

INDIAN SCHOOL MUSCAT
DEPARTMENT OF COMMERCE AND HUMANITIES
STATISTICS FOR ECONOMICS

CHAPTER 2: COLLECTION OF DATA

INTRODUCTION

The purpose of collection of data is to collect evidence for reaching a sound and clear solution to a problem.

When data for a period of varies from year to year they are called variable generally represented by the letters X, Y or Z. The values of these variables are the observation.

For example

The food grain production in India varies between 100 million tonnes in 1970-71 to 220 million tonnes in 2001-02 as shown in the following table. The years are represented by variable X and the production of food grain in India (in million tonnes) is represented by variable Y

Production of Food Grain in India (Million Tonnes)

X Years	Y Production
1970-71	108
1978-79	132
1979-80	108
1990-91	176
1996-97	199
1997-98	194
2001-02	212

WHAT ARE THE SOURCES OF DATA?

Statistical data can be obtained from two sources.

- **Primary Data:** The enumerator (person who collects the data) may collect the data by conducting an enquiry or an investigation. Such data are called Primary Data, OR first hand information.
- **Secondary data:** If the data have been collected and processed (scrutinised and tabulated) by some other agency, they are called Secondary Data. Generally, the published data are secondary data. They can be obtained either from published sources or from any other source, for example, a web site. Use of secondary data saves time and cost.

HOW IS DATA COLLECTED?

Data is collected by conducting a survey by asking questions about a particular product or topic from a large group of people. The purpose of the survey is to collect data. Survey is a method of gathering information from individuals.

PREPARATION OF SURVEY INSTRUMENT

The most common type of instrument used in surveys is questionnaire/ interview schedule. The questionnaire is either self administered by the respondent or administered by the researcher (enumerator) or trained investigator.

While preparing the questionnaire/ interview schedule, the following points must be kept in mind:

- The questionnaire should not be too long. Generally 10 – 15 questions. Long questionnaires discourage people from answering.
- The series of questions should move from general to specific.
- The questions should be precise and clear.
- The questions should not be ambiguous, to enable the respondents to answer quickly, correctly and clearly.
- The question should not use double negatives.
- The question should not be a leading question, which gives a clue about how the respondent should answer.
- The question should not indicate alternatives to the answer.
- The questionnaire may consist of closed ended (or structured) questions or open ended (or unstructured) questions.

Closed ended or structured questions can either be a two-way question or a multiple choice question. Closed -ended questions are easy to use, score and code for analysis, because all the respondents respond from the given options.

Open-ended questions allow for more individualised responses, but they are difficult to interpret and hard to score, since there are a lot of variations in the responses.

MODE OF DATA COLLECTION

There are three basic ways of collecting data:

1. Personal Interviews

This method is used when the researcher has access to all the members. The researcher conducts face to face interviews with the respondents.

Advantages

- Personal contact is made between the respondent and the interviewer.
- The interviewer has the opportunity of explaining the study and answering any query of the respondents.
- The interviewer can request the respondent to expand on answers that are particularly important.
- Misinterpretation and misunderstanding can be avoided.
- Watching the reactions of the respondents can provide supplementary information.

Disadvantages

- It is expensive
- It takes longer time to complete the survey.
- Presence of the researcher may hold back respondents from saying what they really think.

2. Mailing (questionnaire) Surveys

When the data in a survey are collected by mail, the questionnaire is sent to each individual by *mail* with a request to complete and return it by a given date.

Advantages

- It is less expensive.
- It allows the researcher' to have access to people in remote areas, who might be difficult to reach in person or by telephone.
- It does not allow influencing of the respondents by the interviewer.
- It also permits the respondents to take sufficient time to answers to the questions.
- These days' online surveys or surveys through short messaging service

Disadvantages

- There is less opportunity to provide assistance in clarifying instructions
- There is a possibility of misinterpretation of questions.
- Mailing is also likely to produce low response rates due to factors such as returning the questionnaire without completing it, not returning the questionnaire at all, loss of questionnaire in the mail itself, etc.

3. Telephone Interviews

In a telephone interview, the investigator asks questions over the telephone.

Advantages

- They are cheaper than personal interviews and can be conducted in a shorter time.
- They allow the researcher to assist the respondent by clarifying the questions.
- Telephone interview is better in the cases where the respondents are reluctant to answer certain questions in personal interviews.

Advantages	Disadvantages
<ul style="list-style-type: none"> • Highest Response Rate • Allows use of all types of questions • Better for using open-ended questions • Allows clarification of ambiguous questions. 	<ul style="list-style-type: none"> • Most expensive • Possibility of influencing respondents • More time taking.
<ul style="list-style-type: none"> • Least expensive • Only method to reach remote areas • No influence on respondents • Maintains anonymity of respondents • Best for sensitive questions. 	<ul style="list-style-type: none"> • Cannot be used by illiterates • Long response time • Does not allow explanation of unambiguous questions • Reactions cannot be watched.
<ul style="list-style-type: none"> • Relatively low cost • Relatively less influence on respondents • Relatively high response rate. 	<ul style="list-style-type: none"> • Limited use • Reactions cannot be watched • Possibility of influencing respondents.

