Master of Arts

(Psychology)

Research Methodology (1)

Semester-I

Author- Ashok Kumar

SURESH GYAN VIHAR UNIVERSITY

Centre for Distance and Online Education
Mahal, Jagatpura, Jaipur-302025

Published by:

S. B. Prakashan Pvt. Ltd.

WZ-6, Lajwanti Garden, New Delhi: 110046

Tel.: (011) 28520627 | Ph.: 9625993408

Email: info@sbprakashan.com | Web.: www.sbprakashan.com

Designed & Graphic by : S. B. Prakashan Pvt. Ltd.

Printed at :

INDEX

Research Methodology

Learning out comes

Students will be able to understand:

Unit-1

- Recognize the significance of research in academic, professional, and practical contexts.
- Understand the systematic research process, from defining the problem to drawing conclusions.
- Identify and evaluate the criteria that contribute to good research practices.

Unit-2

- Understand the process of collecting secondary data and its utilization in research.
- Define and comprehend the case study method and its application in detailed investigations.
- Explain the survey method and its use in gathering information from a sample population.

Unit-3

- Define and explain non-probability sampling methods.
- Define and explain probability sampling methods.
- Differentiate between non-probability and probability sampling designs.

Unit-4

- Understand the process of organizing data through frequency distribution and graphs.
- Calculate and interpret descriptive statistics, including measures of central tendency and variation.
- Identify and analyze different types of distributions.

Unit-5

- Understand the technique of interpretation in the context of report writing.
- Apply precautions in the interpretation phase to ensure the accuracy and reliability of findings.
- Recognize the significance of report writing in various contexts.

RESEARCH METHODOLOGY
SYLLABUS

UNIT I

Introduction to Research Methodology

Introduction to Research Methodology: Meaning of Research – Objectives of Research – Types of Research – Significance of Research – Research Process – Criteria of Good Research.

UNIT II

Methods of Data Collection

Methods of Data Collection & Analysis of Data: Collection of Primary data: Observation method, Interview method, Questionnaires, schedules, other methods - Collection of Secondary Data – Case study Method – Survey Method.

UNIT III

Sampling Fundamentals

Sampling Fundamentals: Need for sampling – Steps in sample design – Types of sample Designs: Non-probability sampling – Probability sampling – Complex Random Sample Designs: Systematic sampling, Stratified sampling, Cluster sampling, Area sampling, Multistage sampling, Sampling with probability proportional to size, Sequential sampling.

UNIT IV

Measurement and Scaling Techniques

Measurement and Scaling Techniques: Measurement in Research – Measurement scales – Sources of error in measurement – Tests of sound measurement – Technique of developing Measurement tools; Organizing data: Frequency distribution – Graphs – Descriptive statistics: Measures of central tendency – Measures of variation – Types of distributions. Inferential statistics: z test – t test – Analysis of Variance – Correlation– Concepts related to correlation

UNIT V

Report Writing

Report Writing: Technique of Interpretation – Precautions – Significance of Report writing – Steps – Types of Report writing – Mechanics of writing a report.

UNIT I

INTRODUCTION TO RESEARCH METHODOLOGY

STRUCTURE

NOTES 

1.1 LEARNING OBJECTIVE

After learning this unit students will be able to:

- Understand the Research Methodology.
- Understand the Objectives of Research.
- Understand the Types of Research.
- Understand the Research Methods Versus Methodology.
- Understand the Significance of Research Methodologies.
- Understand the Research Process.

1.2 INTRODUCTION TO RESEARCH METHODOLOGY

Meaning of Research

Research may be very broadly defined as systematic gathering of data and information and its analysis for advancement of knowledge in any subject. Research attempts to find answer intellectual and practical questions through application of systematic methods. Webster's Collegiate Dictionary defines research as "studious inquiry or examination; esp.: investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts, or practical application of such new or revised theories or laws". Some people consider research as a movement, a movement from the known to the unknown. It is actually a voyage of discovery. We all possess the vital instinct of inquisitiveness for, when the unknown confronts us, we wonder and our inquisitiveness makes us probe and attain full and fuller understanding of the unknown. This inquisitiveness is the mother of all knowledge and the method, which man employs for obtaining the knowledge of whatever the unknown, can be termed as research. Research is an academic activity and as such the term should be used in a technical sense. According to Clifford Woody research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organizing and evaluating data; making deductions and reaching conclusions; and at last, carefully testing the conclusions to determine whether they fit the formulating hypothesis.

D. Steiner and M. Stephenson in the Encyclopaedia of Social Sciences define research as "the manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art." Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the pursuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research. The systematic approach concerning generalization and the formulation of a theory is also research. As such the term 'research' refers to the systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analyzing the facts and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalizations for some theoretical formulation.

NOTES 

1.3 OBJECTIVES OF RESEARCH

The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has not been discovered as yet. Though each research study has its own specific purpose, we may think of research objectives as falling into a number of following broad groupings:

1. To gain familiarity with a phenomenon or to achieve new insights into it (studies with this object in view are termed as exploratory or formulative research studies);
2. To portray accurately the characteristics of a particular individual, situation or a group (studies with this object in view are known as descriptive research studies);
3. To determine the frequency with which something occurs or with which it is associated with something else (studies with this object in view are known as diagnostic research studies);
4. To test a hypothesis of a causal relationship between variables (such studies are known as hypothesis-testing research studies).

1.4 TYPES OF RESEARCH

Types of research can be classified in many different ways. some major ways of classifying research include the following.

1. Descriptive versus Analytical Research
2. Applied versus Fundamental Research
3. Qualitative versus Quantitative Research
4. Conceptual versus Empirical Research

Descriptive research concentrates on finding facts to ascertain the nature of something as it exists. In contrast analytical research is concerned with determining validity of hypothesis based on analysis of facts collected.

Applied research is carried out to find answers to practical problems to be solved and as an aid in decision making in different areas including product design, process design, and policy making. Fundamental research is carried out as more to satisfy intellectual curiosity, than with the intention of using the research findings for any immediate practical application.

Quantitative research studies such aspects of the research subject which are not quantifiable, and hence not subject to measurement and quantitative analysis. In contrast, quantitative research makes substantial use of measurements and quantitative analysis techniques.

Conceptual research involves investigation of thoughts and ideas and developing new ideas or interpreting the old ones based on logical reasoning. In contrast, empirical research is based on firm verifiable data collected by either observation of facts under natural condition or obtained through experimentation.

NOTES 

Some Other Types of Research: All other types of research are variations of one or more of the above stated approaches, based on either the purpose of research, or the time required to accomplish research, on the environment in which research is done, or on the basis of some other similar factor. Form the point of view of time, we can think of research either as one-time research or longitudinal research.

In the former case, the research is confined to a single time period, whereas in the latter case the research is carried on over several time periods. Research can be field-setting research or laboratory research or simulation research, depending upon the environment in which it is to be carried out. Research can as well be understood as clinical or diagnostic research. Such research follows case-study methods or in-depth approaches to reach the basic causal relations. Such studies usually go deep into the causes of things or events that interest us, using very small samples and very deep probing data-gathering devices.

The research may be exploratory or it may be formalized. The objective of exploratory research is the development of hypotheses rather than their testing, whereas formalized research studies are those with substantial structure and with specific hypotheses to be tested. Historical research is that which utilizes historical sources like documents, remains, etc. to study events or ideas of the past, including the philosophy of persons and groups at any remote point in time. Research can also be classified as conclusion-oriented and decision-oriented.

While doing conclusion, a researcher is free to pick up a problem, redesign the inquiry as he proceeds, and is prepared to conceptualize as he wishes. Decision-oriented research is always for the need of a decision-maker and the researcher in this case is not free to embark upon research according to his own inclination. Operations research is an example of decision-oriented research since it is a scientific method of providing executive departments with a quantitative basis for decisions regarding operations under their control.

Research Approaches

The above description of the types of research brings to light the fact that there are two basic approaches to research, viz., the quantitative approach and the qualitative approach. The former involves the generation of data in a quantitative form which can be subjected to rigorous quantitative analysis in a formal and rigid fashion. This approach can be further sub-classified into inferential, experimental, and simulation approaches to research. The purpose of the inferential approach to research is to form a database from which to infer characteristics or relationships of the population.

This usually means survey research where a sample of the population is studied (questioned or observed) to determine its characteristics, and it is then inferred that the population has the same characteristics. The experimental approach is characterized by much greater control over the research environment and in this case, some variables are manipulated to observe their effect on other variables. The simulation approach involves the construction of an artificial environment within which relevant information and data can be generated. This permits observation of the dynamic behavior of a system (or its sub-system) under controlled conditions.

NOTES 

The term 'simulation' in the context of business and social sciences applications refers to "the operation of a numerical model that represents the structure of the dynamic process. Given the values of initial conditions, parameters, and exogenous variables, simulation is run to represent the behavior of the process over time."5 Simulation approach can also be useful in building models for understanding future conditions.

A qualitative approach to research is concerned with the subjective assessment of attitudes, opinions, and behavior. Research in such a situation is a function of the researcher's insights and impressions. Such an approach to research generates results either in non-quantitative form or in the form which are not subjected to rigorous quantitative analysis. Generally, the techniques of focus group interviews, projective techniques, and depth interviews are used.

1.5 RESEARCH METHODS VERSUS METHODOLOGY

It seems appropriate at this juncture to explain the difference between research methods and research methodology. Research methods may be understood as all those methods/ techniques that are used for the conduction of research. Research methods or techniques*, thus, refer to the methods the researchers use in performing research operations. In other words, all those methods which are used by the researcher during the course of studying his research problem are termed research methods. Since the object of research, particularly applied research, it to arrive at a solution for a given problem, the available data and the unknown aspects of the problem have to be related to each other to make a solution possible. Keeping this in view, research methods can be put into the following three groups:

1. In the first group we include those methods which are concerned with the collection of data. These methods will be used where the data already available are not sufficient to arrive at the required solution;
2. The second group consists of those statistical techniques which are used for establishing relationships between the data and the unknowns;
3. The third group consists of those methods which are used to evaluate the accuracy of the results obtained.

Research methods falling in the above-stated last two groups are generally taken as the analytical tools of research. Research methodology is a way to systematically solve the research problem. It may be understood as a science of studying how research is done scientifically. In it we study the various steps that are generally adopted by a researcher in studying his research problem along with the logic behind them. It is necessary for the researcher to know not only the research methods/techniques but also the methodology. Researchers not only need to know how to develop certain indices or tests, how to calculate the mean, the mode, the median or the standard deviation or chi-square, and how to apply particular research techniques, but they also need to know which of these methods or techniques, are relevant and which are not, and what would they mean and indicate and why.

Researchers also need to understand the assumptions underlying various techniques and they need to know the criteria by which they can decide that certain techniques and

NOTES 

procedures will be applicable to certain problems and others will not. All this means that it is necessary for the researcher to design his methodology for his problem as the same may differ from problem to problem. For example, an architect, who designs a building, has to consciously evaluate the basis of his decisions, i.e., he has to evaluate why and on what basis he selects a particular size, number, and location of doors, windows, and ventilators, uses particular materials and not others and the like. Similarly, in research the scientist has to expose the research decisions to evaluation before they are implemented. He has to specify very clearly and precisely what decisions he selects and why he selects them so that they can be evaluated by others also. From what has been stated above, we can say that research methodology has many dimensions and research methods do constitute a part of the research methodology.

The scope of research methodology is wider than that of research methods. Thus, when we talk of research methodology, we not only talk of the research methods but also consider the logic behind the methods we use in the context of our research study and explain why we are using a particular method or technique and why we are not using others so that research results are capable of being evaluated either by the researcher himself or by others.

Why a research study has been undertaken, how the research problem has been defined, in what way and why the hypothesis has been formulated, what data have been collected and what particular method has been adopted, and why particular technique of analyzing data has been used and a host of similar other questions are usually answered when we talk of research methodology concerning a research problem or study.

1.6 SIGNIFICANCE OF RESEARCH METHODOLOGIES

Research methodology is necessary in sociology as it provides legitimacy and wider acceptance to sociological work. Different methods are used for different purposes of research like Positivist research or Interpretivist research. Research methods used in sociology broadly belong to three categories.

- Quantitative or qualitative methods based on nature of the data and information.
- Micro methods or macro methods based on the scope of the research.
- Methods based on primary data or secondary data.

As per the research requirement, different methods are used. Survey research can be used in large scale research involving large numbers of respondents and investigators. It is the most common example of a macro method that can be used in such a circumstance. On the other hand, in the study of a particular group A micro method like participant observation can be used. A research design follows certain basic steps like choice of topics, collection of facts, representation of facts, hyper hypothesis making testing and validation and so on further research methodology or research design has following basic types. Exploratory research design is used in the initial stages of the research to acquire some preliminary information. The main objective is to find tune the broad problem into a specific problem statement and generate possible hypothesis. It is used to carry initial research to narrow down on the possible alternatives. The methods used in explore exploratory research are

survey of existing literature survey of experienced individuals' analysis of selected case studies.

Descriptive research design provides accurate description of variables relevant to the problem under consideration and generally used for preliminary and exploratory studies. The descriptive study is more formal and less flexible as it involves both qualitative and quantitative information and can be used for both positivist test and non-positivist test research. The commonly used techniques under this category are panel research design or longitudinal research. The panel design involves the continual or periodic information collection from a fixed panel or sample of respondents. The longitudinal analysis involves repeated measurement of the same variables to facilitate a variety of inferences to be drawn about the behaviour of the elements of the panel. Participant observation and field studies use such methods. Cross sectional design is aimed are taking a one-time stock of the situation or the phenomena in which the decision maker is interested. Cross sectional designs give the picture of situation at a given point of time. Opinion polls and market service use such kind of methods.

Experimental Research Design is where the researcher actively tries to change inputs like the situation, circumstances or experience of participants which may lead to a change in behaviour or outcome for the participants of the study. It establishes the causality between dependent and independent variable and test hypothesis. The participants are ideally randomly assigned to different conditions and variable of interest for measured. This is done to eliminate all extraneous variables. Hawthorne studies of Elton Mayo are class examples of such experimental design. This is a method which is most often associated with natural sciences in which we change variables in a controlled environment. This method is mostly used for positivity is to research or quantitative research as it aims to keep prejudices and biases away while doing the research. Experimental research attempts to determine how and why something happens. Experimental research tests the way in which independent variable affect the dependent variable. Due to high objectivity data obtained through such methodology are more reliable.

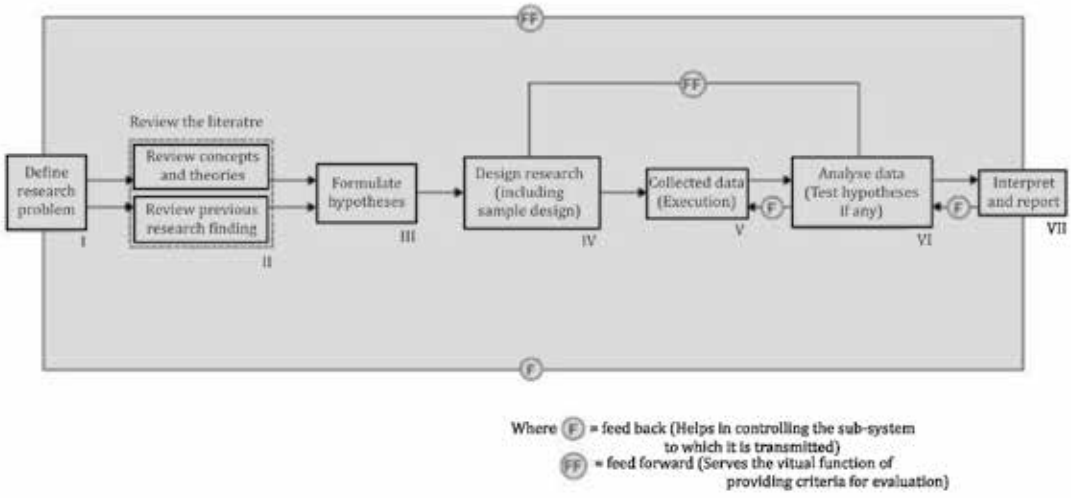
Comparative method is used to compare the social phenomena to arrive at generalized conclusions. It is a method which is suggested as an alternative to the experimental research in sociology but is based on the similar sets of principles. According to Haralambos and Holborn in the Sociology: Themes and Perspectives, 2013, the comparative method is based on what has happened or is happening in society rather than upon situation artificially created by the researcher. It was more popular with the early sociologist. Durkheim was the first sociologist to discuss this method at length in the rules of Sociological Method, 1895. He regarded it as a method of sociology to identify dependent and independent variables. The crime study of suicide is a classic example of use of this methodology. Ginsberg use this method in the study of primitive societies. If a particular social phenomenon is studied in different social contexts and the causes are found out then a cause-and-effect relationship can be established.

