3. Data Collection

Dr. M. Dhanabhakyam

Professor and Dean,
Department Of Commerce,
Bharathiar University,
Coimbatore.

Monish P.

PhD Full time Research Scholar,
UGC Junior Research Fellow (JRF),
Department of Commerce,
Bharathiar University, Coimbatore.

3.1 Introduction:

Data are any facts, figures that is useful for analysis and interpretation. For example gender, age, marital status of the respondents in customer satisfaction study. Every research requires data. While conducting statistical investigation, data collection is considered as the first step. The success of statistical investigation depends on the collection, classification, processing, analysis and meaningful interpretation of data.

3.1.1 Need and Importance of Data:

- Provide meaningful information
- Provide a basis for formulation and testing of hypothesis
- Data is required for construction and selection of measurement and scales
- Provide a basis for analysis and interpretation
- The quality of the research findings is depends on the reliability and validity of the collected data
- Data helps to generalizations
- Data helps to theory building
- Data helps to government institutions and companies for policy formulation

3.1.2 Classification of Data:

Data collected by a researcher can be classified as:

A. Primary Data:

Primary data is the data, which are collected for the first time. This data have original character. The researcher collecting primary data for his study. This data meet the specific requirements of the researcher.

B. Secondary Data:

The secondary data are those data, which are already collected by other persons. The data already available can be called as secondary data. It is already collected and presented by someone else for some other purpose.

This data may or may not suitable for the purpose of other research. Secondary data may be internal secondary data or external secondary data

- **Internal Secondary Data:**

It refers to the data that already available within the company or study unit. Here the data can be generated and collected from the activities within a company. For Example sales records, budget records, market research records, job description, bin card etc.

- **External Secondary Data:**

It refers to the data, which is collected from an external institution.it may be private documents or public documents.

For example life history, diaries, letters, government publications, publications by commercial agencies etc. external secondary data may be published or unpublished data.

3.1.3 Selection of Primary and Secondary Data:

The researcher should decide whether primary or secondary data is suitable for his study. The selection process will depends upon the following factors

- Purpose of the study
- Nature of the research problem
- Accessibility of the information
- Availability of the time and cost
- Degree of accuracy and reliability for the study
- Availability of trained investigators
- Convenience of the researcher

Sources of Data:

The source of data always depends on purpose of the research. Generally it can be classified as primary sources and secondary sources.

a. Primary Sources:

These are the original or first time sources from which a researcher directly collects the essential data. For example data is collected to know the awareness of investors regarding investment opportunities

b. Secondary Sources:

These sources contain data, which has already collected or published for another purpose of study.

Apart from this general classification, sources of data can be also classified as:

- Documentary sources
- Library sources and
- Personal sources

3.1.4 Documentary Sources:

A document can provide vast information. This documentary source can be further classified in to individual documents and public documents.

a. Individual Documents:

These are the documents which are recorded and kept by individuals. A person can records his ideas, thoughts and experience in this individual document.

Types of Individual Documents:

- Life history
- Letter
- Diary and
- Memory
- **Life history:** it contains the biological information of individuals. We can collect biographical and social information from life history.
- **Letter:** it provides intimation about the views of individuals to a particular problem. So it provides an idea about the attitude and perception of the individual. The researcher should be careful in selection of appropriate letters.
- **Diary:** it is the most revealing source of personal information. A person can record certain events, feelings and opinions in his diary. So it can be taken as an evidence for the study which is carried out by a researcher
- **Memories:** some people will write their memories regarding with certain events in their life. These documents provide a strong material for social science studies.

b. Public Documents:

Public documents are kept and record by various public institutions. It can be further classified in to published records and unpublished records:

- Published records: it includes various newspapers, magazines, academic journals, historical documents, survey reports and case study reports.

- Unpublished records: these records mainly deal with the matters of public interest. Hence it is not available to the general public in published form. There is an access limitation. For example meeting proceedings of a company.

