

Chapter 2

Collection of Data

Sources of Data There are two sources of data

- **Primary Source of Data** It implies collection of data from its source of origin.
- **Secondary Source of Data** It implies collection of data from some agency or institution which already happens to have collected the data through statistical survey.

Types of Data There are two types of data

- **Primary Data** Data collected by the investigator for his own purpose for the first time, from beginning to end are called primary data.
- **Secondary Data** These data have already been collected by somebody else, these are available in the form of published or unpublished report.

Principal Differences between Primary and Secondary Data

- Primary data are original and secondary data are already in existence and therefore, are not original.
- Primary data do not need any adjustment, secondary data need to be adjustment to suit the objective of study in hand.
- Primary data are expensive and secondary data are less expensive.

Statistical Methods of Data Collection

(i) Direct Personal Investigation

It is the method by which data are personally collected by the investigator from the information. Merits and demerits of this method are follows.

(a) Merits

- Originality
- Reliability
- Uniformity
- Accuracy
- Related information
- Elastic

(b) Demerits

- Difficult to cover wide areas
- Costly
- Personal bias
- Limited coverage

(ii) Indirect Oral Investigation

It is the method by which information is obtained not from the persons regarding whom the information is needed. It is collected orally from other persons who are expected to possess the necessary information. Merits and demerits of this method are given below

(a) Merits

- Wide coverage
- Expert opinion
- Simple
- Less expensive
- Free from bias

(b) Demerits

- Less accurate
- Doubtful conclusions
- Biased

(iii) Information from Local Sources or Correspondents

Under this method, the investigator appoints local persons or correspondents at different places. Merits and demerits of this method are given below

(a) Merits

- Economical
- Wide coverage
- Continuity
- Suitable for special purpose

(b) Demerits

- Loss of originality
- Lack of uniformity
- Personal bias

- Less accurate
- Delay in collection

(iv) Information Through Questionnaires and Schedules

There are two ways of collecting information on the basis of questionnaire

(a) Mailing Method Under this method questionnaires are mailed to the informants. The method is most suited when

- The area of the study is very wide.
- The informants are educated.

(b) Enumerator's Methods Under this Method enumerator himself fills the schedules after seeking information from the informants. This method is mostly used when

- field of investigation is large.
- the investigation need specialised and skilled investigation.
- the investigators are well versed in the local language and cultural norms of the informants.

(c) Collection of Secondary Data There are two main sources of secondary data

- Published sources
- Unpublished sources

(d) Published Sources Some of the published source of secondary data are

- Government publication
- Semi-government publication
- Reports of committees and commissions
- Publications of trade associations
- Publication of research institutions
- Journals and papers
- Publication of research scholars
- International publication

(e) Unpublished Sources These data are collected by the government organisations and others, generally for their self use or office record.

- In order to assess the reliability, suitability and adequacy of the data, the following points must be kept in mind

- Ability of the collecting organisation
- Objective and scope
- Method of collection
- Time and condition of organisation
- Definition of the unit
- Accuracy

(v) Census 'Method

Census method is that method in which data are collected covering every item of the universe or population relating to the problem under investigation. Merits and demerits of this method are given follows

(a) Merits

- Reliable and accurate
- Less biased
- Extensive information
- Study of diverse characteristic
- Study of complex investigation
- Indirect investigation

(b) Demerits

- Costly
- Large manpower
- Not suitable for large investigation

(vi) Sample Method

It is that method in which data is collected about the sample on a group of items taken from the populations for examination and conclusions are drawn on their basis. Merits and demerits of this method are given below

(a) Merits

- Economical
- Time saving
- Identification of error
- Large investigation
- Administrative convenience
- More scientific

(b) Demerits

- Partial
- Wrong conclusions
- Difficulty in selecting representative sample
- Difficulty in framing a sample
- Specialised knowledge

Methods of Sampling

(i) **Random Sampling** Random sampling is that method of sampling in which each and every item of the universe has equal chance of being selected in the sample. Random sampling may be done in any of the following ways

- Lottery method
- Tables of random number

(ii) **Purposive or Deliberate Sampling** It is that method in which the investigator himself makes the choice of the samples items which in his opinion are the best representative of the universe.

(iii) **Stratified or Mixed Sampling** According to this method of sampling population is divided into different strata having different characteristics and some of the items are selected from each strata, so the entire population gets represented.

(iv) **Systematic Sampling** According to this methods, units of the population are numerically, geographically and alphabetically arranged. Every nth item of the numbered is selected as a sample item.

(v) **Quota Sampling** In this method, the population is divided into different groups or classes according to different characteristics of the population.

(vi) **Convenience Sampling** In this method, sampling is done by the investigator in such a manner that suits his convenience.

Reliability of Sampling Data

It depends mainly on the following factors

- Size of the sample
- Method of sampling
- Bias of correspondents and enumerators
- Training of enumerators

Important agencies at the national level which collect process and tabulate the statistical data. NSSO (National Sample Survey Organisation), RGI (Registrar General of India), DGCIS (Directorate General of Commercial Intelligence and Statistics) and Labour Bureaus.