1.7 RESEARCH PROCESS

Before embarking on the details of research methodology and techniques, it seems

NOTES

appropriate to present a brief overview of the research process. Research process consists of series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps. The chart shown in Figure well illustrates a research process. The chart indicates that the research process consists of a number of closely related activities, as shown through I to VII. But such activities overlap continuously rather than following a strictly prescribed sequence.



At times, the first step determines the nature of the last step to be undertaken. If subsequent procedures have not been taken into account in the early stages, serious difficulties may arise which may even prevent the completion of the study. One should remember that the various steps involved in a research process are not mutually exclusive; nor are they separate and distinct. They do not necessarily follow each other in any specific order and the researcher has to be constantly anticipating at each step in the research process the requirements of the subsequent steps. However, the following order concerning various steps provides a useful procedural guideline regarding the research process:

1. Formulating the Research Problem 2. Extensive Literature Survey
3. Developing the Hypothesis 4. Preparing the Research Design
5. Determining Sample Design 6. Collecting the Data
7. Execution of the Project 8. Analysis of Data
9. Hypothesis Testing 10. Generalizations and Interpretation
11. Preparation of the report or presentation of the results

formal write-up of conclusions reached. A brief description of the above-stated steps will be helpful.

1. **Formulating the research problem:** There are two types of research problems, vi., those which relate to states of nature and those which relate to relationships between variables. At thievery outset the researcher must single out the problem he wants to study, i.e., he must decide the general area of interest or aspect of a subject-matter that he would like to inquire into. Initially the problem may be stated in a broad general way and then the ambiguities, if any, relating to the problem be resolved. Then, the feasibility of a particular solution has to be considered before

a working formulation of the problem can be set up. The formulation of a general topic into a specific research problem, thus, constitutes the first step in a scientific enquiry. Essentially two steps are involved in formulating the research problem, vi., understanding the problem thoroughly, and rephrasing the same into meaningful terms from an analytical point of view. The best way of understanding the problem is to discuss it with one's own colleagues or with those having some expertise in the matter. In an academic institution the researcher can seek the help from a guide who is usually an experienced man and has several research problems in mind. often, the guide puts forth the problem in general terms and it is up to the researcher to narrow it down and phrase the problem in operational terms. In private business units or in governmental, the problem is usually earmarked by the administrative agencies with whom the researcher can discuss as to how the problem originally came about and what considerations are involved in its possible solutions. The researcher must at the same time examine all available literature to get himself acquainted with the selected problem. He may review two types of literature—the conceptual literature concerning the concepts and theories, and the empirical literature consisting of studies made earlier which are similar to the one proposed. The basic outcome of this review will be the knowledge as to what data and other materials are available for operational purposes which will enable the researcher to specify his own research problem in a meaningful context. After this the researcher rephrases the problem into analytical or operational terms i.e., to put the problem in as specific terms as possible.

This task of formulating, or defining, a research problem is a step of greatest importance in the entire research process. The problem to be investigated must be defined unambiguously for that will help discriminating relevant data from irrelevant ones. Care must, however, be taken to verify the objectivity and validity of the background facts concerning the problem. Professor W. A. Mismanage correctly states that the statement of the objective is of basic importance because it determines the data which are to be collected, the characteristics of the data which are relevant, relations which are to be explored, the choice of techniques to be used in these explorations and the form of the final report. If there are certain pertinent terms, the same should be clearly defined along with the task of formulating the problem. In fact, formulation of the problem often follows a sequential pattern where a number of formulations are set up, each formulation more specific than the preceding one, each one phrased in more analytical terms, and each more realistic in terms of the available data and resources.

2. **Extensive literature survey:** Once the problem is formulated, a brief summary of it should be written down. It is compulsory for a research worker writing a thesis for a Ph.D. degree to write synopsis of the topic and submit it to the necessary Committee or the Research Board for approval. At this juncture the researcher should undertake extensive literature survey connected with the problem. For this purpose, the abstracting and indexing journals and published or unpublished bibliographies are the first place to go to. Academic journals, conference proceedings, government reports, books etc., must be tapped depending on the nature of the problem. In this process, it should be remembered that one source will lead to another. The earlier studies, if any, which are similar to the study in and should be carefully studied. A good library will be a great help to the researcher at this stage.

NOTES 

3. **Development of working hypotheses:** After extensive literature survey, researcher should state in clear terms the working hypothesis or hypotheses. Working hypothesis is tentative assumption made in order to draw out and test its logical or empirical consequences. As such the manner in which research hypotheses are developed is particularly important since they provide the focal point for research. They also affect the manner in which tests must be conducted in the analysis of data and indirectly the quality of data which is required for the analysis. In most types of research, the development of working hypothesis plays an important role. Hypothesis should be very specific and limited to the piece of research in hand because it has to be tested. The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track. It sharpens his thinking and focuses attention on the more important facets of the problem. It also indicates the type of data required and the type of methods of data analysis to be used. How does one go about developing working hypotheses? The answer is by using the following approach:

   a. Discussions with colleagues and experts about the problem, its origin and the objectives in seeking a solution;

   b. Examination of data and records, if available, concerning the problem for possible trends, peculiarities and other clues;

   c. Review of similar studies in the area or of the studies on similar problems; and

   d. Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals with a view to secure greater insight into the practical aspects of the problem.

   Thus, working hypotheses arise as a result of a-priori thinking about the subject, examination of the available data and material including related studies and the counsel of experts and interested parties. Working hypotheses are more useful when stated in precise and clearly defined terms. It may as well be remembered that occasionally we may encounter a problem where we do not need working hypotheses, especially in the case of exploratory or formularies researches which do not aim at testing the hypothesis. But as a general rule, specification of working hypotheses in another basic step of the research process in most research problems.

4. **Preparing the research design:** The research problem having been formulated in clear cut terms, the researcher will be required to prepare a research design, i.e., he will have to state the conceptual structure within which research would be conducted. The preparation of such a design facilitates research to be as efficient as possible yielding maximal information. In other words, the function of research design is to provide for the collection of relevant evidence with minimal expenditure of effort, time and money. But how all these can be achieved depends mainly on the research purpose. Research purposes may be grouped into four categories,

   i. Exploration,

   ii. Description,

   iii. Diagnosis, and

   iv. Experimentation.

A flexible research design which provides opportunity for considering many different aspects of a problem is considered appropriate if the purpose of the research study is that of exploration. But when the purpose happens to be an accurate description sofa situation or of an association between variables, the suitable design will be one that minimizes bias and maximizes the reliability of the data collected and analyses. There are several research designs, such as, experimental and non-experimental hypothesis testing. Experimental designs can be either informal designs (such as before-and-after without control-after-only with control, before-and-after with control) or formal designs (such as completely randomized design, randomized block design, Latin square design, simple and complex factorial designs), out of which the researcher must select one for his own project. The preparation of the research design, appropriate for a particular research problem, involves usually the consideration of the following:

i. The means of obtaining the information

ii. The availability and skills of the researcher and his staff (if any)

iii. Explanation of the way in which selected means of obtaining information will be organized and the reasoning leading to the selection

iv. The time available for research

v. The cost factor relating to research, i.e., the finance available for the purpose.

5. **Determining sample design:** All the items under consideration in any field of inquiry constitute 'universe' or 'population'. A complete enumeration of all the items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry when all the items are covered no element of chance is left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observations increases. Moreover, there is no way of checking the element of bias or its extent except through is survey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Not only this, census inquiry is not possible in practice under many circumstances. For instance, blood testing is done only on sample basis. Hence, quite often we select only a few items from the universe for our study purposes. The items so selected constitute what is technically called sample. The researcher must decide the way of selecting a sample or what is popularly known as the sample design. In other words, a sample design is a definite plan determined before any data are actually collected for obtaining a sample from a given population. Thus, the plan to select 12 of a city's 200 drugstores in a certain way constitutes a sample design. Samples can be either probability samples or nonprobability samples. With probability samples each element has a known probability of being included in the sample but the non-probability samples do not allow the researcher to determine this probability. Probability samples are those based on simple random sampling, systematic sampling, stratified sampling, cluster/area sampling whereas non-probability samples are those based on convenience sampling, judgment sampling and quota sampling techniques. A brief mention of the important sample designs is as follows:

   a. **Deliberate sampling:** Deliberate sampling is also known as purposive or nonprobability sampling. This sampling method involves purposive or deliberate selection of particular units of the universe for constituting

NOTES 

a sample which represents the universe. When population elements are selected for inclusion in the sample based on the ease of access, it can be called convenience sampling. If a researcher wishes to secure data from, say, gasoline buyers, he may select a fixed number of petrol stations and may conduct interviews at these stations. This would be an example of convenience sample of gasoline buyers. At times such a procedure may give very biased results particularly when the population is not homogeneous. On the other hand, in judgment sampling the researcher's judgment used for selecting items which he considers as representative of the population. For example, a judgment sample of college students might be taken to secure reactions to a new method of teaching. Judgment sampling is used quite frequently in qualitative research where the desire happens to be to develop hypotheses rather than to generalize to larger populations.

b. **Simple random sampling:** This type of sampling is also known as chance sampling or probability sampling where each and every item in the population has an equal chance of inclusion in the sample and each one of the possible samples, in case of finite universe, has the same probability of being selected. For example, if we have to select a sample of 300items from a universe of 15,000 items, then we can put the names or numbers of all the15,000 items on slips of paper and conduct a lottery. Using the random number tables is another method of random sampling. To select the sample, each item is assigned a number from 1 to 15,000. Then, 300 five digits random numbers are selected from the table. To do this we select some random starting point and then a systematic pattern is used in proceeding through the table. We might start in the 4th row, the second column, and proceed down the column to the bottom of the table, and then move to the top of the next column to the right. When a number exceeds the limit of the numbers in the frame, in our case over 15,000, it is simply passed over and the next number selected does fall within the relevant range. Since the numbers were placed in the table in a completely random fashion, the resulting sample is random. This procedure gives each item an equal probability of being selected. In the case of an infinite population, the selection of each item in a random sample is controlled by the same probability, and that successive selections are independent of one another.

c. **Systematic sampling:** In some instances, the most practical way of sampling is to select every 15th name on a list, every 10th house on one side of a street and so on. Sampling of this type is known as systematic sampling. An element of randomness is usually introduced into this kind of sampling by using random numbers to pick up the unit with which to start. This procedure is useful when sampling frame is available in the form of a list. In such design the selection process starts by picking some random point in the list and then every nth element is selected until the desired number is secured.

d. **Stratified sampling:** If the population from which a sample is to be drawn does not constitute a homogeneous group, then the stratified sampling technique is applied so as to obtain a representative sample. In this technique, the population is stratified into a number of non-overlapping subpopulations or strata, and sample items are selected from each stratum. If the items selected from each stratum are based on simple random sampling the entire

procedure, first stratification, and then simple random sampling, is known as stratified random sampling.

e. Quota sampling: In stratified sampling, the cost of taking random samples from individual strata is often so expensive that interviewers are simply given a quota to be filled from different strata, the actual selection of items for the sample being left to the interviewer's judgment. This is called quota sampling. The size of the quota for each stratum is generally proportionate to the size of that stratum in the population. Quota sampling is thus an important form of non-probability sampling. Quota samples generally happen to be judgment samples rather than random samples.

f. Cluster sampling and area sampling: Cluster sampling involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample. Suppose some departmental store wishes to sample its credit card holders. It has issued its cards to 15,000 customers. The sample size is to be kept say 450. For cluster sampling this list of 15,000 card holders could be formed into 100 clusters of 150 cardholders each. Three clusters might then be selected for the sample randomly. The sample size must often be larger than the simple random sample to ensure the same level of accuracy because is cluster sampling procedural potential for order bias and other sources of error is usually accentuated. The clustering approach can, however, make the sampling procedure relatively easier and increase the efficiency of field work, specially in the case of personal interviews. Area sampling is quite close to cluster sampling and is often talked about when the total geographical area of interest happens to be big one. Under area sampling we first divide the total area into a number of smaller non-overlapping areas, generally called geographical clusters, then a number of these smaller areas are randomly selected, and all units in these small areas are included in the sample. Area sampling is specially helpful where we do not have the list of the population concerned. It also makes the field interviewing more efficient since interviewer can do many interviews at each location.

g. Multi-stage sampling: This is a further development of the idea of cluster sampling. This technique is meant for big inquiries extending to a considerably large geographical area like an entire country. Under multi-stage sampling the first stage may be to select large primary sampling units such as states, then districts, then towns and finally certain families within towns. If the technique of random sampling is applied at all stages, the sampling procedures described as multi-stage random sampling.

h. Sequential sampling: This is somewhat a complex sample design where the ultimate size of the sample is not fixed in advance but is determined according to mathematical decisions on the basis of information yielded as survey progresses. This design is usually adopted under acceptance sampling plan in the context of statistical quality control. In practice, several of the methods of sampling described above may well be used in the same study in which case it can be called mixed sampling. It may be pointed out here that normally one should resort to random sampling so that bias can be eliminated and sampling error can be estimated. But purposive sampling is considered desirable when the universe happens to be small and a known characteristic of it is to be studied intensively. Also, there are conditions under which sample

NOTES 

designs other than random sampling may be considered better for reasons like convenience and low costs. The sample design to be used must be decided by the researcher taking into consideration the nature of the inquiry and other related factors.

6. **Collecting the data:** In dealing with any real-life problem, it is often found that data at hand are inadequate, and hence, it becomes necessary to collect data that are appropriate. There are several ways of collecting the appropriate data which differ considerably in context of money costs, time and other resources at the disposal of the researcher. Primary data can be collected either through experiment or through survey. If the researcher conducts an experiment, he observes some quantitative measurements, or the data, with the help of which he examines the truth contained in his hypothesis. But in the case of a survey, data can be collected by any one or more of the following ways:

   i. **By observation:** This method implies the collection of information by way of investigator's own observation, without interviewing the respondents. The information obtained relates to what is currently happening and is not complicated by either the past behavior or future intentions or attitudes of respondents. This method is no doubt an expensive method and the information provided by this method is also very limited. As such this method is not suitable in inquiries where large samples are concerned.

   ii. **Through personal interview**: The investigator follows a rigid procedure and seeks answers to a set of pre-conceived questions through personal interviews. This method of collecting data is usually carried out in a structured way where output depends upon the ability of the interviewer to a large extent.

   iii. **Through telephone interviews:** This method of collecting information involves contacting the respondents on telephone itself. This is not a very widely used method but it plays an important role in industrial surveys in developed regions, particularly, when the survey has to be accomplished in a very limited time.

   iv. **By mailing of questionnaires:** The researcher and the respondents do come in contact with each other if this method of survey is adopted. Questionnaires are mailed to the respondents with a request to return after completing the same. It is the most extensively used method in various economic and business surveys. Before applying this method, usually Pilot Study for testing the questionnaire is conduced which reveals the weaknesses, if any, of the questionnaire. Questionnaire to be used must be prepared very carefully so that it may prove to be effective in collecting the relevant information.

   v. **Through schedules**: Under this method the enumerators are appointed and given training. They are provided with schedules containing relevant questions. These enumerators go to respondents with these schedules. Data are collected by filling up the schedules by enumerators on the basis of replies given by respondents.

   Much depends upon the capability of enumerators so far as this method is concerned. Some occasional field checks on the work of the enumerators may ensure sincere work. The researcher should select one of these methods of collecting the data taking into consideration the nature of investigation, objective and scope of the inquiry, financial resources, available time and the desired degree

of accuracy. Though he should pay attention to all these factors but much depends upon the ability and experience of the researcher. In this context Dr ALGOL very aptly remarks that in collection of statistical data common sense is the chief requisite and experience is the chief teacher.

7. **Execution of the project:** Execution of the project is a very important step in the research process. If the execution of the project proceeds on correct lines, the data to be collected would be adequate and dependable. The researcher should see that the project is executed in a systematic manner and in time. If the survey is to be conducted by means of structured questionnaires, data can be readily machine processed. In such a situation, questions as well as the possible answers may be coded. If the data are to be collected through interviewers, arrangements should be made for proper selection and training of the interviewers. The training may be given with the help of instruction manuals which explain clearly the job of the interviewers at each step. Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently. A careful watch should be kept for unanticipated factors in order to keep the survey as much realistic as possible. This, in other words, means that steps should be taken to ensure that the survey is under statistical control so that the collected information is in accordance with the pre-defined standard of accuracy. If some of the respondents do not cooperate, some suitable methods should be designed to tackle this problem. One method of dealing with the non-response problem is to make a list of the non-respondents and take a small sub-sample of them, and then with the help of expert's vigorous efforts can be made for securing response.