PILOT SURVEY

Once the questionnaire is ready, it is advisable to conduct a try-out with a small group which is known as **pilot OR Pre-Testing** of the questionnaire.

Advantages

- The pilot survey helps in providing a preliminary idea about the **survey**.
- **It** helps to know the shortcomings and drawbacks of the questions.
- It also helps in assessing the suitability of questions, clarity of instructions, Performance of enumerators and the cost and time involved in the actual survey.

CENSUS AND SAMPLE SURVEYS

Census Survey

A survey, which includes every element of the population, is known as Census or the Method of Complete Enumeration.

The essential feature of this method is that this covers every individual unit in the entire population. You cannot select some and leave out others.

Census of India is carried out every ten years. A house-to- house enquiry is carried out, covering all households in India. Demographic data on birth and death rates, literacy, workforce, life expectancy, size and composition of population, etc. are collected and published by the Registrar General of India.

Sample Survey

Population or the Universe in statistics means totality of the items under study. A population is always all the individuals/items who possess certain characteristics or a set of characteristics according to the purpose of the survey.

The first task in selecting a sample is to identify the population. Once the population is identified, the researcher selects a Representative sample. 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 is capable of providing reasonably accurate information at a much lower cost and shorter time. Most of the surveys are sample surveys.

Advantages of Sampling

- A sample can provide reasonably reliable and accurate information at a lower cost and shorter time.
- More detailed information can be collected by conducting intensive enquiries.
- A smaller team of enumerators is required and it is easier to train than and supervise their work more effectively.
- There are two main types of sampling

Random Sampling

In the random sampling, every individual has an Equal chance of being selected and the individuals who are selected are just like the ones who are not selected. This is also called lottery method. The same could be done using a Random Number Table also.

Random number tables have been generated to guarantee equal probability of selection of every individual unit (by their listed serial number in the sampling frame) in the population. They are available either in a published form or can be generated by using appropriate software packages.

Exit Polls

When an election takes place, the television networks provide election coverage. They also try to predict the results. This is done through exit polls, wherein random samples of voters who exit the polling booths are asked whom they voted for. From the data of the sample of voters, the prediction is made.

Non random Sampling

In a non-random sampling method all the units of the population do not have an equal chance of being selected and convenience or judgment of the investigator plays an important role in selection of the sample. They are mainly selected on the basis of judgment, purpose, convenience or quota and are non-random samples.

CENSUS OF INDIA AND NSSO

Some of the major agencies at the national level are Census of India, National Sample Survey Organisation (NSSO), Central statistical Organisation (CSO), Registrar General of India (RGI), Directorate General of Commercial Intelligence and Statistics (DGCIIS), Labour Bureau etc.

The Census of India provides the most complete and continuous demographic record of population. The Census collects information on various aspects of population such as the size, density, sex ratio, literacy, migration, rural urban distribution etc.

The NSSO was established by the government of India to conduct nation-wide surveys on socio-economic issues. The NSSO does continuous surveys in successive rounds. The data collected by NSSO surveys, on different socio economic subjects, are released through reports and its quarterly journal *Sarvekshana*.

NSSO provides periodic estimates of literacy, school enrolment, utilisation of educational services, employment, unemployment, manufacturing and service sector enterprises, morbidity, maternity, child care, utilisation of the public distribution system etc. Annual survey of industries, conducts crop estimation surveys, collects rural and urban retail prices for compilation of consumer price index numbers.

CONCLUSION

Economic facts, expressed in terms of numbers, are called data. The purpose of data collection is to understand, explain and analyse a problem and causes behind it. Primary data is obtained by conducting a survey. There are various agencies which collect, process, tabulate and publish statistical data. These can re-used as secondary data. However, the choice of source of data and mode of data collection depends on the objective of the study.

EXERCISES

1. What is a variable?
2. Why is data collected?
3. Name the different ways of collecting data.
4. What is Primary data? How is it collected?
5. What are the factors that must be kept in mind while preparing a good questionnaire?
6. Explain personal interview. What are its advantages and disadvantages?
7. Explain Mailing questionnaires. What are its advantages and disadvantages?
8. Distinguish between Primary and Secondary data.
9. What are samples?
10. Define population.
11. What are the different methods of sampling?
12. What are the advantages of sampling?
13. Name the agencies that collect data in India.
14. Name the publication in which collected data is published in India.