Merits of Primary Data:

- High degree of accuracy
- High level of control over data
- Suitable for purpose of the study
- It does not require extra filtering
- We are the owner of the information

Demerits of Primary Data:

- Time consuming
- Cost is comparatively high
- In certain studies, it is not possible to collect primary data
- It requires additional investigators, when the study covers a large geographical area

Merits of Secondary Data:

- Ease of access
- It provide an insight to the total situation
- It is helpful for formulation and testing of hypothesis
- Lower cost
- Helps to generate new insights based on already available information
- Suitable if the sample size is too large
- Anyone can collect data

Demerits of Secondary Data:

- It is the already collected information. So it may not be suitable for the study conducted by other researchers
- Less accuracy
- They are not up to date information
- Lack of control over the quality of data
- Not proprietary information

3.1.5 Collection of Primary Data:

Primary data can be collected through different methods. The selection of methods depends on quality of information, coverage of information, time, cost, convenience of the researcher etc.

The following are the important methods of collecting primary data.

- Observation
- Experimentation
- Interview
- Survey
- Simulation
- Projective techniques
- Socio metry
- Focus group
- Content analysis
- Panel method

3.2 Observation:

Observation is an important method of data collection. It can be used in both physical science and social science. There is no verbal communication with the respondents. Here the investigator or researcher observe or notice the things keenly. Observation can be defined as a systematic viewing of a specific phenomenon or condition in its proper setting for the specific purpose of gathering information for a particular study.

3.2.1 Types of Observation:

A. Simple and Systematic Observation: Observation is conducted without a standard procedure is known as simple observation. In systematic observation, observation is conducted based on standard procedures and training is provided for observers.

B. Subjective and Objective Observation: If the researcher observes his own immediate experience it is known as subjective observation or self-observation. When the researcher observe things, which are not related with him, is called objective observation.

C. Casual and Scientific Observation: In casual observation, there is no well preparation. Researcher observes the things by chance. If the observation is conducted based on certain measurement tools it is known as scientific observation.

D. Factual and Inferential Observation: In factual observation, factual information is collected through observation. In inferential observation, observation is conducted for drawing inferences

E. Direct and Indirect Observation: In direct observation, observer is physically present in the observation process. When the observation is conducted by using mechanical devices like cameras or close circuit TV etc., it is known as indirect observation.

F. Participant and Non-Participant Observation: When the observer participated with the actions of the group under observation, it is known as participant observation. The observer penetrates the thoughts of group members and observes their activities, behaviour and emotions. When the observer does not participate the actions of the study group, it is termed as non-participant observation.

G. Controlled and Non-Controlled Observation: In controlled observation, the observation is conducted based on pre-arranged plans. Here the researcher has control over the situation.in non-controlled observation; there is no well-defined plan. The researchers have no control over the situation.

3.3 Experimental Method:

It is the least used method in collecting primary data. In experimental method the group can be classified as experimental group and control group. Experimental group is the group, which is exposed under specific conditions. Control group is the group which is exposed under normal conditions.

3.3.1 Types of Experiment:

A. Laboratory Experiment: In laboratory experiment the researcher tries to measure the cause and effect relationship between the variables. This experiment is conducted in a well settled laboratory. This method is used in both physical science and social science.

B. Field Experiment: This experiment is conducted at the study unit. In other words it is the experiment conducted in a real life situation.

Here the researcher manipulates the independent variable and testing the hypothesis. This method is more useful in social science research.

3.3 Simulation:

Simulation is a modern technique. It is a realistic representation of an imagined situation. Simulation involves construction of various models representing the real situations. This approach is derived from physical science.

Steps in Simulation:

- Identify the situation to be stimulated
- Decided the objectives of simulation
- Develop a mathematical model based on available information
- Collect input data
- Select the type of simulation
- Operate the simulation with the help of various sets of input data

The simulation may be computer simulation, man simulation or man computer simulation.