8. **Analysis of data:** After the data have been collected, the researcher turns to the task of analyzing them. The analysis of data requires a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences. The unwieldy data should necessarily be condensed into a few manageable groups and tables for further analysis. Thus, researcher should classify the raw data into some purposeful and usable categories. Coding operation is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted. Editing is the procedure that improves the quality of the data for coding. With coding the stage is ready for tabulation. Tabulation is a part of the technical procedure wherein the classified data are put in the form of tables. The mechanical devices can be made use of at this juncture.

   A great deal of data, especially in large inquiries, is tabulated by computers. Computers not only save time but also make it possible to study large number of variables affecting a problem simultaneously. Analysis work after tabulation is generally based on the computation of various percentages, coefficients, etc., by applying various well defined statistical formulae. In the process of analysis, relationships or differences supporting or conflicting with original or new hypotheses should be subjected to tests of significance to determine with what validity data can be said to indicate any conclusion(s).For instance, if there are two samples of weekly wages, each sample being drawn from factories indifferent parts of the same city, giving two different mean values, then our problem may be whether the two mean values are significantly different or the difference is just a matter of chance. Through the use of statistical tests, we can establish whether such a difference is a real one or is the result of random fluctuations. If the difference happens to be real, the inference will be that the two samples

NOTES 

Research come from different universes and if the difference is due to chance, the conclusion would be that the two samples belong to the same universe. Similarly, the technique of analysis of variance can help us in analyzing whether three or more varieties of seeds grown on certain fields yield significantly different results or not. In brief, the researcher can analyze the collected data with the help of various statistical measures.

9. **Hypothesis-testing:** After analyzing the data as stated above, the researcher is in a position to test the hypotheses, if any, he had formulated earlier. Do the facts support the hypotheses or they happen to be contrary? This is the usual question which should be answered while testing hypotheses. Various tests, such as Chi square test, t-test, F-test, have been developed by statisticians for the purpose. The hypotheses may be tested through the use of one or more of such tests, depending upon the nature and object of research inquiry. Hypothesis -testing will result in either accepting the hypothesis or in rejecting it. If the researcher had no hypotheses to start with, generalizations established on the basis of data may be stated as hypotheses to be tested by subsequent researches in times to come.
10. **Generalizations and interpretation:** If a hypothesis is tested and upheld several times, it may be possible for the researcher to arrive at generalization, i.e., to build a theory. As a matter of fact, the real value of research lies in its ability to arrive at certain generalizations. If the researcher had no hypothesis to start with, he might seek to explain his findings on the basis of some theory. It is known as interpretation. The process of interpretation may quite often trigger off new questions which in turn may lead to further researches.
11. **Preparation of the report or the thesis:** Finally, the researcher has to prepare the report of what has been done by him. Writing of report must be done with great care keeping in view the following: The layout of the report should be as follows: i. the preliminary pages; ii. the main text, and iii. the end matter. In its preliminary pages the report should carry title and date followed by acknowledgement sand foreword. Then there should be a table of contents followed by a list of tables and list of graphs and charts, if any, given in the report.

The main text of the report should have the following parts:

Introduction: It should contain a clear statement of the objective of the research and an explanation of the methodology adopted in accomplishing the research. The scope of the study along with various limitations should as well be stated in this part.

Summary of findings: After introduction there would appear a statement of finding sand recommendations in non-technical language. If the findings are extensive, they should be **summarized.**

Main report: The main body of the report should be presented in logical sequence and broken-down into readily identifiable sections.

Conclusion: Towards the end of the main text, researcher should again put down the results of his research clearly and precisely. In fact, it is the final summing up. At the end of the report, appendices should be enlisted in respect of all technical data. Bibliography, i.e., list of books, journals, reports, etc., consulted, should also be given in the end. Index should also be given specially in a published research report. Report should be written

in a concise and objective style in simple language avoiding vague expressions such as 'it seems,' 'there may be', and the like. Charts and illustrations in the main report should be used only if they present the information more clearly and forcibly. Calculated 'confidence limits' must be mentioned and the various constraints experienced in conducting research operations may as well be stated.

1.8 CRITERIA OF GOOD RESEARCH

Whatever may be the types of research and studies; one thing that is important is that they all meet on the common ground of the scientific method employed by them. One expects scientific research to satisfy the following criteria:

1. The purpose of the research should be clearly defined and common concepts bemused.
2. The research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.
3. The procedural design of the research should be carefully planned to yield results that areas objective as possible.
4. The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.
5. The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be appropriate. The validity and reliability of the data should be checked carefully.
6. Conclusions should be confined to those justified by the data of the research and limited to those for which the data provide an adequate basis.
7. Greater confidence in research is warranted if the researcher is experienced, has a good reputation in research, and is a person of integrity.

In other words, we can state the qualities of good research as under:

1. **Good research is systematic:** It means that research is structured with specified steps to be taken in a specified sequence in accordance with the well-defined set of rules. The systematic characterization of the research does not rule out creative thinking but it certainly does reject the use of guessing and intuition in arriving at conclusions.
2. **Good research is logical:** This implies that research is guided by the rules of logical reasoning and the logical process of induction and deduction are of great value in carrying out research. Induction is the process of reasoning from a part to the whole whereas deduction is the process of reasoning from some premise to a conclusion that follows from that very premise. In fact, logical reasoning makes research more meaningful in the context of decision-making.
3. **Good research is empirical:** It implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.
4. **Good research is replicable:** This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

NOTES 

1.9 RESEARCH PROBLEM

What is a Research Problem?

A research problem is a situation that causes the researcher to feel apprehensive, confused, and ill at ease. It is the demarcation of a problem area within a certain context involving the WHO or WHAT, the WHERE, the WHEN, and the WHY of the problem situation. There are many problem situations that may give rise to research. Three sources usually contribute to problem identification. Own experience or the experience of others may be a source of problem supply. A second source could be scientific literature. You may read about certain findings and notice that a certain field was not covered. This could lead to a research problem. Theories could be a third source. Shortcomings in theories could be researched. Research can thus be aimed at clarifying or substantiating an existing theory, clarifying contradictory findings, correcting a faulty methodology, correcting the inadequate or unsuitable use of statistical techniques, reconciling conflicting opinions, or solving existing practical problems.

Techniques Involved in Defining a Problem

As a researcher, you must have often read that defining a problem is the first step in a research process. But, have you ever wondered what is meant by defining a problem? Well, it simply means that the researcher has to lay down certain boundaries within which he/she has to study the problem with a predefined objective in mind. Defining a problem is a herculean task, and this must be done intelligently to avoid confusions that arise in the research operation. Try to follow the below steps systematically to best define a problem:

i. **State the problem in a general way: The** first state the problem in general terms with respect to some practical, scientific or intellectual interest. For this, the researcher may himself read the concerned subject matter thoroughly or take the help of the subject expert. Often, the guide states the problem in general terms; it depends on the researcher if he/she wants to narrow it down to operational terms. The problem stated should also be checked for ambiguity and feasibility.

ii. **Understand the nature of the problem:** The next step is to understand the nature and origin of the problem. The researcher needs to discuss the problem with those related to the subject matter in order to clearly understand the origin of the problem, its nature, objectives, and the environment in which the problem is to be studied.

iii. **Survey the available literature:** All available literature including relevant theories, reports, records, and other relevant literature on the problem needs to be reviewed and examined. This would help the researcher to identify the data available, the techniques that might be used, types of difficulties that may be encountered during the study, possible analytical shortcomings, and even new methods of approach to the present problem.

iv. **Go for discussions for developing ideas:** The researcher may discuss the problem with his/her colleagues and others related to the concerned subject. This helps the researcher to generate new ideas, identify different aspects on the problem, gain suggestions and advice from others, and sharpen his focus on certain aspects within the field. However, discussions should not be limited to the

problem only, but should also be related to the general approach to the problem, techniques that might be used, possible solutions, etc.

v. **Rephrase the research problem into a working proposition:** Finally, the researcher must rephrase the problem into a working proposition. Rephrasing the problem means putting the problem in specific terms that is feasible and may help in the development of working hypotheses. Once the researcher has gone through the above steps systematically, it is easy to rephrase the problem into analytical and operational terms.

1.10 CHAPTER SUMMARY

Research may be very broadly defined as systematic gathering of data and information and its analysis for advancement of knowledge in any subject. Research attempts to find answer intellectual and practical questions through application of systematic methods. Webster's Collegiate Dictionary defines research as "studious inquiry or examination investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts, or practical application of such new or revised theories or laws". Some people consider research as a movement, a movement from the known to the unknown. It is actually a voyage of discovery. The research may be exploratory or it may be formalized. The objective of exploratory research is the development of hypotheses rather than their testing, whereas formalized research studies are those with substantial structure and with specific hypotheses to be tested. It seems appropriate at this juncture to explain the difference between research methods and research methodology. Research methods may be understood as all those methods/ techniques that are used for the conduction of research. Research methods or techniques, thus, refer to the methods the researchers use in performing research operations. In other words, all those methods which are used by the researcher during the course of studying his research problem are termed research methods. As per the research requirement, different methods are used. Survey research can be used in large scale research involving large numbers of respondents and investigators. It is the most common example of a macro method that can be used in such a circumstance. On the other hand, in the study of a particular group A micro method like participant observation can be used. Before embarking on the details of research methodology and techniques, it seems appropriate to present a brief overview of the research process. Research process consists of series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps. A research problem is a situation that causes the researcher to feel apprehensive, confused, and ill at ease. It is the demarcation of a problem area within a certain context involving the WHO or WHAT, the WHERE, the WHEN, and the WHY of the problem situation.

1.11 REVIEW QUESTIONS

SHORT ANSWER TYPE QUESTION

1. **Drive the meaning of research.**
2. **What is the purpose of research?**

NOTES

3. **Elaborate the types of research.**
4. **What do you understand by research approaches?**
5. **Explain research process.**

LONG ANSWER TYPE QUESTIONS

1. **Enlist the criteria of good research**
2. **What is a research problem? Explain briefly.**
3. **Briefly elaborate the significance of research methodologies.**
4. **What are some other types of research methods?**
5. **Explain research methods versus methodology.**

1.12 MULTIPLE CHOICE QUESTIONS

1. **_____, after the data have been collected, the researcher turns to the task of analyzing them.**
   a. Collecting the data
   b. Execution of the project
   c. Quota sampling
   d. Analysis of data
2. **After analyzing the data as stated above, the researcher is in a position to test the _____.**
   a. Hypothesis-testing
   b. Collecting the data
   c. Execution of the project
   d. Quota sampling
3. **_____ it means that research is structured with specified steps to be taken in a specified sequence in accordance with the well-defined set of rules.**
   a. Good research is logical
   b. Good research is systematic
   c. Good research is empirical
   d. Good research is replicable
4. **_____ this implies that research is guided by the rules of logical reasoning and the logical process of induction and deduction are of great value in carrying out research.**
   a. Good research is systematic
   b. Good research is empirical
   c. Good research is logical
   d. Good research is replicable
5. **_____it implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.**
   a. Good research is systematic
   b. Good research is empirical

c. Good research is logical
d. Good research is replicable

6. _____ concentrates on finding facts to ascertain the nature of something as it exists.

a. Descriptive research
b. Applied research
c. Quantitative research
d. Conceptual research

7. _____ is carried out to find answers to practical problems to be solved and as an aid in decision making in different areas including product design, process design, and policy making.

a. Descriptive research
b. Quantitative research
c. Conceptual research
d. Applied research

8. _____ studies such aspects of the research subject which are not quantifiable, and hence not subject to measurement and quantitative analysis.

a. Descriptive research
b. Quantitative research
c. Conceptual research
d. Applied research

9. _____ involves investigation of thoughts and ideas and developing new ideas or interpreting the old ones based on logical reasoning.

a. Descriptive research
b. Quantitative research
c. Conceptual research
d. Applied research

10. _____is a further development of the idea of cluster sampling. This technique is meant for big inquiries extending to a considerably large geographical area like an entire country.

a. Systematic sampling
b. Stratified sampling
c. Quota sampling
d. Multi-stage sampling

♦♦♦♦♦

UNIT II

METHODS OF DATA COLLECTION

STRUCTURE

NOTES

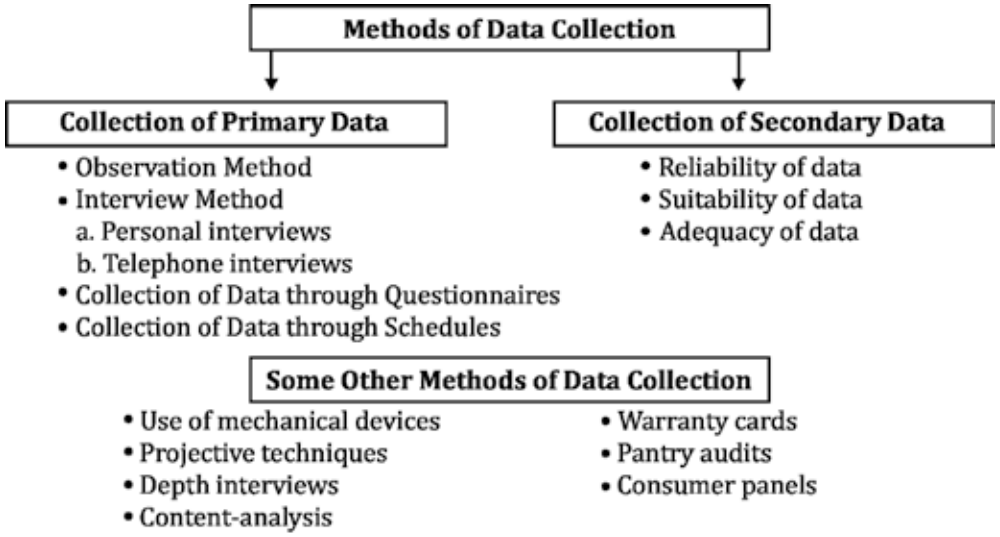
2.1 LEARNING OBJECTIVE

After learning this unit students will be able to:

- Understand the Methods of Data Collection.
- Understand the Case Study Research.
- Understand the Survey Research Definition.

2.2 METHODS OF DATA COLLECTION

The method of data collection to be used for the study, the researcher should keep in mind two types of data viz., primary and secondary. The primary data are those which are collected afresh and for the first time, and thus happen to be original in character. The secondary data, on the other hand, are those which have already been collected by someone else and which have already been passed through the statistical process. The researcher would have to decide which sort of data he would be using (thus collecting) for his study and accordingly he will have to select one or the other method of data collection. The methods of collecting primary and secondary data differ since primary data are to be originally collected, while in the case of secondary data the nature of data collection work is merely that of compilation. We describe the different methods of data collection, with the pros and cons of each method.



Methods of Data Collection

STEPS FOR DATA COLLECTION

Identify issues and opportunities for collecting data: Every tool for collecting data has its own pros and cons. Thus, for deciding the best method, it is important to identify issues and opportunities for collecting data according to the method. It might be helpful to engage in a pilot study to review our tools and sample size.

- **Setting goals and objectives:** The researcher uses data to address his/her research questions and must design his/her methodology accordingly. Thus, every tool used by the researcher must have certain objectives which could be used for addressing these questions after analysis.

NOTES

- **Planning approach and methods:** Researcher would make decisions pertaining to who will be surveyed, how data will be collected, sources and tools for data collection, and duration of the project.
- **Collect data:** While planning the data collection, it is important to understand logistical challenges and prepare accordingly.

COLLECTION OF PRIMARY DATA

We collect primary data during the course of doing experiments in experimental research but in case we do research of the descriptive type and perform surveys, whether sample surveys or census surveys, then we can obtain primary data either through observation or through direct communication with respondents in one form or another or through personal interviews. There are several methods of collecting primary data, particularly in surveys and descriptive researches which are briefly described below;

1. **Observation Method:** The observation method is the most commonly used method especially in studies relating to behavioral sciences. In a way, we all observe things around us, but this sort of observation is not scientific observation. 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. Under the observation method, the information is sought by way of the investigator's own direct observation without asking from the respondent. For instance, in a study relating to consumer behavior, the investigator instead of asking the brand of wristwatch used by the respondent, may himself look at the watch.

