

# Sources of Data Collection | Primary and Secondary Sources

## What is Data?

**Data** is a collection of measurements and facts and a tool that help an individual or a group of individuals to reach a sound conclusion by providing them with some information. It helps the analyst understand, analyze, and interpret different socio-economic problems like unemployment, poverty, inflation, etc. Besides understanding the issues, it also helps in determining the reasons behind the problem to find possible solutions for them. Data not only includes theoretical information but some numerical facts too that can support the information. The collection of data is the first step of the statistical investigation and can be gathered through two different sources, namely, primary sources and secondary sources.



## Table of Content

- Sources of Collection of Data
- Primary and Secondary Data
- Principle Difference between Primary and Secondary Data
- Methods of Collecting Primary Data
- Sources of Collecting Secondary Data

## Sources of Collection of Data

### 1. Primary Source

It is a collection of data from the source of origin. It provides the researcher with first-hand quantitative and raw information related to the statistical study. In short, the primary sources of data give the researcher direct access to the subject of research. For example, statistical data, works of art, and interview transcripts.

## 2. Secondary Source

It is a collection of data from some institutions or agencies that have already collected the data through primary sources. It does not provide the researcher with first-hand quantitative and raw information related to the study. Hence, the secondary source of data collection interprets, describes, or synthesizes the primary sources. For example, reviews, government websites containing surveys or data, academic books, published journals, articles, etc.

Even though primary sources provide more credibility to the collected data because of the presence of evidence, but good research will require both primary and secondary sources of data collection.

## Primary and Secondary Data

### 1. Primary Data

The data collected by the investigator from primary sources for the first time from scratch is known as primary data. This data is collected directly from the source of origin. It is real-time data and is always specific to the researcher's needs. The primary data is available in raw form. The investigator has to spend a long time period in the collection of primary data and hence is expensive also. However, the accuracy and reliability of primary data are more than the secondary data. Some examples of sources for the collection of primary data are observations, surveys, experiments, personal interviews, questionnaires, etc.

### 2. Secondary Data

The data already in existence which has been previously collected by someone else for other purposes is known as secondary data. It does not include any real-time data as the research has already been done on that information. However, the cost of collecting secondary data is less. As the data has already been collected in the past, it can be found in refined form. The accuracy and reliability of secondary data are relatively less than the primary data. The chances of finding the exact information or data specific to the researcher's needs are less. However, the time required to collect secondary data is short and hence is a quick and easy process. Some examples of sources for the collection of secondary data are books, journals, internal records, government records, articles, websites, government publications, etc.

## Principle Difference between Primary and Secondary Data

- **Difference in Objective:** The primary data collected by the investigator is always for the specific objective. Therefore, there is no need to make any adjustments for the purpose of the study. However, the secondary data collected by the investigator has already been collected by someone else for some other purpose. Therefore, the investigator has to make necessary adjustments to the data to suit the main objective of the present study.
- **Difference in Originality:** As the primary data is collected from the beginning from the source of origin, the data is original. However, the secondary data is already present somewhere and hence is not original.

- **Difference in Cost of Collection:** The cost of collecting primary data is higher than the cost of collecting secondary data in terms of time, effort and money. It is because the data is being collected for the first time from the source of origin. However, the cost of collecting secondary data is less as the data is gathered from published or unpublished sources.

## Methods of Collecting Primary Data

**Direct Personal Investigation:** As the name suggests, the method of direct personal investigation involves collecting data personally from the source of origin. In simple words, the investigator makes direct contact with the person from whom he/she wants to obtain information. This method can attain success only when the investigator collecting data is efficient, diligent, tolerant and impartial. For example, direct contact with the household women to obtain information about their daily routine and schedule.

- **Indirect Oral Investigation:** In this method of collecting primary data, the investigator does not make direct contact with the person from whom he/she needs information, instead, they collect the data orally from some other person who has the necessary required information. For example, collecting data of employees from their superiors or managers.
- **Information from Local Sources or Correspondents:** In this method, for the collection of data, the investigator appoints correspondents or local persons at various places, which are then furnished by them to the investigator. With the help of correspondents and local persons, the investigators can cover a wide area.
- **Information through Questionnaires and Schedules:** In this method of collecting primary data, the investigator, while keeping in mind the motive of the study, prepares a questionnaire. The investigator can collect data through the questionnaire in two ways:
  - Mailing Method:** This method involves mailing the questionnaires to the informants for the collection of data. The investigator attaches a letter with the questionnaire in the mail to define the purpose of the study or research. The investigator also assures the informants that their information would be kept secret, and then the informants note the answers to the questionnaire and return the completed file.
  - Enumerator's Method:** This method involves the preparation of a questionnaire according to the purpose of the study or research. However, in this case, the enumerator reaches out to the informants himself with the prepared questionnaire. Enumerators are not the investigators themselves; they are the people who help the investigator in the collection of data.

## Sources of Collecting Secondary Data

### 1. Published Sources

- **Government Publications:** Government publishes different documents which consists of different varieties of information or data published by the Ministries, Central and State Governments in India as their routine activity. As the government publishes these Statistics, they are fairly reliable to the investigator. Examples of Government

publications on Statistics are the Annual Survey of Industries, Statistical Abstract of India, etc.

- **Semi-Government Publications:** Different Semi-Government bodies also publish data related to health, education, deaths and births. These kinds of data are also reliable and used by different informants. Some examples of semi-government bodies are Metropolitan Councils, Municipalities, etc.
- **Publications of Trade Associations:** Various big trade associations collect and publish data from their research and statistical divisions of different trading activities and their aspects. For example, data published by Sugar Mills Association regarding different sugar mills in India.
- **Journals and Papers:** Different newspapers and magazines provide a variety of statistical data in their writings, which are used by different investigators for their studies.
- **International Publications:** Different international organizations like IMF, UNO, ILO, World Bank, etc., publish a variety of statistical information which are used as secondary data.
- **Publications of Research Institutions:** Research institutions and universities also publish their research activities and their findings, which are used by different investigators as secondary data. For example, National Council of Applied Economics, the Indian Statistical Institute, etc.

## 2. Unpublished Sources

Another source of collecting secondary data is unpublished sources. The data in unpublished sources is collected by different government organizations and other organizations. These organizations usually collect data for their self-use and are not published anywhere. For example, research work done by professors, professionals, teachers and records maintained by business and private enterprises.