Research Methodology

Data Collection – Lesson – 4

Content

- **Data and its Type**
- **Data Types Explained**
- **Techniques of Primary Data Collection – Direct Personal Interviews**
- **Techniques of Primary Data Collection – Questionnaires**
- **Techniques of Primary Data Collection – Schedules**
- **Techniques of Primary Data Collection – Observations**
- **Techniques of Primary Data Collection – Focus Groups**
- **Techniques of Primary Data Collection – Experiment**
- **Techniques of Primary Data Collection – Case Study**
- **Sources of Secondary Data Collection**
- **Pros and Cons of Secondary Data Collection**
- **Questionnaire**

- **Steps in Questionnaire**
- **Initial Consideration**
- **Self-Administered or Interviewer Assisted Questionnaire?**
- **Clarification of Concept**
- **Topology of Questionnaire**
- **Questionnaire Design**
- **Open Ended Questions**
- **Closed Ended Questions**
- **Preliminary Decisions in Questionnaire Design**
- **Successful Interviewing**
- **Pre-Testing**
- **Pilot Study**

Data and its Type

- Data is a set of values of subjects with respect to qualitative or quantitative variables.

- Data is raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organized.

- When data is processed, organized, structured or presented in a given context so as to make it useful, it is called information.

- Following are the types of data:
  ► Primary data
  ► Secondary data
  ► Cross-sectional data
  ► Categorical data
  ► Time series data
  ► Spatial data
  ► Ordered data

Data Types Explained

- **Primary Data**
  - ► Primary data is an original and unique data, which is directly collected by the researcher from a source according to his requirements.
  - ► It is the data collected by the investigator himself or herself for a specific purpose.

- **Secondary Data**
  - ► Secondary data refers to the data which has already been collected for a certain purpose and documented somewhere else.
  - ► Data collected by someone else for some other purpose (but being utilized by the investigator for another purpose) is secondary data.

- **Categorical Data**
  - ► Categorical variables represent types of data which may be divided into groups. Examples of categorical variables are race, sex, age group, and educational level.
  - ► The data, which cannot be measured numerically, is called as the categorical data. Categorical data is qualitative in nature.
  - ► The categorical data is also known as attributes.

- **Cross Sectional Data**
  - ► Cross-sectional data is a type of data collected by observing many subjects (such as individuals, firms, countries, or regions) at the same point of time, or without regard to differences in time.
  - ► It is the data for a single time point or single space point.

    This type of data is limited in that it cannot describe changes over time or cause and effect relationships in which one variable affects the other.

- **Time Series Data**
  - ► Time series data occurs wherever the same measurements are recorded on a regular basis.
  - ► Quantities that represent or trace the values taken by a variable over a period such as a month, quarter, or year.
  - ► The values of different phenomenon such as temperature, weight, population, etc. can be recorded over a different period of time.
  - ► The values of the variable remain increasing or decreasing or constant.

- **Spatial Data**
  - ► Also known as geospatial data or geographic information it is the data or information that identifies the geographic location of features and boundaries on Earth, such as natural or constructed features, oceans, and more.

► Spatial data is usually stored as coordinates and topology and is data that can be mapped.

► Spatial data is used in geographical information systems (GIS) and other geolocation or positioning services.

► Spatial data consists of points, lines, polygons and other geographic and geometric data primitives, which can be mapped by location, stored with an object as metadata or used by a communication system to locate end-user devices.

- **Ordered Data**

  ► Data according to ordered categories is called as ordered data.

  ► Ordered data is similar to a categorical variable except that there is a clear ordering of the variables.

Techniques of Primary Data Collection – Direct Personal Interviews

- **Interview method:**
  - Personal or Telephonic / Structured or Unstructured- Uses verbal communication

- **Pros**
  - In-depth information can be collected.
  - Non-response and response bias can be detected.
  - The samples can be controlled.

- **Cons**
  - It is more time-consuming.
  - It is expensive.
  - The interviewer may be biased.

Techniques of Primary Data Collection – Questionnaires

Surveys and questionnaires are 2 similar tools used in collecting primary data. They are a group of questions typed or written down and sent to the sample of study to give responses.

- **Pros**
  - Respondents have adequate time to give responses.
  - It is free from the bias of the interviewer.
  - They are cheaper compared to interviews.

- **Cons**
  - A high rate of non-response bias.
  - It is inflexible and can't be changed once sent.
  - It is a slow process.

Techniques of Primary Data Collection – Schedules

Schedules are similar to Questionnaires with a little difference that schedules are being filled in by the enumerators who are specially appointed for the purpose. While a questionnaire is filled by the informants themselves, enumerators fill the schedule on behalf of the respondent.

- **Pros**
  - High Reponses Rate
  - Can be used for illiterate person as well

- **Cons**
  - Induce bias from enumerator
  - Expensive
  - Small coverage

Techniques of Primary Data Collection – Observations

Observation method is mostly used in studies related to behavioral science. The researcher uses observation as a scientific tool and method of data collection. Observation as a data collection tool is usually systematically planned and subjected to checks and controls.

- **Pros**
  - The data is usually objective.
  - Data is not affected by past or future events.

- **Cons**
  - The information is limited.
  - It is expensive

Techniques of Primary Data Collection – Focus Groups

Focus Groups are gathering of 2 or more people with similar characteristics or who possess common traits. They seek open-ended thoughts and contributions from participants.

A focus group is a primary source of data collection because the data is collected directly from the participant. It is commonly used for market research, where a group of market consumers engage in a discussion with a research moderator.

- **Pros**
  - It incurs a low cost compared to interviews. This is because the interviewer does not have to discuss with each participant individually.
  - It takes lesser time too.

