

1. Introduction

Welcome to the series of E-learning modules on Primary Data. In this module we are going to cover the importance and methods of Primary Data.

At the end of this session, you will be able to:

- Explain the meaning and importance of Primary Data
- Explain the methods of collecting primary data
- Explain the designing of a questionnaire

Introduction:

Statistics is concerned with the scientific method of collecting, organizing, summarizing, presenting and analyzing the data as well as drawing a valid conclusion and making reasonable decisions on the basis of such analysis.

Statistics is used by government, industry, business, agriculture, mining, transport, education, medicine and so on for their planning, setting standards and functional controls.

Thus, the problems pertaining to this field needs to be well defined for which statistics follows a defined step to arrive at decisions. They are:

- Formulation of the problem
- Objective of the study,
- Determining the source of data
- Designing data collection forms
- Conducting the field survey
- Organising the data
- Analysing the data
- Reaching statistical findings
- Presentation of findings

Our discussion in this session focus on the design of collecting the primary data, its methods and designing questionnaire to collect data from the respondents.

Data collection is a very important step in any statistical investigation, as data collected if inaccurate or inadequate or defective in some aspect will undoubtedly be misleading.

Therefore it is important to plan proper collection of the required data.

The word 'data' has been derived from the Latin word 'datum' means 'something given'. It means a piece of information which can be quantitative or qualitative. It is a plural word and means 'facts and statistics collected for analyses'.

Data's are of two types: On the basis of collection of information we classify it as Primary Data and secondary data and on the basis of representation we classify it as Quantitative data and qualitative data.

Primary data are the data collected first hand by the researcher specifically for a purpose to

facilitate the study.

Secondary data are the data in the hands of an individual gathered for some other purpose but utilized for this specific problem.

Qualitative data are the data expressed in the non – numerical form e.g., poor, satisfaction and so on.

Quantitative data are the data expressed in numerical form – e.g. height, weight, income and so on.

Data that has been collected from first-hand-experience is known as primary data. Primary data has not been published yet and is more reliable, authentic and objective. Primary data has not been changed or altered by human beings; therefore its validity is greater than secondary data.

In statistical surveys it is necessary to get information from primary sources and work on primary data: for example, the statistical records of female population in a country cannot be based on newspaper, magazine and other printed sources. One such source is old and secondly they contain limited information as well as they can be misleading and biased.

2. Importance of Primary Data

Importance of Primary Data are:

- Validity
- Authenticity
- Reliability

Validity: Validity is the quality of a research that makes it trustworthy, scientific, logical and acceptable. Using primary data in research can improve the validity of research as it is the first-hand information obtained from a sample that is representative of the target population.

Authenticity: Authenticity is the genuineness of the research. Authenticity can be at stake if the researcher invests personal biases or uses misleading information in the research. Primary research tools and data can become more authentic if the methods chosen to analyze and interpret data are valid and reasonably suitable for the data type. Primary sources are more authentic because the facts have not been overdone. Primary source can be less authentic if the source hides information or alters facts due to some personal reasons.

Reliability: Reliability is the certainty that the research is enough true to be trusted on. For example, if a research study concludes that junk food consumption does not increase the risk of cancer and heart diseases. This conclusion should have to be drawn from a sample whose size, sampling technique and variability is not questionable. Reliability improves with using primary data. In the similar research mentioned above if the researcher uses experimental method and questionnaires the results will be highly reliable. On the other hand, if he relies on the data available in books and on internet he will collect information that does not represent the real facts.

3. Sources of Primary Data - Part 1

Sources of Primary Data:

Collection of primary data is done during the course of experimentation or surveys.

Experiments: Experiments require an artificial or natural setting in which to perform logical study to collect data. Experiments are more suitable for medicine, psychological studies, nutrition and for other scientific studies. In experiments the experimenter has to keep control over the influence of any extraneous variable on the results.

Surveys: Survey refers to the method of securing information concerning a phenomenon under study. It is most commonly used method in social sciences, management, marketing and psychology to some extent.