   The main advantage of this method is that subjective bias is eliminated if the observation is done accurately. Secondly, the information obtained under this method relates to what is currently happening; it is not complicated by either the past behavior or future intentions or attitudes. Thirdly, this method is independent of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents as happens to be the case in the interview or the questionnaire method. This method is particularly suitable in studies that deal with subjects (i.e., respondents) who are not capable of giving verbal reports of their feelings for one reason or the other. However, the observation method has various limitations. Firstly, it is an expensive method. Secondly, the information provided by this method is very limited. Thirdly, sometimes unforeseen factors may interfere with the observational task. At times, the fact that some people are rarely accessible to direct observation creates obstacles for this method to collect data effectively While using this method, the researcher should keep in mind things like What should be observed? How the observations should be recorded? Or how the accuracy of observation can be ensured?

   - **Interview Method:** The interview method of collecting data involves the presentation of oral-verbal stimuli and reply in terms of oral-verbal responses. This method can be used through personal interviews and, if possible, through telephone interviews.

     a. **Personal interviews:** Despite the variations in interview techniques, the major advantages and weaknesses of personal interviews can be

enumerated in a general way. The chief merits of the interview method are as follows:

i. More information and that too in greater depth can be obtained.

ii. Interviewer by his own skill can overcome the resistance, if any, of the respondents; the interview method can be made to yield an almost perfect sample of the general population.

iii. There is greater flexibility under this method as the opportunity to restructure questions is always there, especially in the case of unstructured interviews.

iv. The observation method can as well be applied to recording verbal answers to various questions.

v. Personal information can as well be obtained easily under this method.

vi. Samples can be controlled more effectively as there arises no difficulty of the missing returns; non-response generally remains very low.

vii. The interviewer can usually control which person(s) will answer the questions. This is not possible in the mailed questionnaire approach. If so desired, group discussions may also be held.

viii. The interviewer may catch the informant off-guard and thus may secure the most spontaneous reactions than would be the case if a mailed questionnaire is used.

ix. The language of the interview can be adapted to the ability or educational level of the person interviewed and as such misinterpretations concerning questions can be avoided.

x. The interviewer can collect supplementary information about the respondent's personal characteristics and environment which is often of great value in interpreting results.

But there are also certain weaknesses of the interview method. Among the important weaknesses, mention may be made of the following:

i. It is a very expensive method, especially when a large and widely spread geographical sample is taken.

ii. There remains the possibility of the bias of the interviewer as well as that of the respondent; there also remains the headache of supervision and control of interviewers.

iii. Certain types of respondents such as important officials or executives or people in high-income groups may not be easily approachable under this method and to that extent, the data may prove inadequate.

iv. This method is relatively more time-consuming, especially when the sample is large and recalls upon the respondents are necessary.

v. The presence of the interviewer on the spot may over-stimulate the respondent, sometimes even to the extent that he may give imaginary information just to make the interview interesting.

vi. Under the interview method, the organization required for selecting, training, and supervising the field staff is more complex with formidable problems.

NOTES 

vii. Interviewing at times may also introduce systematic errors.

viii. Effective interview presupposes proper rapport with respondents that would facilitate free and frank responses. This is often a very difficult requirement.

b. Telephone interviews: This method of collecting information consists in contacting respondents on the telephone itself. It is not a very widely used method, but plays an important part in industrial surveys, particularly in developed regions. The chief merits of such a system are:

i. It is more flexible in comparison to the mailing method.

ii. It is faster than other methods i.e., a quick way of obtaining information.

iii. It is cheaper than the personal interviewing method; here the cost per response is relatively low.

iv. Recall is easy; call-backs are simple and economical.

v. There is a higher rate of response than what we have in the mailing method; the non-response is generally very low.

vi. Replies can be recorded without causing embarrassment to respondents.

vii. Interviewers can explain requirements more easily.

viii. At times, access can be gained to respondents who otherwise cannot be contacted for one reason or the other.

ix. No field staff is required.

x. Representative and wider distribution of the sample is possible.

xi. But this system of collecting information is not free from demerits. Some of these may be highlighted.

xii. Little time is given to respondents for considered answers; the interview period is not likely to exceed five minutes in most cases.

xiii. Surveys are restricted to respondents who have telephone facilities.

xiv. Extensive geographical coverage may get restricted by cost considerations.

xv. It is not suitable for intensive surveys where comprehensive answers are required to various questions.

xvi. The possibility of the bias of the interviewer is relatively more.

xvii. Questions have to be short and to the point; probes are difficult to handle.

2. Collection of Data through Questionnaires: This method of data collection is quite popular, particularly in the case of big inquiries. It is being adopted by private individuals, research workers, private and public organizations, and even by governments. In this method, a questionnaire is sent (usually by post) to the persons concerned with a request to answer the questions and return the questionnaire. A questionnaire consists of a number of questions printed or typed in a definite order on a form or set of forms. The questionnaire is mailed to respondents who are expected to read and understand the questions and write down the reply in the space meant for the purpose of the questionnaire itself. The respondents have to answer the questions on their own.

NOTES

The method of collecting data by mailing the questionnaires to respondents is most extensively employed in various economic and business surveys. The merits claimed on behalf of this method are as follows:

i. There is low cost even when the universe is large and is widely spread geographically.

ii. It is free from the bias of the interviewer; answers are in respondents' own words.

iii. Respondents have adequate time to give well-thought-out answers.

iv. Respondents, who are not easily approachable, can also be reached conveniently.

v. Large samples can be made use of and thus the results can be made more dependable and reliable.

The main demerits of this system can also be listed here:

i. Low rate of return of the duly filled in questionnaires; bias due to no-response is often indeterminate.

ii. It can be used only when respondents are educated and cooperating.

iii. The control over the questionnaire may be lost once it is sent.

iv. There is inbuilt inflexibility because of the difficulty of amending the approach once questionnaires have been dispatched.

v. There is also the possibility of ambiguous replies or omission of replies altogether to certain questions; interpretation of omissions is difficult.

vi. It is difficult to know whether willing respondents are truly representative.

vii. This method is likely to be the slowest of all.

3. **Collection of Data through Schedules:** This method of data collection is very much like the collection of data through questionnaires, with little difference which lies in the fact that schedules are being filled in by the enumerators who are specially appointed for the purpose. In certain situations, schedules may be handed over to respondents and enumerators may help them in recording their answers to various questions in the said schedules. Enumerators explain the aims and objects of the investigation and also remove the difficulties which any respondent may feel in understanding the implications of a particular question or the definition or concept of difficult terms.

SOME OTHER METHODS OF DATA COLLECTION

Let us consider some other methods of data collection, particularly used by big business houses in modern times.

1. **Warranty cards:** Warranty cards are usually postal-sized cards that are used by dealers of consumer durables to collect information regarding their products. The information sought is printed in the form of questions on the 'warranty cards' which are placed inside the package along with the product with a request to the consumer to fill in the card and post it back to the dealer.

2. **Distributor or store audits:** Distributor or store audits are performed by distributors as well as manufacturers through their salesmen at regular intervals. Distributors get the retail stores audited through salesmen and use such

NOTES 

information to estimate market size, market share, and seasonal purchasing patterns, and so on. The data are obtained in such audits not by questioning but by observation.

3. **Pantry audits:** The pantry audit technique is used to estimate the consumption of the basket of goods at the consumer level. In this type of audit, the investigator collects an inventory of types, quantities, and prices of commodities consumed. Thus in pantry audit data are recorded from the examination of the consumer's pantry. The usual objective in a pantry audit is to find out what types of consumers buy certain products and certain brands, the assumption being that the contents of the pantry accurately portray consumer's preferences. Quite often, pantry audits are supplemented by direct questioning relating to reasons and circumstances under which particular products were purchased in an attempt to relate these factors to purchasing habits.
4. **Consumer panels:** A consumer panel is essentially a sample of consumers who are interviewed repeatedly over a period of time. A continuing consumer panel is often set up for an indefinite period with a view to collect data on a particular aspect of consumer behavior over time, generally at periodic intervals or maybe meant to serve as a general-purpose panel for researchers on a variety of subjects.
5. **Use of mechanical devices**: The use of mechanical devices has been widely made to collect information by way of indirect means. Eye camera, Pupilometric camera, Psych-galvanometer, Motion picture camera, and Audiometer are the principal devices so far developed and commonly used by modern big business houses, mostly in the developed world for the purpose of collecting the required information.
6. **Projective techniques:** for the collection of data have been developed by psychologists to use projections of respondents for inferring about underlying motives, urges, or intentions which are such that the respondent either resists to reveal them or is unable to figure out himself. In projective techniques, the respondent in supplying information tends unconsciously to project his own attitudes or feelings on the subject under study. Projective techniques play a significant role in motivational researches or in attitude surveys. The use of these techniques requires intensive specialized training.
7. **Depth interviews:** Depth interviews are those interviews that are designed to discover underlying motives and desires and are often used in motivational research. Such interviews are held to explore the needs, desires, and feelings of respondents. In other words, they aim to elicit unconscious as also other types of material relating especially to personality dynamics and motivations. As such, depth interviews require great skill on the part of the interviewer and at the same time involve considerable time. Unless the researcher has the specialized training, depth interviewing should not be attempted.
8. **Content-analysis:** Content-analysis consists of analyzing the contents of documentary materials such as books, magazines, newspapers, and the contents of all other verbal materials which can be either spoken or printed.

A review of research in any area, for instance, involves the analysis of the contents of research articles that have been published. The analysis may be at a relatively simple level or maybe a subtle one. It is at a simple level when we pursue it on the basis of certain characteristics of the document or verbal materials that can be identified and counted.

NOTES 

COLLECTION OF SECONDARY DATA

Secondary data means data that are already available i.e., they refer to the data which have already been collected and analysed by someone else. When the researcher utilizes secondary data, then he has to look into various sources from where he can obtain them. In this case, he is certainly not confronted with the problems that are usually associated with the collection of original data. Secondary data may either be published data or unpublished data. Usually **published data** are available in (a) various publications of the central, state are local governments; (b) various publications of foreign governments or of international bodies and their subsidiary organizations; (c) technical and trade journals; (d) books, magazines and newspapers; (e) reports and publications of various associations connected with business and industry, banks, stock exchanges, etc.; (f) reports prepared by research scholars, universities, economists, etc. in different fields; and (g) public records and statistics, historical documents, and other sources of published information. The sources of **unpublished data** are many; they may be found in diaries, letters, unpublished biographies, and autobiographies and also may be available with scholars and research workers, trade associations, labor bureaus, and other public/ private individuals and organizations.

A researcher must be very careful in using secondary data because it is just possible that the secondary data may be unsuitable or may be inadequate in the context of the problem which the researcher wants to study. By way of caution, the researcher, before using secondary data, must see that they possess the following characteristics:

1. **Reliability of data:** The reliability can be tested by finding out such things about the said data:
   (a) Who collected the data?
   (b) What were the sources of data?
   (c) Were they collected by using proper methods?
   (d) At what time were they collected?
   (e) Was there any bias of the compiler?
   (t) What level of accuracy was desired? Was it achieved?
2. **Suitability of data:** The data that is suitable for one inquiry may not necessarily be found suitable in another inquiry. Hence, if the available data are found to be unsuitable, they should not be used by the researcher. In this context, the researcher must very carefully scrutinize the definition of various terms and units of collection used at the time of collecting the data from the primary source originally. Similarly, the object, scope, and nature of the original inquiry must also be studied. If the researcher finds differences in these, the data will remain unsuitable for the present inquiry and should not be used.
3. **Adequacy of data:** If the level of accuracy achieved in data is found inadequate for the purpose of the present inquiry, they will be considered inadequate and should not be used by the researcher. The data will also be considered inadequate if they are related to an area that may be either narrower or wider than the area of the present inquiry.

NOTES 

SELECTION OF APPROPRIATE METHOD FOR DATA COLLECTION

Thus, there are various methods of data collection. As such the researcher must judiciously select the method/methods for his own study, keeping in view the following factors:

1. **Nature, scope, and object of inquiry:** This constitutes the most important factor affecting the choice of a particular method. The method selected should be such that it suits the type of inquiry that is to be conducted by the researcher. This factor is also important in deciding whether the data already available (secondary data) are to be used or the data not yet available (primary data) are to be collected.
2. **Availability of funds:** The availability of funds for the research project determines to a large extent the method to be used for the collection of data. When funds at the disposal of the researcher are very limited, he will have to select a comparatively cheaper method that may not be as efficient and effective as some other costly method. Finance, in fact, is a big constraint in practice and the researcher has to act within this limitation.
3. **Time factor:** Availability of time has also to be taken into account in deciding a particular method of data collection. Some methods take relatively more time, whereas with others the data can be collected in a comparatively shorter duration. The time at the disposal of the researcher, thus, affects the selection of the method by which the data are to be collected.
4. **Precision required:** Precision required is yet another important factor to be considered at the time of selecting the method of collection of data.

Finally, we can say that the most desirable approach with regard to the selection of the method depends on the nature of the particular problem and on the time and resources (money and personnel) available along with the desired degree of accuracy. But, over and above all this, much depends upon the ability and experience of the researcher.

2.3 CASE STUDY RESEARCH

Introduction

The case study method is a very popular form of qualitative analysis and involves a careful and complete observation of a social unit, be that unit a person, a family, an institution, a cultural group or even the entire community. It is a method of study in depth rather than breadth. The case study places more emphasis on the full analysis of a limited number of events or conditions and their interrelations. The case study deals with the processes that take place and their interrelationship. Thus, case study is essentially an intensive investigation of the particular unit under consideration. The object of the case study method is to locate the factors that account for the behaviour-patterns of the given unit as an integrated totality.

Case study method of Data Collection

According to H. Odum, "The case study method of data collection is a technique by which individual factor whether it be an institution or just an episode in the life of an individual or a group is analysed in its relationship to any other in the group." Thus, a fairly exhaustive study of a person (as to what he does and has done, what he thinks he does and had done and what he expects to do and says he ought to do) or group is called a life or case history. Burgess has used the words "the social microscope" for the case study method."

Pauline V. Young describes case study as "a comprehensive study of a social unit be that unit a person, a group, a social institution, a district or a community." In brief, we can say that case study method is a form of qualitative analysis where in careful and complete observation of an individual or a situation or an institution is done; efforts are made to study each and every aspect of the concerning unit in minute details and then from case data generalisations and inferences are drawn.

Characteristics of Case Study method

The important characteristics of the case study method are as under:

- Under this method the researcher can take one single social unit or more of such units for his study purpose; he may even take a situation to study the same comprehensively.
- Here the selected unit is studied intensively i.e., it is studied in minute details. Generally, the study extends over a long period of time to ascertain the natural history of the unit so as to obtain enough information for drawing correct inferences.
- In the context of this method, we make complete study of the social unit covering all facets. Through this method we try to understand the complex of factors that are operative within a social unit as an integrated totality.
- Under this method the approach happens to be qualitative and not quantitative. Mere quantitative information is not collected. Every possible effort is made to collect information concerning all aspects of life. As such, case study deepens our perception and gives us a clear insight into life. For instance, under this method we not only study how many crimes a man has done but shall peep into the factors that forced him to commit crimes when we are making a case study of a man as a criminal. The objective of the study may be to suggest ways to reform the criminal.
- In respect of the case study method an effort is made to know the mutual inter-relationship of causal factors.
- Under case study method the behaviour pattern of the concerning unit is studied directly and not by an indirect and abstract approach.
- Case study method results in fruitful hypotheses along with the data which may be helpful in testing them, and thus it enables the generalised knowledge to get richer and richer. In its absence, generalised social science may get handicapped.

Evolution and scope

The case study method is a widely used systematic field research technique in sociology these days. The credit for introducing this method to the field of social investigation goes to Frederic Le Play who used it as a hand-maiden to statistics in his studies of family budgets. Herbert Spencer was the first to use case material in his comparative study of different cultures. Dr. William Healy resorted to this method in his study of juvenile delinquency, and considered it as a better method over and above the mere use of statistical data. Similarly, anthropologists, historians, novelists and dramatists have used this method

NOTES

concerning problems pertaining to their areas of interests. Even management experts use case study methods for getting clues to several management problems. In brief, case study method is being used in several disciplines. Not only this, its use is increasing day by day.

Assumptions

The case study method is based on several assumptions. The important assumptions may be listed as follows:

- The assumption of uniformity in the basic human nature in spite of the fact that human
- behaviour may vary according to situations.
- The assumption of studying the natural history of the unit concerned.
- The assumption of comprehensive study of the unit concerned.

Major phases involved

Major phases involved in case study are as follows:

- Recognition and determination of the status of the phenomenon to be investigated or the unit of attention.
- Collection of data, examination and history of the given phenomenon.
- Diagnosis and identification of causal factors as a basis for remedial or developmental treatment.
- Application of remedial measures i.e., treatment and therapy (this phase is often characterized as case work).
- Follow-up programme to determine effectiveness of the treatment applied.