3.4 Interview:

Interview is the face to face contact between two or more persons. In interview method the data is collected through verbal communication. It is an important tool of collection of data in social science research.

In interview method, there is a face to face communication between interviewer and interviewee. The person who is interviewing is called interviewer and the person who is answered in interview is known as interviewee.

3.4.1 Process of Interview:

- Preparation
- Introduction
- Develop a relationship
- Conduct of interview
- Recording the interview
- Closing the interview

A. Preparation: Preparation is essential for conducting an interview. The interviewer should keep a copy of interview schedule and prepare the list of the name and address of the respondents. He should mentally and physically prepare for the interview.

He should also decide about the way of introduction, the questions to be asked, way of handling the session. If he is not prepared well he is not able to collect necessary information from the respondents.

B. Introduction: Interviewer may be a stranger to the respondents. So he should properly introduce himself. There are no universal norms for introduction.

It depends on the situation of the interview. Interviewer should ensure the cooperation of the respondents through his introduction. The following tips can be used in introduction.

- Greet the respondents with a smile
- Explain the purpose of interview
- Properly address the respondents
- Explain the level of confidentiality of information\
- Explain the use of this study

C. Develop Relationship: Before starting the interview, interviewer develops a good relationship with the respondents. Start the interview with a comfortable topic for both.

D. Conduct the Interview: After ensuring good relationship between interviewer and interviewee, questions are asked to the respondents. If the questions are not understood by the respondents, it should be clearly explained by the interviewer once again.

E. Recording the Interview: Each response should be carefully recorded and reported by the interviewer. These recorded notes can be useful for further reference.

F. Closing the Interview: After the interview, close the interview by the interviewer with a smile to the respondents. If the respondents want to know their whole responses, it should be sending to them also.

3.4.2 Types of Interviews:

A. Personal Interview and Group Interview: Personal interview is conducted with a particular person, and focused on the confidential personal aspects of an individual. It is considered as an effective communication tool. In group interview, investigator interviewing two or more persons simultaneously. Data is collected from a number of respondents with a common interest.

B. Formal Interview and Informal Interview: Formal interviews are conducted in formal structure. Here the interviewer collecting information from the respondents with the help of a set of pre-determined questions.

In informal interview the interviewer has freedom to make alternations in the prepared questions. Interviewer can ask questions accidentally.

C. Focused Interview: In this interviewer give focus attention on a particular life event or experience of the respondent. Usually unstructured questions are asked in focused interview

D. Diagnostic Interview: Diagnostic interview is carried out among patients. In this, investigator tries to understand the reasons of the problem. For example a doctor diagnose the disease of a patient.

E. Treatment Interview: The objective of this interview is spread awareness. For example treatment interview conducted with patients to increase their awareness on causes of their disease and scope of further treatment.

F. Telephone Interview: If the investigator used telephone for conducting personal interview, it is termed as telephone interview. The respondent is connected through telephone and researcher gather information through telephone.

G. In-depth Interview: It is mainly used in motivational research. Here the interviewer collects information regarding the underlying motives and desire of the respondents. In depth knowledge about the situation or event can be acquired through in-depth interview.

3.5 Delphi Method:

Delphi technique is used when expert solution or answers is needed. It is a group process involving a group of experts. The facilitator selects a group of experts. Then discussion is going on with these experts on a topic of interest. They provide their view points and solution to the problem.

3.6 Panel Method:

In this method data is collected from the same sample respondents by using a mail or personal interview method. This method is widely used in consumer behaviour and advertisement effectiveness studies.

