

Lecture-3: Methods of Data Collection

Discuss some important sources of Data collection in research. [LQ-2021, LQ-2019, LQ-2017]

What is the significance of observation in research? [SQ-2021]

What is Sampling? [SQ-2021]

- **Research Data is any information** that has been collected, observed, generated or created for the purposes of **conducting any research** i.e., for finding a solution to a problem or answering a question.
- There are **two main sources of data** i.e., **primary and secondary sources**. According to these sources, data are classified as primary and secondary.
- The data published or used by a researcher or an organization which **originally collected them** are called **primary data**. Thus, the primary data are the **first-hand information collected**, compiled and published by an organization for a certain purpose.
- The data published or used by a researcher or an organization **other than the one** which originally collected them are known as **secondary data**. Many official and private publications contain secondary data.
- **Sampling:** Sampling may be defined as the **selection of some part of an aggregate** or totality on the basis of which a **judgement or inference** about the aggregate or totality is made. A sample study is usually **less expensive, faster, more accurate and sometimes the only choice** for data collection, e.g., in a research related to all the **voters' behavior in a city**, samples of voters from a few areas of that city can be helpful to collect data.

I. Methods for the Collection of Primary Data:

i. Direct Personal Observation.

- The observation method is the **most commonly used** method specially in studies relating to **behavioural sciences**.
- Under the observation method, the information is sought **by way of investigator's own direct observation without asking from the respondent**. For instance, in a study relating to **voting behaviour**, the investigator instead of asking the choice of political party by the voters, may himself/herself look at the people's participation in the meetings, processions or rallies of different political parties.
- Observation becomes a scientific tool and the method of data collection for the researcher, when it serves a **formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability**.
- Observation can be **Structured or Unstructured** depending on whether or not, 'the observable units are carefully defined', 'the conditions of observation are standardized'.
- Observation can be of **participant and non-participant types** in the context of social sciences. This distinction depends upon **the observer's sharing or not sharing the life of the group** he/she is observing.
- Observation may also be called **Controlled** (carried out in a laboratory or under controlled conditions) and **Uncontrolled** (in the natural setting).
- This method is **particularly suitable** in studies which deal with subjects (i.e., respondents) **who are not capable of giving verbal reports of** their feelings for one reason or the other.

- But this method has **various limitations**. It is **expensive, time consuming, laborious** and there is also the **danger of being flawed with subjective interpretation**. It may provide very **limited** information. Sometimes **unforeseen factors** may interfere with the observational task. some people are rarely accessible to direct observation.

ii. The Survey Method:

- The **most common method** of generating primary data is through surveys.
- Survey is a **research technique** in which information is gathered from a sample of people using a **questionnaire**.
- The questionnaire was invented by **Sir Francis Galton (1822 - 1911)**.
- The task of **writing a list of questions and designing the exact format** of the printed or written questionnaire is an essential aspect of the development of survey research design.
- Research investigators may choose to contact the respondents **in person, by telephone, by mail, or on the internet**.
- Each of these techniques has advantages and disadvantages. The researcher's task is to **choose the most appropriate one** for collecting the information needed.
- The **advantages** of this method include **low cost, approachability convenience, large universe, freedom from the interviewer's bias and respondents' comfort & convenience**.
- The **disadvantages** of this method include, **time consuming, low rate of response, ambiguous replies, dependence on respondents' abilities, lesser control and an inbuilt inflexibility of method**.

iii. The Interview Method:

- Interview is the **most common method used in social sciences** research, e.g. while doing research on Pak-US Relations, Interviews from ex-foreign ministers, ambassadors, authors or other relevant people.
- Interviews basically consist of **asking questions, listening to individuals and recording** their responses.
- This method can be used through **personal interviews**, through the **use of telephonic or computer/internet resources**.
- Interviews can be **structured or unstructured** on the basis of rigidity of procedure, form & order of questions and freedom available to the interviewer. A mixture of these can be called as semi-structured interview.
- **Chief Merits** of the Interview Method are '**prompt and quick response, 'more information', 'greater depth', 'supplementary information', 'greater flexibility', 'room for the use of observation method', 'better control', 'avoidance of misinterpretations', etc.**
- **Some weaknesses** of the interview method are '**relatively expensive', 'time-consuming', 'dependence on proper rapport with respondents', 'possibility of the bias of interviewer as well as that of the respondent', 'over-stimulation or extra cautiousness of the respondent', etc.**

II. Methods for the Collection of Secondary Data:

- Published Printed Sources:** There are varieties of published printed sources. Their **credibility** depends on many factors. For example, on the **writer, publishing company and time and date** when

published. **New sources are preferred** and old sources should be avoided as new technology and researches bring new facts into light.

- **Books:** Books are the **most authentic** one in secondary sources. The use of books starts **even before the selection of the topic**. After selection of topics books **provide insight** on how much work has already been done on the same topic and **a better literature review** can be prepared from them.
- **Journals/periodicals:** Journals and periodicals are becoming **more important** as they provide up-to-date information which at times books cannot. Moreover, journals can give **information on the very specific topic** of research rather than the general topics.
- **Magazines/Newspapers:** Magazines are also effective but **not very reliable**. Newspapers on the other hand are more reliable and in some cases the information can only be obtained from newspapers as in the case of some political studies.
- **Government Records:** Government records are very important for marketing, management, humanities and social science research, e.g., **Census Data/population** statistics, health records, educational institutes' records etc.
- **Non-Governmental & Private Sector Records:** NGOs' survey data e.g., Human Rights' Reports of Amnesty International are considered as a good source of secondary data as in the case of some other private companies' records, e.g., the records of firms of overseas investments in the Panama Scandal.

ii. Published Electronic Sources: As **internet is becoming more advance, fast and reachable** to the masses, secondary data is **readily available**, even that which is not available in printed form. The **credibility of internet has increased** today because almost every journal and book is available online, either free or with price.

- **e-journals** are more commonly available than printed journals. Latest journals are difficult to retrieve without subscription but if your university has an e-library you can view any journal, print it and those that are not available you can make an order for them.
- **General Websites:** Generally, websites do not contain very reliable information so their content should be checked for the **reliability** before quoting from them.
- **Weblogs:** Weblogs are **actually diaries** written by different people. These diaries are as reliable to use as personal written diaries.

iii. Unpublished Personal Records: Some unpublished data may also be useful in some cases.

- **Diaries:** Diaries are personal records and are rarely available but they might be very useful. **The Anne Frank's diary** is the most famous example of this. That diary contained the most accurate records of **Nazi wars**.
- **Letters:** Letters like diaries are also **a rich source** but should be checked for their reliability before using them.
- **Private Audio/Video Recordings** can also be useful sources of secondary data if their reliability is established.