Advantages

There are several advantages of the case study method that follow from the various characteristics outlined above. Mention may be made here of the important advantages.

- Being an exhaustive study of a social unit, the case study method enables us to understand fully the behaviour pattern of the concerned unit. In the words of Charles Horton Cooley, "case study deepens our perception and gives us a clearer insight into life…. It gets at behaviour directly and not by an indirect and abstract approach."
- Through case study a researcher can obtain a real and enlightened record of personal experiences which would reveal man's inner strivings, tensions and motivations that drive him to action along with the forces that direct him to adopt a certain pattern of behaviour.
- This method enables the researcher to trace out the natural history of the social unit and its relationship with the social factors and the forces involved in its surrounding environment.
- It helps in formulating relevant hypotheses along with the data which may be helpful in testing them. Case studies, thus, enable the generalised knowledge to

get richer and richer.

- The method facilitates intensive study of social units which is generally not possible if we use either the observation method or the method of collecting information through schedules. This is the reason why case study method is being frequently used, particularly in social researches.
- Information collected under the case study method helps a lot to the researcher in the task of constructing the appropriate questionnaire or schedule for the said task requires thorough knowledge of the concerning universe.
- The researcher can use one or more of the several research methods under the case study method depending upon the prevalent circumstances. In other words, the use of different methods such as depth interviews, questionnaires, documents, study reports of individuals, letters, and the like is possible under case study method.
- Case study method has proved beneficial in determining the nature of units to be studied along with the nature of the universe. This is the reason why at times the case study method is alternatively known as "mode of organising data".
- This method is a means to well understand the past of a social unit because of its emphasis of historical analysis. Besides, it is also a technique to suggest measures for improvement in the context of the present environment of the concerned social units.
- Case studies constitute the perfect type of sociological material as they represent a real record of personal experiences which very often escape the attention of most of the skilled researchers using other techniques.
- Case study method enhances the experience of the researcher and this in turn increases his analysing ability and skill.
- This method makes possible the study of social changes. On account of the minute study of the different facets of a social unit, the researcher can well understand the social change then and now. This also facilitates the drawing of inferences and helps in maintaining the continuity of the research process. In fact, it may be considered the gateway to and at the same time the final destination of abstract knowledge.
- Case study techniques are indispensable for therapeutic and administrative purposes. They are also of immense value in taking decisions regarding several management problems. Case data are quite useful for diagnosis, therapy and other practical case problems.

Limitations

Important limitations of the case study method may as well be highlighted.

- Case situations are seldom comparable and as such the information gathered in case studies is often not comparable. Since the subject under case study tells history in his own words, logical concepts and units of scientific classification have to be read into it or out of it by the investigator.

NOTES 

- Read Bain does not consider the case data as significant scientific data since they do not provide knowledge of the "impersonal, universal, non-ethical, non-practical, repetitive aspects of phenomena." Real information is often not collected because the subjectivity of the researcher does enter in the collection of information in a case study.
- The danger of false generalisation is always there in view of the fact that no set rules are followed in collection of the information and only few units are studied.
- It consumes more time and requires lot of expenditure. More time is needed under case study method since one study the natural history cycles of social units and that too minutely.
- The case data are often vitiated because the subject, according to Read Bain, may write what he thinks the investigator wants; and the greater the rapport, the more subjective the whole process is.
- Case study method is based on several assumptions which may not be very realistic at times, and as such the usefulness of case data is always subject to doubt.
- Case study method can be used only in a limited sphere., it is not possible to use it in case of a big society. Sampling is also not possible under a case study method.
- Response of the investigator is an important limitation of the case study method. He often thinks that he has full knowledge of the unit and can himself answer about it. In case the same is not true, then consequences follow. In fact, this is more the fault of the researcher rather than that of the case method.

2.4 SURVEY RESEARCH DEFINITION

Survey Research is defined as the process of conducting research using surveys that researchers send to survey respondents. The data collected from surveys is then statistically analyzed to draw meaningful research conclusions.

In the 21st century, every organization's eager to understand what their customers think about their products or services and make better business decisions. Researchers can conduct research in multiple ways, but surveys are proven to be one of the most effective and trustworthy research methods. An online survey is a method for extracting information about a significant business matter from an individual or a group of individuals. It consists of structured survey questions that motivate the participants to respond, Creditable survey research can give these businesses access to a vast information bank. Organizations in media, other companies, and even governments rely on survey research to obtain accurate data.

The traditional definition of survey research is a quantitative method for collecting information from a pool of respondents by asking multiple survey questions. This research type includes the recruitment of individuals, collection, and analysis of data. It's useful for researchers who aim at communicating new features or trends to their respondents. Generally, it's the primary step towards obtaining quick information about mainstream topics and conducting more rigorous and detailed quantitative research methods like

surveys/polls or qualitative research methods like focus groups/on-call interviews can follow. There are many situations where researchers can conduct research using a blend of both qualitative and quantitative strategies.

SURVEY RESEARCH METHODS

Survey research methods can be derived based on two critical factors: Survey research tool and time involved to conduct research. There are three main survey research methods, divided based on the medium of conducting survey research:

- **Online/ Email:** Online survey research is one of the most popular survey research methods today. The cost involved in online survey research is extremely minimal, and the responses gathered are highly accurate.
- **Phone:** Survey research conducted over the telephone (CATI) can be useful in collecting data from a more extensive section of the target population. There are chances that the money invested in phone surveys will be higher than other mediums, and the time required will be higher.
- **Face-to-face:** Researchers conduct face-to-face in-depth interviews in situations where there is a complicated problem to solve. The response rate for this method is the highest, but it can be costly.

Further, based on the time taken, survey research can be classified into two methods:

- **Longitudinal survey research:** Longitudinal survey research involves conducting survey research over a continuum of time and spread across years and decades. The data collected using this survey research method from one time period to another is qualitative or quantitative. Respondent behavior, preferences, attitudes are continuously observed over time to analyze reasons for a change in behavior or preferences. For example, suppose a researcher intends to learn about the eating habits of teenagers. In that case, he/she will follow a sample of teenagers over a considerable period to ensure that the collected information is reliable. Often, cross-sectional survey research follows a longitudinal study.
- **Cross-sectional survey research:** Researchers conduct a cross-sectional survey to collect insights from a target audience at a particular time interval. This survey research method is implemented in various sectors such as retail, education, healthcare, SME businesses, etc. Cross-sectional survey research can either be descriptive or analytical. It is quick and helps researchers collected information in a brief period. Researchers rely on cross-sectional survey research method in situations where descriptive analysis of a subject is required.

Survey research also is bifurcated according to the sampling methods used to form samples for research: Probability and Non-probability sampling. Every individual of a population should be considered equally to be a part of the survey research sample. Probability sampling is a sampling method in which the researcher chooses the elements based on probability theory. The are various probability research methods such as simple random sampling, systematic sampling, cluster sampling, stratified random sampling, etc. Non-probability sampling is a sampling method where the researcher uses his/her knowledge and experience to form samples. The various non-probability sampling techniques are

NOTES 

convenience sampling, snowball sampling, consecutive sampling, judgemental sampling, and quota sampling.

PROCESS OF IMPLEMENTING SURVEY RESEARCH METHODS

- **Decide survey questions:** Brainstorm and put together valid survey questions that are grammatically and logically appropriate. Understanding the objective and expected outcomes of the survey helps a lot. There are many surveys where details of responses are not as important as gaining insights about what customers prefer from the provided options. In such situations, a researcher can include multiple-choice questions or closed-ended questions. Whereas, if researchers need to obtain details about specific issues, they can consist of open-ended questions to the questionnaire. Ideally, the surveys should include a smart balance of open-ended and closed-ended questions. Use survey questions like Likert Scale, Semantic Scale, Net Promoter Score question, etc. to avoid fence-sitting.
- **Finalize a target audience:** Send out relevant surveys as per the target audience and filter out irrelevant questions as per the requirement. The survey research will be instrumental in case the target population decides a sample. This way, results can be according to the desired market and be generalized to the entire population.
- **Send out surveys via decided mediums:** Distribute the surveys to the target audience and patiently wait for the feedback and comments- this is the most crucial step of the survey research. The survey needs to be scheduled, keeping in mind the nature of the target audience and its regions. Surveys can be conducted via email, embedded in a website, shared via social media, etc. to gain maximum responses.
- **Analyze survey results:** Analyze the feedback in real-time and identify patterns in the responses which might lead to a much-needed breakthrough for your organization. GAP, TURF, Conjoint analysis, Cross tabulation, and many such survey feedback analysis methods can be used to spot and shed light on respondent behavior. Researchers can use the results to implement corrective measures to improve customer/employee satisfaction.

REASONS TO CONDUCT SURVEY RESEARCH

The most crucial and integral reason for conducting market research using surveys is that you can collect answers regarding specific, essential questions. You can ask these questions in multiple formats as per the target audience and the intent of the survey. Before designing a study, every organization must figure out the objective of carrying this out so that the study can be structured, planned, and executed to perfection.

- **Questions that need to be on your mind while designing a survey are:**
- What is the primary aim of conducting the survey?
- How do you plan to utilize the collected survey data?
- What type of decisions you plan to take based on the points mentioned above?

- There are three critical reasons why an organization must conduct survey research.
- **Understand respondent behavior to get solutions to your queries:** If you've carefully curated a survey, the respondents will provide insights about what they like about your organization as well as suggestions for improvement. To motivate them to respond, you must be very vocal about how secure their responses will be and how you will utilize the answers. This will push them to be 100% honest about their feedback, opinions, and comments. Online surveys or mobile surveys have proved their privacy, and due to this, more and more respondents feel free to put forth their feedback through these mediums.
- **Present a medium for discussion:** A survey can be the perfect platform for respondents to provide criticism or applause for an organization. Important topics like product quality or quality of customer service etc. can be put on the table for discussion. A way you can do it is by including open-ended questions where the respondents can write their thoughts. This will make it easy for you to correlate your survey to what you intend to do with your product or service.
- **Strategy for never-ending improvements:** An organization can establish the target audience's attributes from the pilot phase of survey research. Researchers can use the criticism and feedback received from this survey to improve the product/services. Once the company successfully makes the improvements, it can send out another survey to measure the change in feedback keeping the pilot phase the benchmark. By doing this activity, the organization can track what was effectively improved and what still needs improvement.

SURVEY RESEARCH SCALES

There are four main scales for measurement of variables:

- **Nominal Scale:** A nominal scale associates numbers with variables for mere naming or labelling, and the numbers usually have no other relevance. It is the most basic of the four levels of measurement.
- **Ordinal Scale:** The ordinal scale has an innate order within the variables along with labels. It establishes the rank between the variables of a scale but not the difference value between the variables.
- **Interval Scale:** The interval scale is a step ahead in comparison to the other two scales. Along with establishing a rank and name of variables, the scale also makes known the difference between the two variables. The only drawback is that there is no fixed start point of the scale, i.e., the actual zero value is absent.
- **Ratio Scale:** The ratio scale is the most advanced measurement scale, which has variables that are labeled in order and have a calculated difference between variables. In addition to what interval scale orders, this scale has a fixed starting point, i.e., the actual zero value is present.

NOTES 

BENEFITS OF SURVEY RESEARCH

In case survey research is used for all the right purposes and is implemented properly, marketers can benefit by gaining useful, trustworthy data that they can use to better the ROI of the organization.

Other benefits of survey research are

- **Minimum investment:** Mobile surveys and online surveys have minimal finance invested per respondent. Even with the gifts and other incentives provided to the people who participate in the study, online surveys are extremely economical compared to the paper-based surveys.
- **Versatile sources for response collection:** You can conduct surveys via various mediums like online and mobile surveys. You can further classify them into qualitative mediums like focus groups, interviews, and quantitative mediums like customer-centric surveys. Due to the offline survey response collection option, researchers can conduct surveys in remote areas with limited internet connectivity. This can make data collection and analysis more convenient and extensive.
- **Reliable for respondents:** Surveys are extremely secure as the respondent details and responses are kept safeguarded. This anonymity makes respondents answer the survey questions candidly and with absolute honesty. An organization seeking to receive explicit responses for its survey research must mention that it will be confidential.

SURVEY RESEARCH DESIGN

Researchers implement a survey research design in cases where there is a limited cost involved, and there is a need to access details easily. This method is often used by small and large organizations to understand and analyze new trends, market demands, and opinions. Collecting information through a tactfully designed survey research can be much more effective and productive than a casually conducted survey.

There are five stages of survey research design:

- **Decide an aim of the research:** There can be multiple reasons for a researcher to conduct a survey, but they need to decide a purpose for research. This is the primary stage of survey research as it can mold the entire path of a survey, impacting its results.
- **Filter the sample from target population:** Who to target? is an essential question that a researcher should answer and keep in mind while conducting research. The precision of the results is driven by who the members of a sample are and how useful their opinions are. The quality of respondents in a sample is essential for the results received for research and not the quantity. If a researcher seeks to understand whether a product feature will work well with their target market, he/she can conduct survey research with a group of market experts for that product or technology.
- **Zero-in on a survey method:** Many qualitative and quantitative research methods

can be discussed and decided. Focus groups, online interviews, surveys, polls, questionnaires, etc. can be carried out with the pre-decided sample of individuals.

- **Design the questionnaire:** What will the content of the survey be? A researcher is required to answer this question to be able to design it effectively. What will the content of the cover letter be? Or what are the survey questions of this questionnaire? Understand the target market thoroughly to create a questionnaire that targets a sample to gain insights about a survey research topic.
- **Send out surveys and analyze results:** Once the researcher decides on which questions to include in a study, they can send it across to the selected sample. Answers obtained for this survey can be analysed to make product-related or marketing-related decisions.

2.5 CHAPTER SUMMARY

The method of data collection to be used for the study, the researcher should keep in mind two types of data viz., primary and secondary. The primary data are those which are collected afresh and for the first time, and thus happen to be original in character. The secondary data, on the other hand, are those which have already been collected by someone else and which have already been passed through the statistical process. Identify issues and opportunities for collecting data: Every tool for collecting data has its own pros and cons. Thus, for deciding the best method, it is important to identify issues and opportunities for collecting data according to the method. It might be helpful to engage in a pilot study to review our tools and sample size. We collect primary data during the course of doing experiments in experimental research but in case we do research of the descriptive type and perform surveys, whether sample surveys or census surveys, then we can obtain primary data either through observation or through direct communication with respondents in one form or another or through personal interviews. Secondary data means data that are already available i.e., they refer to the data which have already been collected and analysed by someone else. When the researcher utilizes secondary data, then he has to look into various sources from where he can obtain them. The case study method is a very popular form of qualitative analysis and involves a careful and complete observation of a social unit, be that unit a person, a family, an institution, a cultural group or even the entire community. Survey Research is defined as the process of conducting research using surveys that researchers send to survey respondents. The data collected from surveys is then statistically analysed to draw meaningful research conclusions

2.6 REVIEW QUESTIONS

SHORT ANSWER TYPE QUESTIONS

1. **Explain collection of primary data.**
2. **Write some other methods of data collection.**
3. **What do you understand by collection of secondary data?**
4. **Enlist the selection of appropriate method for data collection.**
5. **Name the five stages of survey research design.**

NOTES 

LONG ANSWER TYPE QUESTIONS

1. **What are the characteristics of case study method?**
2. **What is the process of implementing survey research methods?**
3. **What is the reasons to conduct survey research?**
4. **Briefly explain survey research scales.**
5. **What are the benefits of survey research?**

2.7 MULTIPLE CHOICE QUESTIONS

1. **The _____ are those which are collected afresh and for the first time, and thus happen to be original in character.**
   a. Primary Data
   b. Secondary Data
   c. Both of the above
   d. None of the above
2. **The _____, on the other hand, are those which have already been collected by someone else and which have already been passed through the statistical process.**
   a. Primary Data
   b. Secondary Data
   c. Both of the above
   d. None of the above
3. **The _____ is the most commonly used method especially in studies relating to behavioral sciences.**
   a. Personal interviews
   b. Telephone interviews
   c. observation method
   d. Interview Method
4. **The _____ of collecting data involves the presentation of oral-verbal stimuli and reply in terms of oral-verbal responses.**
   a. Personal interviews
   b. Telephone interviews
   c. observation method
   d. Interview Method
5. **_____ are usually postal-sized cards that are used by dealers of consumer durables to collect information regarding their products.**
   a. Warranty cards
   b. Distributor or store audits
   c. Pantry audits
   d. None of the above

NOTES 

6. _____ are performed by distributors as well as manufacturers through their salesmen at regular intervals.

a. Warranty cards
b. Distributor or store audits
c. Pantry audits
d. None of the above

7. The _____ technique is used to estimate the consumption of the basket of goods at the consumer level.

a. Warranty cards
b. Distributor or store audits
c. Pantry audits
d. None of the above

8. _____ survey research conducted over the telephone (CATI) can be useful in collecting data from a more extensive section of the target population

a. Face-to-face
b. Phone
c. Online/ Email
d. All of the above

9. A _____ associates numbers with variables for mere naming or labelling, and the numbers usually have no other relevance.

a. Interval Scale
b. Ratio Scale
c. Ordinal Scale
d. Nominal Scale

10. The _____ has an innate order within the variables along with labels.

a. Ordinal Scale
b. Interval Scale
c. Ratio Scale
d. Nominal Scale

♦♦♦♦♦

UNIT III

SAMPLING FUNDAMENTALS

STRUCTURE

NOTES 

3.1 LEARNING OBJECTIVE

After learning this unit students will be able to:

- Understand the Sampling Fundamentals.
- Understand the Types of Sampling.
- Understand the Sample Design Process.
- Understand the Errors in Sampling.

