

Module – 2

Data Collection and Processing

PRIMARY DATA:

It refers to first-hand information which is collected to solve specific problem. The primary data provides accurate data as it is collected first-hand. The researcher can obtain in-depth information.

The methods of collecting primary data includes:

1. **Observation method** - where the researcher obtains data by observation such situations or behaviours. For instance, researcher may study consumer buying behaviour at shopping centres or malls.
2. **Experimentation method** - This method is mostly used in the case of scientific research study. With the help of experimentation, the researcher may study the cause-effect relationship between two or more variables.
3. **Survey method** - It is used when the sample size is generally large. Respondents are asked certain to give replies against certain elements. Generally, open-ended questions are asked. Example includes census survey of population.
4. **Interview method** - It is similar to survey method. This is specially used in the case of commercial research. Respondents are asked open ended and close ended questions. Under interview method, in-depth information can be obtained from respondents.

The factors influencing primary data collection includes:

- **The time factor** - when the researcher has comparatively more time to conduct research as compared to that of research through secondary data.
- **The cost factor** - when the researcher has comparatively more funds to conduct research as compared to that of research through secondary data.
- **The objectives of research** - when the objectives of research are qualitative in nature, primary data is collected.
- **Availability of research staff** - when adequate number of research staff is available, the research may collect primary data.
- **Availability of respondents** - when adequate number of respondents are available, the researcher collects data from primary sources.

The main features of primary data are as follows:

1. **First Hand Information:** It is collected directly from the respondents. For example, the college may take up primary data collection source from their students to study the facilities provided by them. The College may collect

Module – 2

Data Collection and Processing

the data through survey and/or interview. The data collected through this source is original and more specific.

2. **Time-consuming:** This type of data collection is a time- consuming and a lengthy process. It involves:
 - Deciding the types and size of respondents
 - Scheduling the survey/interview
 - Gathering the responses
 - Data analysis and interpretation, etc.

3. **Expensive:** Primary data collection is an expensive affair; as il involves:
 - Training the research staff
 - Preparation of questionnaire
 - Printing of questionnaire, etc.

4. **Paper Work:** Primary data collection involves a lot of paper work in respect of.
 - Preparation of questionnaire
 - Coding and tabulation of data
 - Analysis of data, etc.

5. **Types/ Methods:** There are various methods of collecting primary data, such as:
 - Survey method
 - Observation method
 - Experimental method

6. **Availability:** The primary data is not available but needs to be gathered by the researcher with the help of his research staff. The researcher has to spend a lot of time, effort and money to collect data from primary sources. The primary data sources supplement the secondary data.

7. **Reliability:** The primary data is reliable. The researcher can check the reliability of the data. It is collected first hand. Latest data is collected from the respondents.

Module – 2

Data Collection and Processing

- 8. Accuracy:** The data collected from primary sources is more accurate. The researcher can check the accuracy of data at the time of interviews. He can check the accuracy by cross- questioning the respondents, whenever required.

SECONDARY DATA

The secondary data is collected from readily available sources such as published journals, magazines, newspapers, books, directories, etc. The secondary data can be collected at a quicker pace as compared to primary data. The cost of collecting secondary data is comparatively low as compared to that of primary data. However, the secondary data is less accurate and less reliable and therefore may affect the quality of decision-making.

The main sources of collecting secondary data include:

(a) Internal Sources such as:

- Purchase and sales records
- Debtors and creditors records
- Employees records relating to performance, compensation service records, etc.
- Financial statements such as cash flow/fund flow statements, balance sheets, profit and loss statements, etc.

(b) External Sources such as:

- Government publications and records.
- Publications of Reserve Bank of India.
- Publications of Directorate General of Commercial Intelligence and Statistics.
- Publications of Directorate of Foreign Trade
- Publications from Export Promotion Councils, Commodity Boards, etc.

The main factors influencing the choice of secondary data include:

- 1. Time Factor** - when the researcher has limited time to collect the data.
- 2. Cost Factor** - when the researcher has limited funds at the disposal to conduct research.

Module – 2

Data Collection and Processing

3. **Objectives of research** - when the data to be collected is most of quantitative in nature for general survey,
4. **Availability of Resources** - when a researcher such as arcades researcher- has limited resources to conduct research.