The sources for collecting primary data which could be valid reliable and authentic are limited the most commonly used methods are:

- Observation
- Interview
- Questionnaire
- Schedules and
- Others like
 - Warranty card,
 - Distributor audits,
 - Pantry audits,
 - Consumer panel,
 - Mechanical devices,
 - Projective techniques,
 - Depth interview and
 - Content analysis

Observation:

Observation method of collecting data is a scientific tool where the investigator sought information through direct observation of the respondent. For example the investigator may observe the brand of oil a respondent buys and recorded it as information.

The observation on the basis of purpose is classified as structured observation and unstructured observation:

- Structured observation are observation where in the observation is having a specific unit, style, standardized condition and the selection of the pertinent data to be observed
- Unstructured observation are observations which takes place without a specific unit, style, standardized condition and the selection of the pertinent data to be observed though in advance

On the basis of Participation of the observer: Participant observation and non participant observation.

- ✧ Participant observation are observation where the observer observes by making himself more or less a member of the group he is observing so that he can experience what the members of the group experience
- ✧ Non participant observation is an observation where the observer observes as a detached emissary without the attempt to experience through what the members of the group experience

On the basis of the Environment: Controlled and uncontrolled observation.

- ✧ Controlled observation is observation recorded in a pre arranged, planned, environment involving experimental procedures
- ✧ Uncontrolled observations generally are observation recorded in a natural setting

Let us now discuss about the Merits and Demerits of Observations:

Merits are:

- Subjective bias is eliminated
- Information is related to current happenings
- Independent of the respondents willingness to participate or not
- Record the natural behaviour of the group
- Easy verification of the truth of the statement

Demerits are:

- ✧ It is an expensive method
- ✧ Information provided is limited
- ✧ Unforeseen factors can interfere the observation
- ✧ Observer may lose the objectivity of observation
- ✧ Problem of observation control is not solved
- ✧ It narrow downs the investigators range of experience

Interview:

Interview is a face-to-face conversation with the respondent. In interview the main problem arises when the respondent deliberately hides information otherwise it is an in depth source of information. It is a method of collecting primary data through oral – verbal stimuli presentation and reply in terms of oral – verbal responses.

The interviewer can not only record the statements the interviewee speaks but he can observe the body language, expressions and other reactions to the questions too. This enables the interviewer to draw conclusions easily.

Interview can be carried out instantly or through pre-determined questions. If there are structured pre determine questions we call it as structured interview and if there is flexibility in asking the questions we call it as unstructured interview.

This method can be used through personal interview, telephonic interview, focused interview, clinical interview and non-directive interview.

Personal interview is a face to face contact with the person from whom the information needed to be gathered.

Collecting information by contacting the respondents on telephone itself is called telephonic interview. It is used widely in industrial surveys.

Interview can also be focused interview when attention is given on the experience of the respondent.

Clinical interview is related to the feelings or motivations of the individual.

Non-directive interview when the respondent is just encouraged to talk about a topic with bare minimum questions.

Merits of Interview are:

- ⤴ More information and in greater depth can be collected
- ⤴ Can overcome resistance if any by the respondent.
- ⤴ Greater flexibility as questions can be restructured
- ⤴ Personal information can be easily obtained
- ⤴ Samples can be controlled effectively
- ⤴ Record more spontaneous results
- ⤴ Flexibility with language can be adopted which can reach the education level of the respondent as well as avoid misinterpretation.
- ⤴ Can collect supplementary information

Demerits of Interview are:

- ⤴ It is a very expensive method when large and widely spread samples are taken
- ⤴ Bias of the interviewer as well as that of the respondent can be exist
- ⤴ Respondents may not be easily approachable
- ⤴ Relatively time consuming when large and widely spread samples are taken
- ⤴ Presence of the interviewer on the spot may over stimulate the respondent
- ⤴ High cost as it involves selecting, training and supervising the interviewer
- ⤴ Effective response is possible only when there is a proper rapport between the interviewer and the respondent

4. Sources of Primary Data - Part 2

Questionnaire: It is the most commonly used method in survey. Questionnaires are a list of questions which either has open-ended or close -ended questions for which the respondent give answers. Questionnaire can be conducted via telephone, mail, live in a public area, or in an institute, through electronic mail or through fax and other methods.