3.2 INTRODUCTION TO SAMPLING FUNDAMENTALS

In order to carry out a research study, you have to first acquire relevant information on the subject. In other words, you have to collect data. This data is required to test your 'hypotheses or generalizations that you have made for the time being. The process of selection demands thorough understanding of the concept of population, sample and various sampling techniques. In this Unit, we shall familiarize you with the concepts of sample and population. We shall also discuss the characteristics of a good sample and the various methods of sampling. Sampling has been an age-old practice in everyday life. Whenever we want to buy a huge quantity of a commodity, we decide about the total lot by simply examining a small fraction of it. It has been established that the sample survey if planned properly, can give very precise information. Since in surveys a part of the population is only surveyed and inference is drawn about the whole population, the results likely to be different from the population values. But the advantage with the sample survey is that this type of error can be measured and controlled and it can be eliminated to great extent by employing properly trained persons in surveys. The other advantage of sample surveys is that it is less time consuming and involves less cost. Usually, the population is too large for the researcher to attempt to survey all of its members. A small, but carefully chosen sample can be used to represent the population. The sample reflects the characteristics of the population from which it is drawn.

SAMPLING: MEANING AND CONCEPTS

Meaning of Sampling

According to Levin and Rubin, statisticians use the word, population, to refer not only to people, but to all items that have been chosen for study. They use the word, sample, to describe a portion chosen from the population. According to Croach and Housden, a sample is a limited number taken from a large group for testing and analysis, on the assumption that the sample can be taken as representative for the whole group. According to Boyce, sampling makes an estimate about some of the characteristics of a population. To sample is to make a judgment or a decision about something after experiencing just part of it.

Concepts in Sampling

For clarity and brevity, some concepts and preliminaries of sampling theory, which are used in the study material, are discussed below.

- **Sampling Units and Population:** a unit may be taken as a well-defined and identifiable element or a group of elements on which observations can be made.

NOTES 

The aggregate of these units is termed as population and the population is said to be finite, if the units are countable. The population is sub-divided into suitable small units known as sampling units for the purpose of sampling. Sampling units may consist of one or more elementary units and each elementary unit belongs to one and one sampling unit.

- **Sampling Frame:** a sampling frame is a list of sampling units with identification particulars indicating the location of the sampling units. A sampling frame represents the population under investigation, and it is the base of drawing a sample. As far as possible, it should be up-to–date, i.e., free from omissions and duplications.
- **Sample:** a fraction of the population is said to constitute a sample. The number of units included in the sample is known as the size of the sample.
- **Sampling Fraction:** the ratio of the sample size, n, to the population size, N, is known as sampling fraction and it is denoted by (n / N).
- **Sampling Procedure/Method:** this is the method of selecting a sample from a population.
- **Census:** this denotes all the elements or units of a population which are used to explain the features of population. It usually refers to complete enumeration of all persons in the population.
- **Population Parameter and Sample Estimator:** any function of the values of units in the population, such as population mean or population variance, is termed a population parameter. There can only be one set of values for a population and the population values are treated as constant. However, the function of the values of the units in the sample, such as sample mean and sample variance, is known as a statistic. The value of the mean and variance differ from sample to sample and, therefore, it is a random variable.

NEEDS OF SAMPLING IN RESEARCH METHODOLOGY

- Sampling is used in practice for a variety of reasons such as:
- Sampling can save time and money. A sample study is usually less expensive than a census study and produces results at a relatively faster speed.
- Sampling may enable more accurate measurements for a sample study is generally conducted by trained and experienced investigators.
- Sampling remains the only way when population contains infinitely many members.
- Sampling remains the only choice when a test involves the destruction of the item under study.
- Sampling usually enables to estimate the sampling errors and, thus, assists in obtaining information concerning some characteristic of the population.

ADVANTAGES OF SAMPLING

Some of the key advantages of sampling are:

- It costs less
- Takes less time
- Data are acquired quickly
- Fewer mistakes are likely
- A more detailed study can be done.

3.3 TYPES OF SAMPLING

There are broadly two types of sampling: i) Probability sampling ii) non-probability sampling

1. **Probability Sampling**
   A probability sample is one in which each element of the population has a known, non-zero chance of being included in the sample. Probability methods include simple random sampling, systematic sampling, and stratified sampling.

   a. **Simple Random** Sample

      The random sample entails that each and every individual in a population has an equal chance of being included in the sample and that the selection of one individual is in no way dependent upon the selection of another person. The two popularly used methods in random sampling are

      - Draw of lottery
      - Using a random number table
      - In lottery draw, for example,if we have to select a sample of 25 students from a total of 600 students in a college, then we make separate slips of paper for 600 students and put them in a box and thoroughly mix them. After that, a person is asked to pick up one slip. Here, the probability of each of the student being selected in the sample is 1/600. This procedure is continued till the sample size is acquired.
      - Another method of simple random sampling is to use a random number table for drawing 25 students from a total of 600 students. The procedure for using a random number table follows.

        i. Number each element in the sample frame from 001 to 600.

        ii. Decide a random starting point in the table. Any point will do. Say second row in the second column (Appendix 1).

        iii. Look at the first digits at that point, because there are three digits in 600.

        iv. Then, if the number is less than 600, include it in the sample; if not then look for a number where the first three digits are less than 600.

        v. From that point you can move in any direction. Select only three-digit numbers that are less than 600, until you have 25 such numbers.

        Note: You can move in any direction in the random number table because every digit has been placed in the table at random. For example, here if we start from the second row in the third column, then,

NOTES 

the random numbers are: 31684; 09865; 14491; 34691, continuing till 25 samples are selected.

b. Systematic Random

Sample Designing a Systematic Random Sample is sometimes quite difficult and time consuming and therefore, Systematic Random Sample, like Simple Random Sample, also uses a list of all members of the population in its sampling frame. However, instead of using random numbers to select the sample elements, the researcher applies a skip interval to the list to produce a sample of the required size.

K**=**

K = skip interval

N = Universe size

n = Sample size

For example, if we have to select a sample of 100 persons from a universe of 1000 population, then the skip is 10. In this case one number between 1 and 10 has to be selected. Suppose 5 is selected, then the first sample would be 5th and the next one 15^{th}, 25^{th}, 35^{th}, 45^{th}, and so on. One of the advantages of this method is that it is more convenient than other methods and simple to design. Again, it is used with very large populations.

c. Stratified Random

Sample Sampling In Stratified Random Sampling, the target population of N units is first divided into k subpopulations of N1, N2, ..., Nk units. These populations are non-overlapping and together they comprise the whole population. So that N1 + N2 + ... + Nk = N

The sub-populations are called strata. The number in each stratum should be known. A sample is drawn from each stratum independently. The sample sizes within 'k' strata are denoted by n1, n2, ..., nk respectively. If the total sample size 'n' is to be drawn from the target population than n1 + n2 + ... + nk = n

If a simple random sample is drawn in each stratum, the whole procedure is described as stratified random sampling.

Stratified random sampling requires more than making a list of elements (and estimating the number of elements on the list). It also involves ordering that list by sub groups (or strata) and then, to do sampling randomly or systematically within those sub groups. This method of sampling is used for the following reasons.

- It can reduce the errors in the statistical estimates calculated from the sample.
- It allows you to create a sample that is exactly representative of the various sub groups in the population that you find to be of special interest.

For example, the selected village may have households of SC, ST, OBCs, Others, Minority. The village population first may be divided in to smaller sub groups of different sections of population (stratum) and, thus, the village sample may consist of households from each stratum so that sample may contain all the important characteristics of the village population. In the case of SRS, the

sample of all strata/ sub groups sometimes may not be included or covered adequately.

- This method helps in conducting and managing a large-scale survey to be conducted in a country like India. The agency conducting the survey may have field offices in different locations; each one can supervise the survey for a part of the population.
- The basic idea is that it sub-divides the heterogeneous population into homogeneous sub-populations. If each stratum is homogenous in itself, a precise estimate of any stratum mean can be obtained from a small sample, thus, saving a lot of time and cost.

There are two types of stratified samples.

A proportionate stratified sample selects the number of elements from each stratum so that the stratum sample size (n1, n2 ,.....…., nk) is proportional to their respective stratum population size (N1 , N2 , …….., Nk). Consider the following examples:

- A selected village may have households of SC (10%), ST (5%), OBCs (45%), Others (30%), Minority (10%). A village sample of 100 may constitute the households of various casts in the above proportion/ percentage so that the sample may contain all important characteristics of village population.
- Hospital patients are stratified according to age, dividing the population into those who are aged 50years or above, and, those who are under 50. If there are twice as many people aged 50 or above admitted to the hospital as those under 50, a proportionate stratified sample will include twice as many people aged 50 or above.

  A **disproportionate stratified sample** selects the number of elements from each stratum so that the stratum sample size is not proportional to the stratum population size. The most common reason for selecting this type of sample is when you want to study a relatively rare but important subpopulation, such as younger patients suffering from heart disease. Proportionate stratification may result in too few elements being selected so that little, if any, statistical analysis can be done. Consequently, even if these patients represent only 1% of the population, you might decide to make them 10% of the final sample. However, once we combine values of all stratums, the size of the higher selected proportion needs to be readjusted which is called weighted estimate.

d. Probability Proportion to Size (PPS) Sample

It has been observed that the elementary units of the population vary in size. Such ancillary information about the size of the unit can be utilized in selecting the sample so as to get better and efficient estimates of the population parameter. For example, villages with larger geographical area are likely to have larger area under food crops; therefore, in estimating the production, it would be desirable to adopt a sampling scheme in which villages are selected with probability proportional to geographical area. When units vary in their size and the variable under study is directly related with the size of the unit, the probabilities may be assigned proportional to the size of the unit.

NOTES

Probability Proportion to Size (PPS) Sampling assures higher probability of selection to sampling unit which are larger in size. This technique was initially used in estimation of crop production, fruits production etc because productivity is directly related with the size of field. In social science surveys also characteristics of village population is influenced by the size of population. The procedure of selecting the sample is described below.

Suppose you have to select 5 villages from the list of 10 using PPS sampling. First arrange all villages in ascending or descending order of population size as may be seen in column 2 of the table 1. Then, in the third column, find the cumulative sum of population size and in the fourth column, assign them range of serial numbers as shown below in the table.

Village population Size

S. No.	Village Population Size	Commulative Sum of Population Size	Comulative Population Size Interval
1.	**2**	**3**	**4**
1.	200	200	0001-0200
2.	250	450	0201-0450
3.	300	750	0451-0750
4.	350	1100	0751-1100
5.	400	1500	1101-1500
6.	450	1950	1501-1950
7.	500	2450	1951-2450
8.	550	3000	2451-3000
9.	600	3600	3001-3600
10.	650	4250	3601-4250
Total	**4250**		

Please notice that the total population of all villages in the target population is a four-digit number (4250). Therefore, initially, a random number in four digits, which is less than or equal to the total population of all villages (4250), is selected from the random number table. For example, it is 0331 which will correspond to serial number 2. Next random number is 4320; therefore, it may be discarded. The next number selected is 1296; therefore, it will correspond to serial number 5. The next random numbers may be 1553, 2402 and 3640 which will correspond to serial numbers 6, 8, and 10 respectively. In this way, selected villages will be serial numbers 2, 5, 6, 8, 10.

e. Cluster Sample

Cluster sampling is a sampling technique used when natural groupings are evident in a statistical population. It is often used in marketing research. In this technique, the total population is divided into these known groups (or clusters) and a sample of the groups is selected. Then the required information is collected from the elements within each selected group. This may be done for every element in these groups, or a sub sample of elements may be selected within each of these groups. The technique works best when most of the variation in the population is within the groups, not between them. Briefly, the procedure for selecting a cluster sample is given below.

- The population is divided into N groups, called clusters.
- The researcher randomly selects n clusters to include in the sample.
- The number of observations within each cluster is known: M = M1 + M2 + M3 + ... + MN
- Each element of the population can be assigned to one, and only one, cluster.

Cluster sampling should be used only when it is economically justified - when reduced costs can be used to overcome losses in precision. This is most likely to occur in the following situations.

- Constructing a complete list of population elements is difficult, costly, or impossible. For example, it may not be possible to list all elementary units of the populations, for example all households in village, block, etc. However, it would be possible to randomly select a subset of villages, blocks (stage 1 of cluster sampling) and, then, interview the head of family in a house of the selected cluster (stage 2).
- The population is concentrated in natural clusters (city blocks, schools, hospitals, etc.). For example, to conduct personal interviews of operating room nurses, it might make sense to randomly select a sample of hospitals (stage 1 of cluster sampling) and then interview all of the operating room nurses at that hospital. Using cluster sampling, the interviewer could conduct many interviews in a single day at a single hospital. Simple random sampling, in contrast, might require the interviewer to spend all day travelling to conduct a single interview at a single hospital.

As discussed above, in the cluster sampling method, the primary selecting unit is not a household, rather a natural cluster of households, viz., hamlets in villages, or, created clusters, viz., schools, malls, etc., may be decided. The first list of clusters may be selected using the SRS or the PPS sampling techniques. Then, from each selected cluster, all units, or, some of the units, may be selected as per the required sample size using Stratified Random Sampling or the Systematic Random Sampling techniques. This sampling technique is quite popular in evaluation surveys in health – it is also called the 30 Cluster Sampling Technique. This is also a rapid method of data collection as the researcher can collect more data in less time due to the decrease in transportation time as compared with other sampling techniques.

2. Non-Probability Sampling

A non-probability sample is one in which a case in a sample is chosen in such a manner that it gives you information for the sample itself and makes it possible to generalize the findings for the population with certain degree of precision. Such a sample is also called a purposive sample. This kind of sampling is primarily used to collect information on market surveys to know the attitude, opinion, behaviour, reactions of individuals. There are many types of non-probability samples, including snowball sampling, convenience, purposive/ judgment, quota sampling, etc.

a. Convenience Sample

The convenience sample is so called because it is relatively easy to obtain and contact. In this method the investigators are usually asked to select the

NOTES

people for the interview in accordance to the instructions from the researcher. The benefit of a convenience sample is that the interviewer can usually get interviews done quickly and cheaply. Convenience sampling is appropriate for exploratory research.

b. Judgments Sample

Sampling A judgment sample is similar to that of convenience sample. In a judgment sample, the researcher selects samples that are believed to represent the population. The selection of samples is based on the knowledge of the population and the characteristics which the sample is to represent. It is less costly and very useful for forecasting.