The main features of secondary data are as follows:

1. **Recorded Data:** Secondary data is collected from recorded and published/unpublished paper source. The data has already been collected by somebody for some purpose other than the project on hand.
2. **Easy to Collect:** The data is readily available and hence easy to collect in the most economical way. The researcher needs to put less efforts to collect secondary data as compared to primary data.
3. **Quantitative Data:** Secondary data is called quantitative data. It can be used for substantiating the primary data.
4. **Comprehensive:** Secondary data is available on every subject and even in those subjects where collection of primary data is difficult. It can supply volumes of large data.
5. **Sources:** Secondary data may be collected from primary source or secondary source. For example, the government conducts primary census survey and its report is called census report. Newspapers and magazines may publish a part of this report which is called secondary source. The researcher may go through the census report which is the primary source.
6. **Relevance:** Secondary data is a readymade information which is meant for general purpose. The researcher must evaluate the secondary data with reference to its accuracy, relevancy and reliability.
7. **Availability:** Secondary data is available internally (from the company's own records, accounts, etc.) or externally (books, magazines, published reports, etc.).
8. **Less Time Consuming and Less Expensive:** Secondary data requires less time to collect the data as compared to primary as it is readily available.

Module – 2

Data Collection and Processing

Secondly, it is less expensive because there is no need to conduct surveys or other methods of primary data collection.

METHODS OF PRIMARY DATA COLLECTION

The primary data collection methods are as follows:

1. Observation Method
2. Experimental Method
3. Interview Method
4. Survey Method

1. Observation Method:

The researcher obtains information of the subjects (persons) under study with the help of observation rather than by way of interviewing. For instance, a researcher studying customer buying behaviour at shopping malls; then he would visit the shopping malls and observe the behaviour of the customers in terms of:

- What they buy/prefer?
- How much they buy?
- How they react to sales promotion schemes offered by the shop owners?
- Whether or not they haggle for the price?
- Whether or not they complain on their previous purchases?

Advantages:

- There is no respondent bias. The respondents are not interviewed. They are only observed, and they may not be aware that they are observed.
- First-hand information about peoples/customers reactions.
- The information collected is reliable and accurate, because the info is collected at the point of action or reaction.

Disadvantages:

- Time consuming, as lot of time is required for observing the actions/reactions of the people under study.
- It is expensive method, as competent and trained staff is required to record the observations of the subjects under study.

Module – 2

Data Collection and Processing

- In depth interviewing is not possible, and therefore, the researcher may not get complete information.
- There may be bias of the researcher, as he may record certain observations as per his own judgement or feelings.

Types:

A. Structured and Unstructured Observation:

Structured Observation is based on the assumption that for the data collection to be most effective every pertinent detail should be worked out in advance. The researcher identifies the elements to be observed well in advance. The observation is done as planned with very little deviation. The time allotted for observation to each respondent and latent information sought is pre-decided. This type of observation is controlled by the researcher.

Unstructured observation is relatively non-planned by the researcher and observes the elements as and when the events take place. It is a more flexible approach for the researcher but needs to be competent to not to lose the direction of the research topic.

B. Disguised and Undisguised Observation:

In case of disguised observation, the researcher acts as an observer as well as a participant or a member of the sample group. However, he will conceal his identity and purpose so that the sample group behaves naturally. The researcher may adopt disguised observation techniques, when he believes that:

- i. The sample group may not give consent for observation.
- ii. The sample group may manipulate their behaviour due to the presence of the observer/ researcher.

In case of undisguised observation, the researcher acts as an observer and is physically present as a part of the phenomenon or sample group with formal consent. The researcher does not conceal his identity and may or may not conceal his identity and may not disclose the research purpose. The researcher/ observer is free to record the behaviour and reactions of the sample group.

Module – 2

Data Collection and Processing

C. Mechanical Observation:

When the researcher uses mechanical devices like cameras, tape recorders, video tape, light meter, stop watch and other electronic devices, then it is called as mechanical observation. Generally, when the physical presence of the observer/ researcher may not be possible, then, in such a situation, mechanical observation technique is adopted. Also, at times, even with the presence of the observer researcher, such mechanical devices may be used to supplement the observer's/ researcher's records and perception.