Classification:

- ⤴ Structured –non-disguised
- ⤴ Structured disguised
- ⤴ Non-structured non-disguised
- ⤴ Non- structure disguised

Structured questionnaires are those in which the questions are listed in a formal prearranged order and asked by the investigator in the same order.

Non structured questionnaires are those in which the questions are not structured and is left to the investigators discretion to ask the respondent.

Disguised questionnaires are those in which the objective of the questionnaire is not revealed to the respondent.

Non Disguised questionnaires are those in which the objective of the questionnaires are revealed to the respondent.

Designing a Questionnaire

A collection of data in the primary data format should be extremely careful in deciding what information is to be collected, how many questions to be formulated, sequence of the questions, its wordings, layout of the questionnaire.

A good questionnaire should also consider the time and effort to be spent and the objective to be achieved.

The designing of an effective questionnaire in a systematic manner is possible by keeping the following points in mind:

- Types of information to be collected
- Types of questions
- Phrasing of the questions
- Order of the questions
- Number of questions
- Layout of the questions

Types of information to be collected:

The specific information to be collected needs to be pre -determined so that questions can be framed in such a way that it will be able to help us in collecting facts that are relevant, vital, and accurate and manages cost. If there are questions that are irrelevant then the data collected can be inaccurate, raise in total cost and time consumption.

Types of Questions:

This the second important point to be considered. We need to understand that the questions are classified into three types: Open ended question, dichotomous question and multiple choice questions.

Types of Questions-Open Ended Question:

Open ended questions are questions where the respondent can feel free to answer. This type of questions are best suitable when the investigator needs to understand the top most priority in the minds of the respondent also the answer can be of any length and contain a detail answer. However there are certain limitations while using open ended questions – it is difficult to note down all the points, if they are more than two interviewers then there is likely chance of element of bias as each interviewer would record the information based on their understanding.

For example: In case you would like to give any additional information about the product please do so in the space given below.

Types of Questions-A Dichotomous Question:

A dichotomous question are question which have only two answers in the form of 'yes' or 'no', 'true' or 'false' 'use or 'don't use', etc. These types of questions are most convenient to indicate their choice, least bothersome, minimum time, easy to edit, tabulate and interpret. The limitations are it would not be able to get information that needs to be probe like attitude, actions or feelings, fixed alternative answers, forcing a statement of opinion.

For example: Do you currently use this product?

Yes ☐ No ☐

Types of Questions-A Multiple Choice Question:

Multiple choice questions are questions having more than two choices. In this type of question all possible choices can be exhausted and the respondent can choose the appropriate answer. The advantage is that such question can give a list of alternatives helping in getting a clarity of the question, easy handling and inexpensive. The limitations are more time is required for editing, tabulation, and interpretation.

For example: If you have already used the product how would you rate it?

Extremely good ☐ Good ☐ Satisfactory ☐ Bad ☐

Phrasing of the Question:

Here we focus on preparing a questionnaire attention must be paid to the wordings of the questions as reliable and meaningful information depends on it to a large extent. The word used in the questionnaire should be simple and familiar to all respondents. While preparing the questionnaire care should be taken to avoid words that are ambiguous, danger words, words with emotional connotations, use of phrases that reflect the prestige of the respondent and question words that bias the answers. In fact usage of questioning words and formulation is an art that is learnt out of practice.

Order of the Question:

In order to make the questionnaire effective and to ensure quality to the information collected care should be taken to sequence the questions such that:

- It reduces the chances of questions being misunderstood.
- Relation of one question to the other should be apparent.
- Easy answering questions in the beginning as it will influence the attitude of the respondent.
- Questions should be framed in such a way that it arouses human interest.
- Should conform to the respondent's way of thinking.
- Help in building rapport with the respondent.
- The question should be in the form that it moves from a general to a more specific sequence.
- The question sequence should be in such a way that it helps in understanding that the later question is a function of the previous question also.

Number of Questions to be Asked:

It should be kept in mind that the length of the questionnaire should be decided in such a way that the interest of the respondent is sustained till the last moment to obtain the required information.