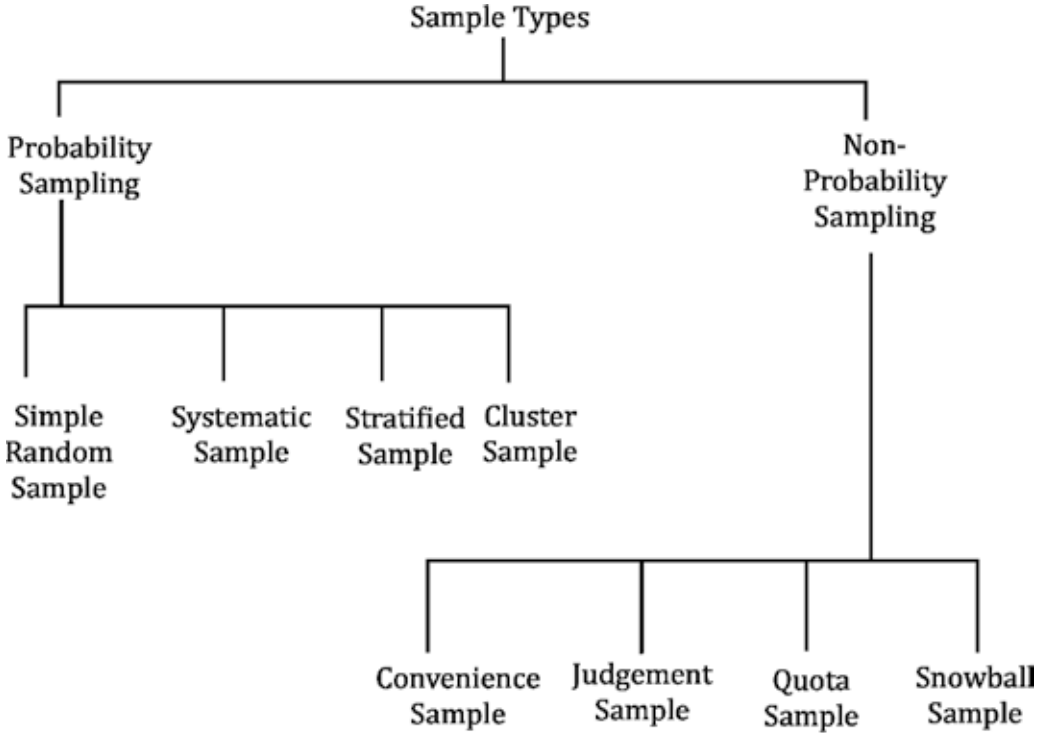
c. Quota Sample

Quota sampling is like stratified sampling. In quota sampling, the population is categorized into several strata which consist of an expected size, and the samples are considered to be important for the population they represent. The advantages of quota sample are that it involves a short time duration, is less costly, and gives moderate representation to a heterogeneous population.

d. Snowball Sample

This is one of the important types of non-probability sampling. In snowball sampling, the investigator encourages the respondents to give the names of other acquaintances and it continues growing in size and chains until the research purpose is achieved. It is also, therefore, known as networking, chain, or referred sampling method. It is very useful in the study of networking and is less costly

A comprehensive overview of the various types of sampling can be seen in figure 4.1



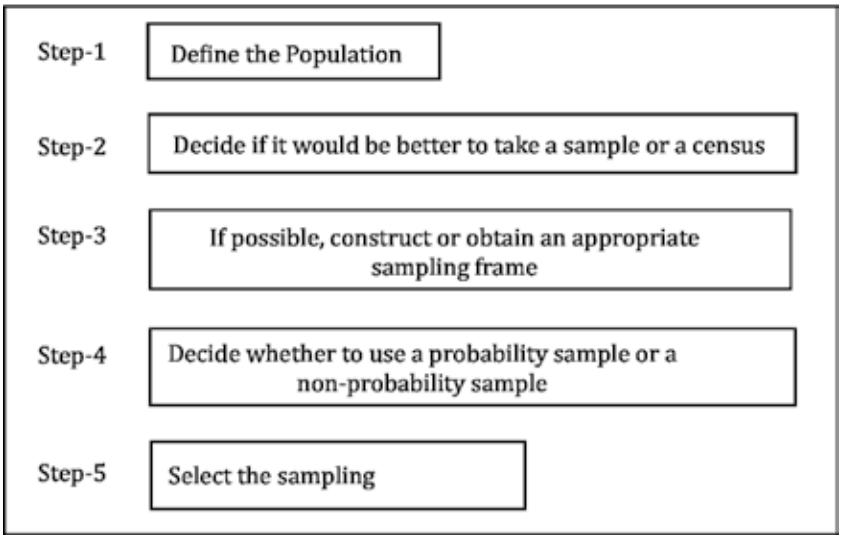
Types of Sample

NOTES

3.4 SAMPLE DESIGN PROCESS

The sample design process follows five steps as given in Box-1

Sample Design Process



Step-1: Define the Population

We use the word, population, frequently in our day-to-day conversations, for example, 'The population of India', or, 'The population of Punjab', or, 'The population of Kerela'. However, the meaning of the word, population, in research is different from what we use in day-to-day conversation. A research population may be defined as 'a clearly defined group of entities that have some characteristics in common'. This means the kind of people on whom we wish to base our research project. Sometimes, in research, we use the word, universe, instead of population. In a research project, our intention is to learn or infer something about the population. Whether we would use a sample or a population has to be clearly defined. For example, if we want to conduct a study on road safety, then the task of defining the population for a survey would be whether we should

- Interview only the people who drive two wheelers
- Interview only the people who drive four wheelers
- Interview the pedestrians
- Interview only who are hand rickshaw pullers or ride bi-cycle.

Therefore, judging a population is the starting of the sampling process

Step-2: Decide whether to take a sample or a census

After judging the population, the next step in the sampling process is to decide whether to take a sample or a population in your research project. In a census, usually every member of the population is interviewed. While in a sample method only selected members of the population are Sampling included. From the census we obtain data that are called population parameter, and from the sample we obtain statistics in a parameter. A parameter is a measurement of a characteristic of a population, while a sample statistic

NOTES 

is used as an estimate of a population parameter. Usually, in individual research, we use census when the population is small, and sample when the population is large.

Step-4: Decide whether to use a probability sample or a non-probability sample

The fourth step in the sampling process is whether to use probability sampling or non-probability sampling. In a probability sample, the sample elements are chosen by random selection, while in nonprobability sampling, each sample element is chosen according to whether the researcher decides that it should be included or not.

Step-5: Select the sampling method

Last, but not least in the sampling process, is the selection of the sampling method. In the probability sampling method, the following four principal kinds of probability sampling are used: the simple random sample, the systematic sample, the stratified sample, and the cluster sample. The main non-probability sampling methods are the convenience sample, the judgment sample, the quota sample, and the snowball sample.

SAMPLING FRAME

A sampling frame is a listing of all the elements from which you will draw the sample. In the ideal situation, the sampling frame will include all elementary units in the target population. A list of employees in an organization can create a sampling frame that exactly matches the population of interest. You should try to ensure that the sampling frame has the following characteristics. • It is actually created from the target population. • It is as complete a list as possible of the elements in the population.

3.5 ERRORS IN SAMPLING

Many mistakes and errors in social science research happen because of misleading and biased sampling. A sample which does not represent the population is called a biased sample. According to Yule and Kendal, "Bias may be due to imperfect instruments, the personal qualities of the observer, defective techniques and other cases. Like experimental error, it is difficult to eliminate entirely, but usually may be reduced to relatively small dimensions by taking proper care." There are two types of errors such as sampling errors and non-sampling errors. These are discussed below:

1. **Sampling Error**

   By definition, when you have collected a sample from a population, you have less than complete information about the population. This, in turn, means that there is a chance that the sample statistics you calculate, (for example, the mean of a variable, a frequency distribution, etc.) may not be an unbiased estimate of the population parameter.

   The error in the sample estimate is not an intrinsic impediment to analysis. For probability samples, sampling theory allows you to calculate the expected amount of error given a particular sample size, sampling method, and the specific statistic of interest. In general terms, the sampling error for a statistic can be defined as:

$$\text{Standard error} = \sqrt{\frac{Variance}{n}} = \frac{\text{sd}}{\sqrt{\text{n}}}$$

Where n refers to the number of respondents (sample size).

As the sample size increases, the standard error of a statistic decreases; as Sampling the variance, or dispersion, of a statistic increase, so does its sampling error. Sampling error decreases rapidly as the sample size increases from a few hundred to about 1000 respondents. However, there is rarely any reason to select larger samples while comparing the increased cost of survey with reduction in sampling error (see 'Calculating the Sample Size', in next section). The formula for the standard error of a proportion is simple and easy to apply:

Standard error = $\sqrt{\frac{p \times (1-p)}{n}}$

Here, p represents the proportion of successes (favourable response, those who received the benefits), {q = (1-p)} represents the proportion of failures (those who did not receive the benefits), and n is the total number of respondents. The standard error of a statistic is greatest when p and (1-p) are equal, which occurs when each is 0.50, or 50%, of the sample.

2. Non-Sampling Error

Before discussing how to determine sample size, we will briefly review other sources of error in surveys. When you read a news article that reports the results of a national poll, the error in the estimates is always listed, derived, generally speaking, from Equation 6.2. However, experienced survey researchers know that errors due to other sources are typically greater than the error due to sampling alone. Following are some other types of errors.

- Measurement errors, caused by poorly written questions, poorly designed questionnaires, respondent errors in completing questionnaires, and so on.
- Non-response errors, caused because the respondents are not a representative subset of the population.
- Data coding errors, caused, by errors in coding and entering the data

Of these error sources, the first two are typically more severe. In mail surveys, non-response error is often the most serious problem.

There are two critical characteristics of these non-sampling errors. First, as mentioned above, their sum is often greater than the sampling error. Second, and more insidious, these errors are often impossible to estimate for any one survey, especially measurement and non-response errors. Consequently, using Equation 6.1 and Equation 6.2 to estimate the error in a statistic often provides a false sense of security. Experienced survey researchers take this fact into account by being more cautions in discussing survey results than the sampling error alone would indicate, and you should do the same. Ideally, the other sources of error would balance themselves out so that errors in one direction negate errors in the other directions, but you cannot assume that this is the case.

3.5 CHAPTER SUMMARY

According to Levin and Rubin, statisticians use the word, population, to refer not only to people, but to all items that have been chosen for study. They use the word, sample, to describe a portion chosen from the population. A probability sample is one in which each

NOTES

element of the population has a known, non-zero chance of being included in the sample. Sample Designing a Systematic Random Sample is sometimes quite difficult and time consuming and therefore, Systematic Random Sample, like Simple Random Sample, also uses a list of all members of the population in its sampling frame. sampling unit which are larger in size. This technique was initially used in estimation of crop production, fruits production etc because productivity is directly related with the size of field. A sampling frame is a listing of all the elements from which you will draw the sample. In the ideal situation, the sampling frame will include all elementary units in the target population. A list of employees in an organization can create a sampling frame that exactly matches the population of interest. Many mistakes and errors in social science research happen because of misleading and biased sampling. A sample which does not represent the population is called a biased sample. There are two critical characteristics of these non-sampling errors. First, as mentioned above, their sum is often greater than the sampling error. Second, and more insidious, these errors are often impossible to estimate for any one survey, especially measurement and non-response errors.

3.6 REVIEW QUESTIONS

SHORT ANSWER TYPE QUESTIONS

1. **Drive the meaning of sampling.**
2. **Explain the concepts in sampling**
3. **Explain sampling frame.**
4. **What are the advantages of sampling?**
5. **Write a short note on probability sampling.**

LONG ANSWER TYPE QUESTIONS

1. **Elaborate the types of stratified samples.**
2. **Write a brief note on non-probability sampling.**
3. **Sample design process follows five steps. Explain.**
4. **Enlist the needs of sampling in research methodology.**
5. **What do you understand by errors in sampling?**

3.7 MULTIPLE CHOICE QUESTION

1. **According to _____, a sample is a limited number taken from a large group for testing and analysis, on the assumption that the sample can be taken as representative for the whole group.**
   a. Levin and Rubin
   b. Croach and Housden
   c. Boyce
   d. None of the above
2. **In _____ a unit may be taken as a well-defined and identifiable element or a group of elements on which observations can be made.**

a. Sampling Frame
b. Sample
c. Sampling Units and Population
d. Census

3. In _____ a sampling frame is a list of sampling units with identification particulars indicating the location of the sampling units.
a. Sample
b. Sampling Frame
c. Sampling Units and Population
d. Census

4. In _____ a fraction of the population is said to constitute a sample. The number of units included in the sample is known as the size of the sample.
a. Sampling Frame
b. Sampling Units and Population
c. Census
d. Sample

5. In _____ this denotes all the elements or units of a population which are used to explain the features of population. It usually refers to complete enumeration of all persons in the population.
a. Sampling Frame
b. Sampling Units and Population
c. Census
d. Sample

6. A _____ is one in which each element of the population has a known, non-zero chance of being included in the sample.
a. Probability Sample
b. Non-Probability
c. Both of the above
d. None of the above

7. The _____ entails that each and every individual in a population has an equal chance of being included in the sample and that the selection of one individual is in no way dependent upon the selection of another person.
a. Systematic Random
b. Stratified Random
c. Cluster Sample
d. Random sample

8. _____ is a sampling technique used when natural groupings are evident in a statistical population.
a. Systematic Random
b. Cluster Sample

NOTES 

c. Stratified Random
d. random sample

9. A _____ is one in which a case in a sample is chosen in such a manner that it gives you information for the sample itself and makes it possible to generalize the findings for the population with certain degree of precision.

a. Probability Sample
b. Non-Probability
c. Both of the above
d. None of the above

10. In _____, the researcher selects samples that are believed to represent the population.

a. Quota Sample
b. Convenience Sample
c. Judgment Sample
d. Snowball Sample

♦♦♦♦♦

UNIT IV

MEASUREMENT AND SCALING TECHNIQUES

STRUCTURE

NOTES 

4.1 LEARNING OBJECTIVE

After learning this unit students will be able to:

- Understand the Measurement and Scaling Techniques.
- Understand the Measurement and Scaling & Issues in Attitude Measurement.
- Understand the Levels of Measurement Scales & Types of Scaling Techniques.
- Understand the Selection of an Appropriate Scaling Technique & Organizing Data: Frequency Distribution.
- Understand the Descriptive Statistics: Measures of Central Tendency & Inferential Statistics: Z Test – T Test.
- Understand the Analysis of Variance & Correlational Research.

4.2 INTRODUCTION TO MEASUREMENT AND SCALING TECHNIQUES

As we discussed earlier, the data consists of quantitative variables like price, income, sales etc., and qualitative variables like knowledge, performance, character etc. The qualitative information must be converted into numerical form for further analysis. This is possible through measurement and scaling techniques. A common feature of survey-based research is to have respondent's feelings, attitudes, opinions, etc. in some measurable form. For example, a bank manager may be interested in knowing the opinion of the customers about the services provided by the bank. Similarly, a fast-food company having a network in a city may be interested in assessing the quality and service provided by them. As a researcher you may be interested in knowing the attitude of the people towards the government announcement of a metro rail in Delhi. In this unit we will discuss the issues related to measurement, different levels of measurement scales, various types of scaling techniques and also selection of an appropriate scaling technique.

4.3 MEASUREMENT AND SCALING

Before we proceed further it will be worthwhile to understand the following two terms: (a) Measurement, and (b) Scaling.

a. **Measurement:** Measurement is the process of observing and recording the observations that are collected as part of research. The recording of the observations may be in terms of numbers or other symbols to characteristics of objects according to certain prescribed rules. The respondent's, characteristics are feelings, attitudes, opinions etc. For example, you may assign '1' for Male and '2' for Female respondents. In response to a question on whether he/she is using the ATM provided by a particular bank branch, the respondent may say 'yes' or 'no'. You may wish to assign the number '1' for the response yes and '2' for the response no. We assign numbers to these characteristics for two reasons. First, the numbers facilitate further statistical analysis of data obtained. Second, numbers facilitate the communication of measurement rules and results. The most important aspect of measurement is the specification of rules for assigning numbers to characteristics. The rules for assigning numbers should be standardised and applied uniformly. This must not change over time or objects.

NOTES

b. Scaling: Scaling is the assignment of objects to numbers or semantics according to a rule. In scaling, the objects are text statements, usually statements of attitude, opinion, or feeling. For example, consider a scale locating customers of a bank according to the characteristic "agreement to the satisfactory quality of service provided by the branch". Each customer interviewed may respond with a semantic like 'strongly agree', or 'somewhat agree', or 'somewhat disagree', or 'strongly disagree'. We may even assign each of the responses a number. For example, we may assign strongly agree as '1', agree as '2' disagree as '3', and strongly disagree as '4'. Therefore, each of the respondents may assign 1, 2, 3 or 4.

SOURCES OF ERROR IN MEASUREMENT

Measurement should be precise and unambiguous in an ideal research study. This objective, however, is often not met with in entirety. As such the researcher must be aware about the sources of error in measurement. The following are the possible sources of error in measurement.