2. Experimental Method

The experimentation method is mostly used in the case of scientific research study. With the help of experimentation, the research may like to study the cause-effect relationship between two or more variables.

The experimentation method can be also used in development of a new product. For instance, a firm may like to introduce fuel efficient bike or some other product. Therefore, R & D experiments may be required so as to develop the fuel-efficient product. The newly developed product may be subject to trials before it is launched in the market.

An experiment should have two types of validity:

- i. **Internal validity:** When the effects of extraneous variables are eliminated and when the observed effects of dependent variable can be solely attributed to the experimented dependent variable then the experiment is said to have internal validity.
- ii. **External validity:** When the results of the experiments can be generalised and can be extended to other settings and discipline then, the experiment is known to have external validity.

Advantages:

- It provides first-hand information.
- It provides reliable and relevant information.
- The researcher may be in a position to develop new techniques/ methods.

Disadvantages:

Module – 2

Data Collection and Processing

- It may be very expensive as lot of money may be required for experimentation.
- At times, lot of time and effort is required on the part of the researcher.
- The delay in results may generate frustration on the part of the researcher.

Types:

Field Experiments:

Field experiments are conducted in a real-life situation in which the researcher manipulates the independent variable to test the hypothesis. These experiments are quite expensive and are prolonged over a long period of time. They provide a high degree of external validity i.e. the findings of field experiments can be extended to other settings. Field experiments are used to find solutions to practical social problem as like:

- Changing organization culture
- Improving the effectiveness of mass education
- Testing new products in markets, etc.

Laboratory Experiments:

Laboratory experiments are highly structured and conducted in a setting specifically created for the research purpose. The researcher has a complete control on extraneous variables and hence provides high degree of internal validity. However, these experiments may not be generalized to real life situations. Laboratory experiments are useful in physical sciences as well as social sciences.

3. Interview Method:

Interview method of data collection involves face-to-face interaction of the interviewer/ researcher with the respondents. The purpose of interview is to collect as much information as possible from the respondents.

- "Interview is an effective informal non-verbal conversation initiated for specific purpose and focused on certain planned content areas" - P.V. Young
- "Interviewing is fundamentally a process of social interaction ie exchange of thoughts, ideas, views, opinions, and experiences. It is not the words spoken which matter but also the gestures, glances, facial

Module – 2

Data Collection and Processing

expressions, pauses, modulation of voice, rate of speech, etc. which also matter a lot." Goode and Hatt

Advantages:

The advantages of Interview Method are as follows.

- 1. Reliability:** The primary data collected through interview method provides first-hand information to the researcher. Therefore, this type of data is more accurate and reliable. However, the accuracy and reliability of primary data largely depends on the quality of interview method. To get accurate and reliable data, the field interviewers must be properly trained.

- 2. Detailed Information:** Interview method can provide detailed information. The researcher can obtain in-depth information by asking relevant questions. The interviewer may ask additional questions depending upon the responses and information required. Helps in Hypothesis.

- 3. Formulation:** The use of documentary sources helps in formulation of research hypothesis. When an investigator has more than one hypothesis in hand, primary data supports the selection of correct hypothesis.

- 4. Flexibility:** The interview method permits flexibility in collecting the data. The interviewer can restructure or modify the questions depending upon the situation. The interviewer may delete or add certain questions depending upon the situation. The interviewer can also ask questions in a certain language with which the respondent is more familiar with.

- 5. Personal Touch:** Personal interview provides personal touch in collecting the data. The interviewer can develop a good rapport with the respondent. The personal touch can enable the interviewer to obtain frank responses.

- 6. Helps to Overcome Resistance:** Survey/Interview method helps to overcome resistance from the respondents to respond favourably. The interviewer can assure the respondent that his/ her identity may not be revealed, if so desired. The interviewer can use persuasion skills to collect relevant information.

