

Chapter

Collection of Data

Collection: It is the process of collecting data or information from different sources. A place from data obtained is called a data source. Statistical data can be obtained from two sources. They are primary and secondary sources.

Distinguish between primary and secondary data

PRIMARY DATA	SECONDARY DATA
It is first hand in use.	It is second hand in use.
It is collected from an inquiry.	It is collected from published sources.
It is in the form of raw materials.	It is in the form of finished products.

What are the qualities of a good questionnaire?

The main qualities of a good questionnaire are the following.

- The questionnaire should not be too long.
- A series of questions should move from general to specific.
- The questions should not be ambiguous. They should enable the respondents to answer quickly, correctly, and clearly.
- The questions should not use double negatives.
- The questions should not be a leading question, which gives a clue about how to answer.
- The questions should be precise and clear.
- The questionnaires should include both open ended and closed ended questions.

MODE OR METHODS OF DATA COLLECTION

There are three basics of collecting primary data. They are the following

- 1) **Personal interview:** this method is used when the researcher has to access all the members. The main merits and demerits of the method are the following.

ADVANTAGES	DISADVANTAGES

Highest response rate	Most expensive
Allows all types of questions	More time taking
Better for using open-ended questions.	Possibility of influencing respondents
Allows clarifications to ambiguous questions	

- 2) **Mailing questionnaire:** The data in a survey are collected by mail, such a method of database collection is called mailing questionnaire. Its advantages and disadvantages are as follows.

ADVANTAGES	DISADVANTAGES
Least expensive	Cannot be used by illiterates
The only method to reach remote areas	Long response time
No influence on respondents	Does not allow clarifications to ambiguous questions
Maintains anonymity of respondents	Reactions can't be watched
Best for sensitive questions	

- 3) **Telephone interview:** In a telephone interview, the investigator asks questions over the telephone. Its advantages and disadvantages are as follows.

ADVANTAGES	DISADVANTAGES
Relatively low cost	Limited use
Relatively less influence on respondents	Reactions can't be watched
The relatively high response rate	Possibility of influencing the respondents

PILOT SURVEY

When the questionnaire is ready, it is advisable to conduct a try out with a small group of The population is known as a pilot survey or pre-testing of the questionnaire. It's advantages Are the following

- It helps us to access the suitability of the questionnaire.
- It helps us to identify the drawbacks and shortcomings of the questionnaire.
- It helps us to calculate survey cost and time

- It helps us to pre-testing of the questionnaire.

CENSUS OR COMPLETE ENUMERATION: A survey which includes every element of the population is called census of complete enumeration. In India census are conducted by the Registrar General of India(R.G.I.). In India census, which was carried out every ten years. The last census of India was held in the year 2011. According to the 2011 census 'India's population is 121crores. The main advantages and disadvantages of the census are the following.

ADVANTAGES	DISADVANTAGES
Additional information is to be obtained	More time taking
More reliable information obtained	More expensive
Covers the entire population	More enumerators needed

- **SAMPLE SURVEY:** Data or information is collected from samples only; such a method of data collection is called a sample survey. A sample refers to a group or section of the population from which information is to be obtained. A good sample is generally smaller than the population and gives reasonably accurate information about the population. In India, sample surveys are conducted by **NSSO**. Most of the surveys are sample surveys because of the following reasons.
 - i. It provides reasonably reliable and accurate information.
 - ii. It needed lower cost and shorter time
 - iii. More detailed information can be collected
 - iv. A smaller team of enumerators is needed

According to the selection of samples, a .sample survey was divided into two. They are the following.

RANDOM SAMPLING: In random sampling, every individual has an equal chance of being selected as a sample. The individuals who are selected are just like the ones who are not selected. In random sampling, samples are selected with the help of random number tables or Lott's so this method is also called the lottery method.

NON-RANDOM SAMPLING.: In non-random sampling, all the units of the population do not have an equal chance of being selected as a sample. The convenience or judgment of the

investigator plays an important role in the selection of the sample.

SAMPLING ERRORS: The difference between the actual value of the parameter of the population and its estimate is the sampling error. It is possible to reduce the magnitude of sampling error by taking a larger sample.

NON SAMPLING ERRORS: Nonsampling errors are more serious than sampling errors because sampling error can be minimized by taking a larger sample. It is difficult to minimize nonsampling errors, even taking a large sample. some of the non-sampling errors are:

- Errors in data acquisition: This type of error arises from the recording of incorrect responses.
- Nonresponse errors: It occurs if an interviewer is unable to contact a person listed in the sample.
- Sampling bias: Sampling bias occurs when the investigator performs biased in the selection of samples.

List out some data collecting agencies in India.

- Central statistical organization. **(CSO)**
- National sample survey organization. **(NSSO)**
- Registrar General of India **(RGI)**
- Directorate general of commercial intelligence and statistics **(DGCIIS)**
- Labour bureau.