3.7 Survey Method:

Survey is a descriptive research method. Survey can be conducted for collecting responses from the sample respondents. Data can be collected through using a questionnaire or schedule. This method requires expertise, careful planning, and accuracy in interpretation of the results

3.7.1 Characteristics of Survey Method:

- It is considered as a field study
- This method can cover huge population
- This method can cover wide geographical area
- This method is applicable for intensive studies
- It is useful for describing a phenomenon and analyse the cause and effect relationship

3.7.2 Types of surveys:

The surveys can be broadly classified in to cross sectional survey and longitudinal survey

A. Cross Sectional Survey: This survey is used to collect data from the sample respondents at a single point of time. For example survey conducted to know employee job satisfaction during the year 2020. Here we can collect the data, conduct analysis and find out the results at a single point of time.

B. Longitudinal Survey: In longitudinal survey, researcher collects data from the respondents over a period of time. For example surveys conducted to know the effectiveness of advertisement before and after an advertisement campaign is conducted.

Methods of Survey:

Survey methods can be classified as census survey method and sample survey method.

Census Survey Method:

In census survey, the researcher collects data from all units of the population. Census method helps to increase accuracy of collected data. This method requires more time, cost and resources.in certain cases it is not possible to collect data from the entire units of the population. So sample survey method can be used in that situation.

Merits of Census Survey Method:

- More accurate survey method, because this method covers each and every units of the population
- More reliable method
- Collected data can be used for various surveys and depth analysis
- Errors can be easily detected

Demerits of Census Survey Method:

- It requires more time and money
- It requires large number of trained enumerators
- It is not possible to conduct a census survey, if there is an infinite population

Sample Survey Method:

In sample survey method, the researcher selected a smaller representation of the population, known as sample. The entire study is conducted based on sample respondents. The process of selecting samples from the population is known as sampling. Various sampling technics are available for selecting samples from the population.

Merits of Sample Survey Method:

- Very useful to collect primary data regarding personal information, attitudes and behaviour of the respondents
- It helps to make generalizations based on samples
- This method is comparatively less time consuming and less expensive process than census survey method

Demerits of Sample Survey Method:

- The success of this method depends on cooperation of the respondents
- There is limit in collection of information from sample respondents
- It is not suitable for historical studies
- There is a chance for sampling errors
- There is problem of reliability and validity

3.8 Projective Techniques:

This method encourages the respondents to reveal their unconscious feelings, attitude and emotions. In marketing research it is used to access the unconscious mind of the customers. In this various tests can be used to project the personality traits of the respondents.

Projective techniques are widely used in psychology to study the psychological and unconscious behaviour of people. The following are important projective techniques.

A. Word Association Test: In this test, researcher says a list of words and the participant is to respond with the first word that comes to their mind. This test is used to identify the words which have high association. Usually 50 to 100 words are distributed to respondents.

B. Sentence Completion Test: In is the extension of word association test. In this researchers distributed a list of incomplete sentences and ask respondents to complete them. The researchers can analysing the feelings of the respondents through observing the way they complete the sentence.

C. Story Completion Test: In this test researcher distributed stories to the respondents instead of sentence. Then the respondents are asked to complete the story.

D. Verbal Projection Test: In verbal projection test, respondents are asked to explain about the situation or comment on what other people do. Researcher observes the explanation and find out the motives and behavioural patterns of the respondents.

E. Cartoon Projection Test: In this test researcher shows various characters in a situation or event and ask respondents to describe what the characters are doing and explain about the situation in cartoon.

F. Pictorial Techniques: Many pictorial techniques are also used under projective techniques such as thematic appreciation test, Rorschach in blot test, holtzman inkblot test, rosensweig test etc.

3.9 Sociometry:

Sociometry is the technique which is used to measuring the degree of relatedness among the peoples in society. Sociometry is the best technique for reducing conflict and improves communication in a group. This method is developed by Jacob Levi Moreno in 1934. Sociometry is a powerful technique for improving the interrelationship between the members in group.

