

Data Collection

What is Data Collection?

In Statistics, data collection is a process of gathering information from all the relevant sources to find a solution to the research problem. It helps to evaluate the outcome of the problem. The data collection methods allow a person to conclude an answer to the relevant question. Most of the organizations use data collection methods to make assumptions about future probabilities and trends. Once the data is collected, it is necessary to undergo the data organization process.

The main sources of the data collections methods are “Data”. Data can be classified into two types, namely primary data and secondary data. The primary importance of data collection in any research or business process is that it helps to determine many important things about the company, particularly the performance. So, the data collection process plays an important role in all the streams. Depending on the type of data, the data collection method is divided into two categories namely,

Primary Data Collection methods

Secondary Data Collection methods

In this article, the different types of data collection methods and their advantages and limitations are explained.

Primary Data Collection Methods

Primary data or raw data is a type of information that is obtained directly from the first-hand source through experiments, surveys or observations. The primary data collection method is further classified into two types. They are

Quantitative Data Collection Methods

Qualitative Data Collection Methods

Let us discuss the different methods performed to collect the data under these two data collection methods.

Quantitative Data Collection Methods

It is based on mathematical calculations using various formats like close-ended questions, correlation and regression methods, mean, median or mode measures. This method is cheaper than qualitative data collection methods and it can be applied in a short duration of time.

Qualitative Data Collection Methods

It does not involve any mathematical calculations. This method is closely associated with elements that are not quantifiable. This qualitative data collection method includes interviews, questionnaires, observations, case studies, etc. There are several methods to collect this type of data. They are

Observation Method

Observation method is used when the study relates to behavioural science. This method is planned systematically. It is subject to many controls and checks. The different types of observations are:

Structured and unstructured observation

Controlled and uncontrolled observation

Participant, non-participant and disguised observation

Interview Method

The method of collecting data in terms of verbal responses. It is achieved in two ways, such as

Personal Interview – In this method, a person known as an interviewer is required to ask questions face to face to the other person. The personal interview can be structured or unstructured, direct investigation, focused conversation, etc.

Telephonic Interview – In this method, an interviewer obtains information by contacting people on the telephone to ask the questions or views, verbally.

Questionnaire Method

In this method, the set of questions are mailed to the respondent. They should read, reply and subsequently return the questionnaire. The questions are printed in the definite order on the form. A good survey should have the following features:

Short and simple

Should follow a logical sequence

Provide adequate space for answers

Avoid technical terms

Should have good physical appearance such as colour, quality of the paper to attract the attention of the respondent

Schedules

This method is similar to the questionnaire method with a slight difference. The enumerations are specially appointed for the purpose of filling the schedules. It explains the aims and objects of the investigation and may remove misunderstandings, if any have come up. Enumerators should be trained to perform their job with hard work and patience.

Secondary Data Collection Methods

Secondary data is data collected by someone other than the actual user. It means that the information is already available, and someone analyses it. The secondary data includes magazines, newspapers, books, journals, etc. It may be either published data or unpublished data.

Published data are available in various resources including

- Government publications
- Public records
- Historical and statistical documents
- Business documents
- Technical and trade journals

Unpublished data includes

- Diaries
- Letters
- Unpublished biographies, etc.