Module – 2

Data Collection and Processing

Disadvantages:

The disadvantages of interview method are as follows:

1. **Time Consuming:** The interview method is time consuming, as lot of time is required to interview the respondents and collect the data.
2. **Expensive:** The interview method is expensive as there is a need to appoint interview. Salary and perks to be paid to the staff. However, in the case of secondary data, there is no need for interview.
3. **Paper Work:** There is lot of paper work involved, because the interview requires questionnaires to be filled in. Also, there is good deal of paper work in respect of processing of data.
4. **Respondent Bias:** The respondents may not give proper responses. They may withhold certain data or may provide fake responses in respect of certain sensitive aspects such as salary/ Income, investments, payment of taxes.
5. **Interviewer Bias:** There is a possibility of interviewer bias. The interviewer may fill the questionnaire on his own, or he may edit the questionnaire as he likes.
6. **Sampling Problems:** It is difficult to select a proper sample especially when the universe is large, and covers a large area

Types:

Personal Interview:

It is a face-to-face interaction between the interviewer and the respondent. The interviewer may ask the questions, and the respondent accordingly responds. There are various types of personal interviews:

- **Formal and Informal** - In formal interview certain formalities are followed such as prior appointment of the interviewee, In an informal interview, the interviewer may not follow formalities in data collection.

Module – 2

Data Collection and Processing

- **Structured and Unstructured** - In structured interview, the interviewer asks certain questions designed well in advance. In unstructured interview, no specific questionnaire is followed. The interviewer is given freedom to ask certain questions depending on the situation. Structured questionnaire is normally not followed.
- **Individual and Group-In individual** - interview, only one person is interviewed at a time, and in a group interview, a group ed respondents may be asked questions at the same time and at the same place.
- **General or Specific Interview** - In general interview, the questions asked may be general in nature. For example, genera comments on advertising media. However, in the case of specific interview, specific questions may be asked such as specific comments on T.V. advertising.
- **Directive and non-directive interview** - In directive interview the interviewer gives directions to the interviewee to response to certain questions, In non-directive interview, the respondents are given freedom to respond to certain topic without much questioning on the part of the interviewer.

Focused Group Interview:

In this type of Interview, the researcher focuses on the experiences of the group of respondents and its possible effects. The researcher interviewer structure the interview well in advance but has the flexibility in asking the sequence of questions and deciding the manner of asking the same. A definite framework of topics is decided in advance. This type of interview helps to develop the hypotheses relating to specific experiences of the respondents. The effectiveness of the conducted focused group interview depends upon:

- Line of thought of the interviewer
- Direction of interview resorted by the interviewer

In-depth Interview:

It is a semi-structured approach wherein details relating to the respondents' opinions, emotions, attitudes, convictions, etc. are sought. It is designed

Module – 2

Data Collection and Processing

intensively to examine the personal feelings and emotions of the respondents. The purpose of the interview is to get a true picture of the respondent through deep probing into his mind. This type of interview requires a mature understanding of human behaviour on the part of the interviewer/ researcher.

4. Survey Method:

The survey can be census survey or sample survey. In case of census Survey, the entire universe is contacted to collect the data For example, a survey on job satisfaction of doctors in Government hospitals in the city of Mumbai, then each and every doctor in Govt hospitals in Mumbai needs to be contacted to collect the data. But because of limitations of time, and money, it may not be possible to contact each and every doctor, and therefore, the researcher may adopt sample survey method. In sample survey method, only a part of the universe is selected to collect the data. But the sample should be so selected that it represents the universe

"Survey research is an organized attempt to analyse, interpret and report the present status of social institution group or areas." -F.L. Whitney

Types:

- 1. Telephone Survey:** It is a method of conducting interview by telephoning the respondents. A series of questions are asked on phone and the answers are recorded. It is a very popular method extensively used in western countries.

- 2. Mail Survey:** It is another method of data collection. A questionnaire is prepared containing a list of questions to solicit information from selected respond respondents. This questionnaire is sent through post or advertised in a newspaper or magazine, explaining the purpose of the questionnaire and a request to complete and return it to the researcher. A reply-paid envelope may also be given to the respondent wherever possible to encourage the respondents to promptly send the answers.