Objectives of Sociometry:

- To explore social choice patterns of people
- To reduce the conflicts in a group
- To facilitate constructive changes in individuals
- To defines the roles and responsibilities of members in a group
- To improve the interpersonal relationship between people
- To increase group cohesiveness

3.10 Focus Group:

It is a technique in which a skilled expert or facilitator leads a small group of participants in an unstructured discussion about a specific topic. The moderator is the leader of focus group discussions.

Uses of focus group discussion:

- To assist for the preparation of questionnaire
- To assess the needs of people
- To identify the factors while people making decisions
- To test existing policies and procedures
- To recruit new employees in an organization
- To evaluate the effectiveness of a marketing program

Steps in focus group discussion:

- Determine purpose of the focus group
- Selection of the participants
- Intimation to participants
- Providing fee and refreshment
- Selection of facilitator
- Formulation of protocol for the session
- Proper recording of discussion
- Conduct discussion and follow up

3.11 Content Analysis:

Under content analysis, the most important contents of any documentary materials such as books, magazines, articles are analysed. In content analysis the researcher categorize the entire data. This technique is most popular in social studies.

Steps in content analysis:

a. To identify the universe of content that is analysed under the study
b. To find out the specific characteristics to be measured under content analysis and then the universe should be divided on the basis of these characteristics.
c. Categories it on the basis of objectives of the research
d. Quantify this data in a suitable form, for example ranking method.

Tools for collecting primary data:

Various tools are used for collecting primary data from the respondents. Among these tools the widely used tools are questionnaire and schedule. Both these tools are adopted by private individuals, research scholars and various organizations for collecting primary data from their respondents. The key difference between questionnaire and schedule is that, questionnaire is filled by the respondents and schedule is filled by the enumerator.

A. Questionnaire:

A questionnaire is filled by respondents and it consists of a number of questions printed in a definite order on a proper form. The success of questionnaire is depends on the reliability and validity of the collected information.

- **Types of Questionnaire:**

a. Structured and Unstructured Questionnaire: Structured questionnaire is definite and concrete in nature. In the questions are presented in a structured format. Same set of standardized questions are distributed to all respondents.in unstructured questionnaire, questions are not arranged in a well manner and some of the questions are partially incomplete.

b. Fact and Opinion Questionnaire: In fact questionnaire is designed to collect factual information from the respondents.in opinion questionnaire respondent's opinion, perception or attitude regarding a particular phenomenon is collected.

c. Open form and Closed form Questionnaire: In open form questionnaire, the respondents can freely record their Reponses.

There are no restrictions regarding responses. In closed form questionnaire, short responses from respondents are collected. Most of the questions are dichotomous questions.

Guidelines for the preparation of questionnaire:

While preparing questionnaire, researcher should consider some points. The points are listed below.

- Use the language that is comfortable to the respondent
- Avoid lengthy questions
- Avoid questions having double meaning
- Forecast the frame of responses from the respondents
- Protect ego of the respondents
- Decide whether direct or indirect questions are more needed
- Decide whether general or specific questions are more needed
- Avoid biased questions
- Avoid misleading questions
- Ensure questions should be limited to single idea or concept at a time

Merits of Questionnaire:

- Economical
- Uniformity
- Free from bias
- Convenient tool
- Reliable tool

Demerits of Questionnaire:

- Limited use
- Lower response rate
- Misinterpretation of data

B. Schedule:

A schedule is a set of questions which are asked and filled by the enumerator through face to face interaction with respondents. The schedule may contains both direct and indirect questions

Steps in schedule preparation process:

- The problem of the study should be divided in to various portions. Researcher must understand the problem clearly before division
- These portions or aspects of the problem are further subdivided in to various components
- The next step is preparation of questions related with the study
- All the questions are well ordered and properly arranged in schedule

Merits of Schedule:

- Most suitable tool, when the respondents are illiterate
- Suitable tool when there is a small non-response rate
- The collected data through using schedule is more reliable

Demerits of Schedule:

- Costly tool
- In certain situations more enumerators are needed to collect data
- The success of this tool depends on efficiency of enumerators