐ Up to 10 + 2

☐ Up to Graduation

☐ Higher than Graduation

3. How many hours did you need for studying the units?

Unit Number	9	10	11	12
Number of hours				

4. Provide your feedback unit wise by putting a tick (✓) mark in the appropriate box based on your experience get at the time of studying the units of this block.

Item	Unit No	Excellent	Very Good	Good	Poor	Very Poor	Give specific example(s), in case of poor and very poor
Unit structure	9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Way of presenting the content	9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Conceptual clarity	9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Language and style used	9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Examples, exercises used	9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

5. Number of diagrams and flow charts used

Unit 9 ☐ Sufficient ☐ Insufficient

Unit 10 ☐ Sufficient ☐ Insufficient

Unit 11 ☐ Sufficient ☐ Insufficient

Unit 12 ☐ Sufficient ☐ Insufficient

6. Any other comments

Mail to  
Course Coordinator (MST-001)  
Statistics Discipline, SOS  
IGNOU, Maidan Garhi  
New Delhi – 110068, India