- 3. E-mail Survey:** When the surveys are e-mailed to the potential respondents to get responses, then, it is called as e-mail survey. Not many

Module – 2

Data Collection and Processing

researchers use e-mail survey as a method of data collection. Email survey may be an easy method to collect data. However, it is not advisable to use this technique due to the following reasons:

- Every potential respondent may not be using e-mail or a computer.
- The respondent may not be the true representation of the targeted population.
- The responses may not be reliable.

4. Internet Survey: An internet survey is a questionnaire that the target can fill and respond over the internet. Internet surveys are usually created as web forms with a database to store and retrieve the answers. A statistical software tools are provided for data analysis. The researcher can gather the responses as soon as the respondents submit the web forms. Internet surveys can be incorporated on the website of the company. A pop-up survey can also be created on the company website.

The advantages of internet surveys are:

- It can involve larger sample at a lower cost.
- The information can be easily and quickly collected through internet surveys.
- It is a convenient method of data collection.

Generally, the response rate of internet surveys is not favourable. The researcher may send reminder mails to the potential respondents to respond to the web forms. Also, to increase the response rate, the researcher may offer incentives like cash prize or some other incentives.

5. Social media and Media Listening: Social media listening is also called as social media monitoring and active web listening. It is a process of identifying and evaluating what is being said of a person, company, product and/or brand on internet.

In large organizations, social media listening tools can mine text for specific key words on social networking websites, blogs, discussion forums and other social media. Monitoring software transposes specific keywords in unstructured data into numerical values. These numerical values are linked to structured data in a database so the data can be analysed with traditional data mining techniques.

Module – 2

Data Collection and Processing

The top five free social media listening tools are:

- Hootsuite
- TweetReach
- Klout
- Social Mention
- Twazzup

Social media listening can benefit the organization in the following ways:

- Effective brand management
- Evaluation of marketing campaigns accurately
- Discover market gaps
- Respond to positive and negative posts
- Analysis of frequent conversation topics
- Assessment of competitor's strategies
- Interaction with customers through social media platform.

SCHEDULES:

A schedule is similar to questionnaire. A schedule is generally used in the case of large survey such as population census survey. It contains a list of questions which the respondents need to answer and the research staff records the responses given by the respondents.

Purposes: The main purposes of schedule are:

1. To provide a standardized tool for observation or interview in order to attain objectivity,
2. To act as memory tickler i.e., the schedule keeps the memory of the interviewer/observer refreshed and keeps him reminded of the different aspects that are to be particularly observed, and
3. To facilitate the work of tabulation and analysis.

Types of Schedules:

There are several kinds of schedule:

Module – 2

Data Collection and Processing

1. **Rating Schedule** is a schedule used to obtain opinions, preferences etc, relating to statements in the schedule about the phenomenon/problem under study. The schedule consists of positive and negative statements of opinion on the phenomenon.
2. **Documents Schedule** is used to collect data from recorded evidences and/or case histories. In this case, the respondent has to fill up the blanks based on the records or documents provided to the respondent.
3. **Survey Schedule** is like questionnaire with a set of questions relating to the research problem.
4. **Observation Schedule** is a schedule used when observational method of data collection is used.
5. **Structured or unstructured** interview schedules are used for collecting data when interview method is used for collecting data from the respondents.

Framing of a Schedule:

1. **Study all Aspects of the Problem:** The problem under study should first of all be split up into various aspects. The determination of these aspects will depend upon clear understanding of the problem under study.
2. **Clarity:** Care should be taken to see that the questions convey the exact sense. Respondents will be willing to supply information without any hesitation, bias or distortion of facts, if questions are exact and clear.
3. **Sequencing of Questions:** In order to obtain well-organised information, it is necessary that the questions should be presented to the respondents in a logical sequence. It has been observed that the change in the order of questions affects the answers adversely.
4. **Pre-testing of Schedule:** There is a need to conduct a pre-test of the schedule before it is presented to the respondents. Any vagueness and missing of vital aspects can be located with the help of pre-test of schedule.
5. **Division of Schedule:** The schedule may be divided into adequate number of divisions:
 - Introductory part,
 - Instructional part,
 - Issues related part.
6. **Appropriate Form of Questions:** Use appropriate forms of questions at appropriate places. Open ended, close ended, pictorial, Yes/No (Questions), multiple choice questions, etc. can be used.