- **Cons**
  - Response bias is a problem in this case because a participant might be subjective to what people will think about sharing a sincere opinion.
  - Group thinking does not clearly mirror individual opinions.

Techniques of Primary Data Collection – Experiment

An experiment is a structured study where the researchers attempt to understand the causes, effects, and processes involved in a particular process. This data collection method is usually controlled by the researcher, who determines which subject is used, how they are grouped and the treatment they receive.

- **Pros**
  - It is usually objective since the data recorded are results of a process.
  - Non-response bias is eliminated.

- **Cons**
  - Incorrect data may be recorded due to human error.
  - It is expensive.

Techniques of Primary Data Collection – Case Study

Case study research is a qualitative research method that is used to examine contemporary real-life situations and apply the findings of the case to the problem under study. Case studies involve a detailed contextual analysis of a limited number of events or conditions and their relationships. It provides the basis for the application of ideas and extension of methods. It helps a researcher to understand a complex issue or object and add strength to what is already known through previous research.

- **Steps for Case Study:**
  - ► Identify and define the research question
  - ► Select the cases
  - ► Collect the data
  - ► Evaluate and analyze the data
  - ► Presentation of Results

Sources of Secondary Data Collection

Sources of secondary data includes books, personal sources, journal, newspaper, website, government record etc. Secondary data are known to be readily available compared to that of primary data. It requires very little research and need for manpower to use these sources. With the advent of electronic media and the internet, secondary data sources have become more easily accessible. Some of these sources are highlighted below.

- ► Books
- ► Published Sources
- ► Unpublished Sources
- ► Journals
- ► Newspaper
- ► Websites
- ► Government Records

Pros and Cons of Secondary Data Collection

- **Pros**
  - ► Ease of Access
  - ► Inexpensive
  - ► Time-Saving
  - ► Longitudinal and Comparative Studies

- **Cons**
  - ► Data Quality
  - ► Irrelevant Data
  - ► Exaggerated Data
  - ► Outdated Information

Questionnaire

Formalized set of questions for obtaining information from respondents

"Questionnaire is a method not an end in itself"

Steps in Questionnaire

Step 1: Initial Consideration

Step 2: Clarification of Concepts

Step 3: Typology of a Questionnaire

Step 4: Pre-Testing of a Questionnaire

Step 5: Administering a Questionnaire

Initial Consideration

Clarify nature of research problem & objectives

Identify information need & sources of information

Define target population & sampling frame

Determine sampling approach, sample size and expected response rate

Method of data collection

Self Administered or Interviewer Assisted Questionnaire?

Respondent Capabilities

Educational background

Vocabulary level

Prior experience in completing questionnaires

Age

Cultural issues

Clarification of Concept

Ensure the concepts can be clearly defined

Select the variables / indicators to represent the concept

Determine the level of measurement

Topology of Questionnaire

Determine the type of questions to include and their order

Check the wording

Decide on the grouping of questions and overall length of questionnaire

Determine the structure & length of questionnaire

Questionnaire Design

Two types of questions:

- **Open ended**

- **Closed ended**
    Single answer
    Multiple answer
    Rank order
    Numeric
    Likert type
    Semantic Differential

Open Ended Questions

Typically used in exploratory studies

Allows respondents freedom of choice

Respondent must be articulate & willing to spend time giving a full answer

Data is narrative in form – difficult to code & analyze

Possible researcher bias in interpretation

Can be analyzed using content analysis

Closed Ended Questions

Typically used in quantitative studies

Assumption – researcher has knowledge to pre-specify response categories

Data – pre coded and easy to analyze

Difficult to design but simplifies analysis

Limited range of response options

Preliminary Decisions in Questionnaire Design

There are nine steps involved in the development of a questionnaire:

1.   Decide the information required.
2.   Define the target respondents.
3.   Choose the method(s) of reaching your target respondents.
4.   Decide on question content.
5.   Develop the question wording.
6.   Put questions into a meaningful order and format.
7.   Check the length of the questionnaire.
8.   Pre-test the questionnaire.
9.   Develop the final survey form.

Successful Interviewing

- Make sure the research question is clear.
- Develop a check list of the questions to be asked during the interview.
- Express clearly the purpose of the interview.
- Start with a neutral question to facilitate free flow of information.
- Use open-ended questions so that the respondent can choose his answer.
- Limit the content of each question with a single idea to avoid confusion.
- Reduce questions that give responses of `yes' or `no', because they give limited information.
- Do not influence the respondent by asking leading questions.
- If you have not understood the response ask the respondent to repeat and clarify.
- Do not assume answers.
- Do not pass judgements.
- Avoid irrelevant discussions.
- When you change the tack, inform the respondent that you are doing so.
- Keep the interview short.
- At the end of the interview summarise the points reported and ask the respondent if the summary is correct.

Pre-Testing

- Pre-testing a research instrument entails a critical examination of the understanding of each question and its meaning as understood by a respondent. A pre-test should be carried out under actual field conditions on a group of people similar to your study population. The purpose is not to collect data but to identify problems that the potential respondents might have in either understanding or interpreting a question. Your aim is to identify if there are problems in understanding the way a question has been worded, the appropriateness of the meaning it communicates, whether different respondents interpret a question differently, and to establish whether their interpretation is different to what you were trying to convey

Pilot Study

- A Pilot Study (also known as feasibility study), is a small-scale implementation of a larger study or a part of a larger study. It is a trial run of entire study from start to finish that increases the likelihood of success of main study. It is conducted to test the entire research process from a methodological standpoint in actual field conditions.

End of Lesson 4