5. Sources of Primary Data - Part 3

Layout of the Questionnaire:

To create a favourable impression in the mind of the respondent care should be taken to see that the questionnaire is:

- ⤴ Neatly printed
- ⤴ An individual page does not have too many question
- ⤴ Proper spacing between questions
- ⤴ Bold and underline the important wordings
- ⤴ If lengthy questionnaire then reduce the size by providing two columns in a page
- ⤴ Using finer type and easy looks
- ⤴ It is better to print the questionnaire in a superior quality paper so that writing by pen or pencil is smooth.

Essentials of a good questionnaire:

We can summarise the following points for preparing an effective questionnaire:

- ⤴ Size of the questionnaire should be short and simple
- ⤴ Logical sequence – moving from easy to more difficult
- ⤴ Question could be open ended, dichotomous or multiple choice
- ⤴ Include control questions to indicate the reliability of the respondent
- ⤴ Adequate space for answers to be provided.
- ⤴ Space for indicating uncertainties 'like do not know' or 'no preference.'
- ⤴ Brief instructions with regard to filling up the questionnaire.
- ⤴ Attractive looking questionnaire in a quality paper to attract the attention of the respondents.

Things to be avoided while preparing a questionnaire:

- ⤴ Avoid usage of technical terms and vague expressions
- ⤴ Avoid usage of difficult questions as its analyses is difficult
- ⤴ Avoid questions affecting the sentiments of the respondent

Merits of Questionnaire are:

- ⤴ Low cost when the universe is large and is widely spread geographically
- ⤴ Free from the bias of the interviewer
- ⤴ Adequate time to give well thought answers
- ⤴ Respondent easily not approachable can also be reached conveniently
- ⤴ Usage of large samples resulting in more dependable and reliable data

Demerits of Questionnaire are:

- ⤴ Low rate of return of filled questionnaire due to no response
- ⤴ Can be used only for educated and cooperating respondents
- ⤴ Control over questionnaire can be lost
- ⤴ Inbuilt inflexibility as once the questionnaires are dispatched
- ⤴ Possibility of ambiguous replies, omissions and misinterpretations of questions.
- ⤴ Slowest method and
- ⤴ Difficulty in finding out whether the respondents are true representatives.

Schedule:

Schedule is a set of questions which is generally filled by an enumerator who is specially appointed for this purpose. The enumerator along with the schedules go to respondents put the questions from the schedule and records the replies in the space meant for the same in the schedule. In certain cases the schedule is handed to the respondent and the enumerator helps in recording their answers.

Others

Others includes: Warranty card, Distributor Audit, Pantry Audit, Consumer panel, Mechanical devices, Projective techniques, Depth interview.

Warranty card: These are postal sized cards which have questions in the printed form placed inside the package along with the product with a request to fill in the card and post it back to the dealer.

Distributor Audit: This is a method where the retail stores are audited by the salesman and use such information to estimate the market size, market share, seasonal purchasing pattern and so on.

Pantry Audit: This is a technique where information is collected by examining the consumer's pantry and data is used to understand the consumers buying behavior.

Consumer panel: This is a sample of consumers who are interviewed repeatedly over a period of time and information is collected.

Mechanical devices: This is a method in which devices like the eye camera, pupil metric camera, and psycho galvanometer are used for the purpose of collecting the required information.

Projective techniques: These are the techniques developed by the psychologists to collect information from the respondent unconsciously to project his own attitudes or feelings on the subject under the study. It plays an important role in motivational researches or in attitude surveys. The important projective techniques are word association tests, sentence completion tests, story completion tests, verbal projections tests, play techniques, quizzes tests and examination.

Depth interview: These are interviews held to explore the needs, desire and feelings of the respondents. It requires a great skill on the part of the interviewer and involves considerable time.

Here's a summary of our learning in this session:

- Meaning and importance of Primary Data
- Methods of collecting primary data
 - Observation
 - Interview
 - Questionnaire
 - Schedules and
 - Others

- Designing of a questionnaire
 - Types of information to be collected
 - Types of questions
 - Phrasing of the questions
 - Order of the questions
 - Number of questions
 - Layout of the questions