- **Respondent:** At times the respondent may be reluctant to express strong negative feelings or it is just possible that he may have very little knowledge but may not admit his ignorance. All this reluctance is likely to result in an interview of 'guesses.' Transient factors like fatigue, boredom, anxiety, etc. may limit the ability of the respondent to respond accurately and fully.
- **Situation:** Situational factors may also come in the way of correct measurement. Any condition which places a strain on interview can have serious effects on the interviewer-respondent rapport. For instance, if someone else is present, he can distort responses by joining in or merely by being present. If the respondent feels that anonymity is not assured, he may be reluctant to express certain feelings.
- **Measurer:** The interviewer can distort responses by rewording or reordering questions. His behavior, style and looks may encourage or discourage certain replies from respondents. Careless mechanical processing may distort the findings. Errors may also creep in because of incorrect coding, faulty tabulation and/or statistical calculations, particularly in the data-analysis stage.
- **Instrument:** Error may arise because of the defective measuring instrument. The use of complex words, beyond the comprehension of the respondent, ambiguous meanings, poor printing, inadequate space for replies, response choice omissions, etc. are a few things that make the measuring instrument defective and may result in measurement errors. Another type of instrument deficiency is the poor sampling of the universe of items of concern. Researcher must know that correct measurement depends on successfully meeting all of the problems listed above. He must, to the extent possible, try to eliminate, neutralize or otherwise deal with all the possible sources of error so that the final results may not be contaminated.

TECHNIQUE OF DEVELOPING MEASUREMENT TOOLS

The technique of developing measurement tools involves a four-stage process, consisting of the following:

- Concept development

NOTES 

- Specification of concept dimensions
- Selection of indicators
- Formation of index

The first and foremost step is that of concept development which means that the researcher should arrive at an understanding of the major concepts pertaining to his study. This step of concept development is more apparent in theoretical studies than in the more pragmatic research, where the fundamental concepts are often already established. The second step requires the researcher to specify the dimensions of the concepts that he developed in the first stage. This task may either be accomplished by deduction i.e., by adopting a more or less intuitive approach or by empirical correlation of the individual dimensions with the total concept and/or the other concepts.

For instance, one may think of several dimensions such as product reputation, customer treatment, corporate leadership, concern for individuals, sense of social responsibility and so forth when one is thinking about the image of a certain company. Once the dimensions of a concept have been specified, the researcher must develop indicators for measuring each concept element. Indicators are specific questions, scales, or other devices by which respondent's knowledge, opinion, expectation, etc., are measured. As there is seldom a perfect measure of a concept, the researcher should consider several alternatives for the purpose. The use of more than one indicator gives stability to the scores and it also improves their validity.

The last step is that of combining the various indicators into an index, i.e., formation of an index. When we have several dimensions of a concept or different measurements of a dimension, we may need to combine them into a single index. One simple way for getting an overall index is to provide scale values to the responses and then sum up the corresponding scores. Such an overall index would provide a better measurement tool than a single indicator because of the fact that an "individual indicator has only a probability relation to what we really want to know." This way we must obtain an overall index for the various concepts concerning the research study.

4.4 ISSUES IN ATTITUDE MEASUREMENT

When a researcher is interested in measuring the attitudes, feelings or opinions of respondents he/she should be clear about the following:

- What is to be measured?
- Who is to be measured?
- The choices available in data collection techniques

The first issue that the researcher must consider is 'what is to be measured'? The definition of the problem, based on our judgments or prior research indicates the concept to be investigated. For example, we may be interested in measuring the performance of a fast-food company. We may require a precise definition of the concept on how it will be measured. Also, there may be more than one way that we can measure a particular concept. For example, in measuring the performance of a fast-food company we may use

a number of measures to indicate the performance of the company. We may use sales volume in terms of value of sales or number of customers or spread of network of the company as measures of performance. Further, the measurement of concepts requires assigning numbers to the attitudes, feelings or opinions. The key question here is that on what basis do we assign the numbers to the concept. For example, the task is to measure the agreement of customers of a fast-food company on the opinion of whether the food served by the company is tasty, we create five categories: (1) strongly agree, (2) agree, (3) undecided, (4) disagree, (5) strongly disagree. Then we may measure the response of respondents. Suppose if a respondent states 'disagree' with the statement that 'the food is tasty', the measurement is 4.

The second important issue in measurement is that, who is to be measured? That means who are the people we are interested in. The characteristics of the people such as age, sex, education, income, location, profession, etc. may have a bearing on the choice of measurement. The measurement procedure must be designed keeping in mind the characteristics of the respondents under consideration. The third issue in measurement is the choice of the data collection techniques. In Unit II, you have already learnt various methods of data collection. Normally, questionnaires are used for measuring attitudes, opinions or feelings.

4.5 LEVELS OF MEASUREMENT SCALES

The level of measurement refers to the relationship among the values that are assigned to the attributes, feelings or opinions for a variable. For example, the variable 'whether the taste of fast food is good' has a number of attributes, namely, very good, good, neither good nor bad, bad and very bad. For the purpose of analysing the results of this variable, we may assign the values 1, 2, 3, 4 and 5 to the five attributes respectively. The level of measurement describes the relationship among these five values. Here, we are simply using the numbers as shorter placeholders for the lengthier text terms. We don't mean that higher values mean 'more' of something or lower values mean 'less' of something. We don't assume that 'good' which has a value of 2 is twice of 'very good' which has a value of 1. We don't even assume that 'very good' which is assigned the value '1' has more preference than 'good' which is assigned the value '2'. We simply use the values as a shorter name for the attributes, opinions, or feelings. The assigned values of attributes allow the researcher more scope for further processing of data and statistical analysis.

Typically, there are four levels of measurement scales or methods of assigning numbers: (a) Nominal scale, (b) Ordinal scale, (c) Interval scale, and (d) Ratio scale.

a. **Nominal Scale** is the crudest among all measurement scales but it is also the simplest scale. In this scale the different scores on a measurement simply indicate different categories. The nominal scale does not express any values or relationships between variables. For example, labelling men as '1' and women as '2' which is the most common way of labelling gender for data recording purpose does not mean women are 'twice something or other' than men. Nor it suggests that men are somehow 'better' than women. Another example of nominal scale is to classify the respondent's income into three groups: the highest income as group 1. The middle income as group 2, and the low-income as group 3. The nominal

NOTES

scale is often referred to as a categorical scale. The assigned numbers have no arithmetic properties and act only as labels. The only statistical operation that can be performed on nominal scales is a frequency count. We cannot determine an average except mode. In designing and developing a questionnaire, it is important that the response categories must include all possible responses. In order to have an exhaustive number of responses, you might have to include a category such as 'others', 'uncertain', 'don't know', or 'can't remember' so that the respondents will not distort their information by forcing their responses in one of the categories provided. Also, you should be careful and be sure that the categories provided are mutually exclusive so that they do not overlap or get duplicated in any way.

b. Ordinal Scale involves the ranking of items along the continuum of the characteristic being scaled. In this scale, the items are classified according to whether they have more or less of a characteristic. For example, you may wish to ask the TV viewers to rank the TV channels according to their preference and the responses may look like this as given below:

TV Channel	Viewers Preferences
Doordarshan-1	1
Star Plus	2
NDTV News	3
Aaaj Tak TV	4

The main characteristic of the ordinal scale is that the categories have a logical or ordered relationship. This type of scale permits the measurement of degrees of difference, (that is, 'more' or 'less') but not the specific number of differences (that is, how much 'more' or 'less'). This scale is very common in marketing, satisfaction and attitudinal research. Another example is that a fast-food home delivery shop may wish to ask its customers:

How would you rate the service of our staff?
(10 Excellent (2) Very Good (3) Good (4) Poor (5) Worst

Suppose respondent X gave the response 'Excellent' and respondent Y gave the response 'Good', we may say that respondent X thought that the service provided better than respondent Y to be thought. But we don't know how much better and even we can't say that both respondents have the same understanding of what constitutes 'good service'. In marketing research, ordinal scales are used to measure relative attitudes, opinions, and preferences. Here we rank the attitudes, opinions and preferences from best to worst or from worst to best. However, the amount of difference between the ranks cannot be found out. Using ordinal scale data, we can perform statistical analysis like Median and Mode, but not the Mean.

c. Interval Scale is a scale in which the numbers are used to rank attributes such that numerically equal distances on the scale represent equal distance in the characteristic being measured. An interval scale contains all the information of an ordinal scale, but it also one allows to compare the difference/distance between attributes. For example, the difference between '1' and '2' is equal to the difference between '3' and '4'. Further, the difference between '2' and '4' is twice the difference between '1' and '2'. However, in an interval scale, the zero point is arbitrary and is not true zero. This, of course, has implications for the type of data

manipulation and analysis. We can carry out on data collected in this form. It is possible to add or subtract a constant to all of the scale values without affecting the form of the scale but one cannot multiply or divide the values. Measuring temperature is an example of interval scale. We cannot say 400 C is twice as hot as 200 C. The reason for this is that 00 C does not mean that there is no temperature, but a relative point on the Centigrade Scale. Due to lack of an absolute zero point, the interval scale does not allow the conclusion that 400 C is twice as hot as 200 C.

Interval scales may be either in numeric or semantic formats. The following are two more examples of interval scales one in numeric format and another in semantic format.

i. Example of Interval Scale in Numeric Format

Food supplied is:						Indicate your score on the concerned blank and circle the appropriate numner on each line.
Fresh	1	2	3	4	5	
Tastes good	1	2	3	4	5	
Value for money	1	2	3	4	5	
Attractive packaging	1	2	3	4	5	
Prompt time delivery	1	2	3	4	5	

ii. Example of Interval Scale in Semantic Format

Please indicate your views on the food supplied by XXX Fast Food Shop by scoring them on a five points scale from 1 to 5 (that is, 1=Excellent, 2=Very Good, 3=Good, 4=Poor, 5=Worst). Indicate your views by ticking the appropriate responses below:

Food supplied is:	**Excellent**	**Very Good**	**Good**	**Poor**	**Worst**
Fresh					
Tastes good					
Value for money					
Attractive packaging					
Prompt time delivery					

The interval scales allow the calculation of averages like Mean, Median and Mode and dispersion like Range and Standard Deviation.

d. Ratio Scale is the highest level of measurement scales. This has the properties of an interval scale together with a fixed (absolute) zero point. The absolute zero point allows us to construct a meaningful ratio. Examples of ratio scales include weights, lengths and times. In the marketing research, most counts are ratio scales. For example, the number of customers of a bank's ATM in the last three months is a ratio scale. This is because you can compare this with previous three months. Ratio scales permit the researcher to compare both differences in scores and relative magnitude of scores. For example, the difference between 10 and 15 minutes is the same as the difference between 25 and 30 minutes and 30 minutes is twice as long as 15 minutes. Most financial research that deals with rupee values utilizes ratio scales. However, for most behavioural research, interval scales are typically the highest form of measurement. Most statistical data analysis procedures do not distinguish between the interval and ratio properties of the measurement scales

NOTES

and it is sufficient to say that all the statistical operations that can be performed on interval scale can also be performed on ratio scales.

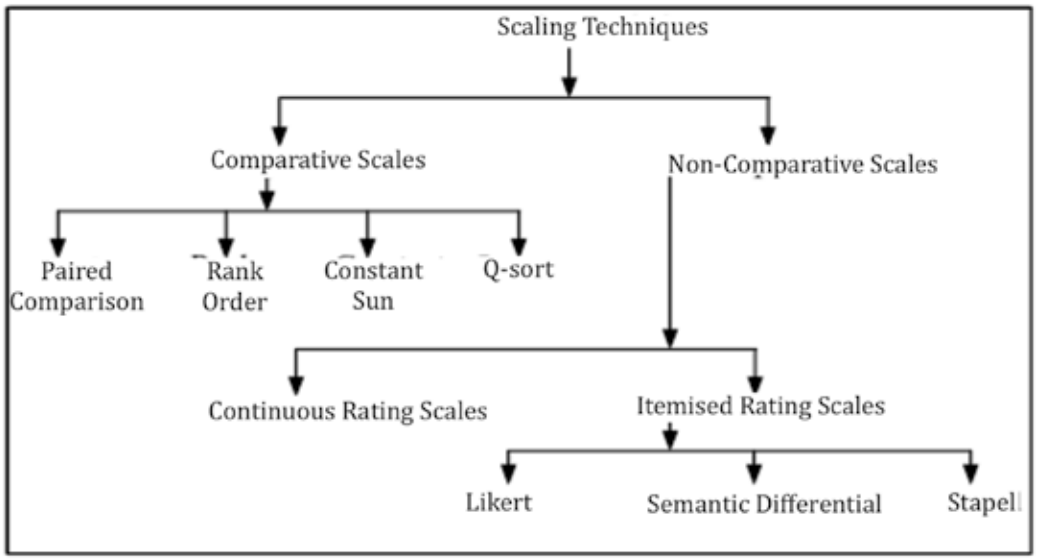
Now you must be wondering why you should know the level of measurement. Knowing the level of measurement helps you to decide on how to interpret the data. For example, when you know that a measure is nominal then you know that the numerical values are just short codes for longer textual names. Also, knowing the level of measurement helps you to decide what statistical analysis is appropriate on the values that were assigned. For example, if you know that a measure is nominal, then you would not need to find mean of the data values or perform a t-test on the data. (t-test will be discussed in Unit-16 in the course).

It is important to recognise that there is a hierarchy implied in the levels of measurement. At lower levels of measurement, assumptions tend to be less restrictive and data analyses tend to be less sensitive. At each level up the hierarchy, the current level includes all the qualities of the one below it and adds something new. In general, it is desirable to have a higher level of measurement (that is, interval or ratio) rather than a lower one (that is, nominal or ordinal).

4.6 TYPES OF SCALING TECHNIQUES

The various types of scaling techniques used in research can be classified into two categories: (a) comparative scales, and (b) non-comparative scales. In comparative scaling, the respondent is asked to compare one object with another. For example, the researcher can ask the respondents whether they prefer brand A or brand B of a detergent. On the other hand, in noncomparative scaling respondents need only evaluate a single object. Their evaluation is independent of the other object which the researcher is studying. Respondents using a non-comparative scale employ whatever rating standard seems appropriate to them. Non-comparative techniques consist of continuous and itemized rating scales. Figure 5.1 shows the classification of these scaling techniques.

Scaling Techniques



NOTES

1. Comparative Scales

The comparative scales can further be divided into the following four types of scaling techniques: (a) Paired Comparison Scale, (b) Rank Order Scale, (c) Constant Sum Scale, and (d) Q-sort Scale.

a. Paired Comparison Scale: This is a comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object (rate between two objects at a time) according to some criterion. The data obtained are ordinal in nature. For example, there are four types of cold drinks - Coke, Pepsi, Sprite, and Limca. The respondents can prefer Pepsi to Coke or Coke to Sprite, etc. In all we can have the following six comparisons.

Coke–Pepsi

Coke–Sprite

Coke–Limca

Pepsi–Sprite

Pepsi–Limca

Sprite–Limca

In general, with n brands we have paired comparisons. The following is the data recording format using the paired comparisons.

Brand	Coke	Pepsi	Sprite	Limca
Coke	-	√		
Pepsi		-		
Sprite	√	√		
Limca	√	√	√	-
No. of times preferred	2	3	1	0

A √ in a particular box means that the brand in that column was preferred over the brand in the corresponding row. In the above recording, Coke was preferred over Sprite, Coke over Limca, in this case the number of times coke preferred was 2 times. Similarly, Pepsi over Coke, Pepsi over Sprite, Pepsi over Limca, in this case Pepsi was 3 time preferred. Thus, the number of times a brand was preferred is obtained by summing the √ s in each column. The following table gives paired comparison of data (assumed) for four brands of cold drinks.

Brand	Coke	Pepsi	Sprite	Limca
Coke	-	0.90	0.64	0.14
Pepsi	0.10	-	0.32	0.02
Sprite	0.36	0.68	-	0.15
Limca	0.86	0.98	0.85	-

The entries in the boxes represent the proportion of respondents preferring 'column brand' and to 'row' brand. For example, 90% prefer Pepsi to Coke and only 10% prefer Coke to Pepsi, etc.

Paired comparison is useful when the number of brands is limited, since it requires direct comparison and overt choice. One of the disadvantages of paired

NOTES

comparison scale is violation of the assumption of transitivity may occur. For example, in our example (Table 5.1) the respondent preferred Coke 2 times, Pepsi 3 times, Sprite 1 time, and Limca 0 times. That means, preference-wise, Pepsi >Coke, Coke >Sprite, and Sprite >Limca. However, the number of times Sprite was preferred should not be that of Coke. In other words, if A>B and B >C then C >A should not be possible. Also, the order in which the objects are presented may bias the results. The number of items/brands for comparison should not be too many. As the number of items increases, the number of comparisons increases geometrically. If the number of comparisons is too large, the respondents may become fatigued and no longer be able to carefully discriminate among them. The other limitation of paired comparison is that this scale has little resemblance to the market situation, which involves selection from multiple alternatives. Also, respondents may prefer one item over certain others, but they may not like it in an absolute sense.

b. Rank Order Scale: This is another type of comparative scaling technique in which respondents are presented with several items simultaneously and asked to rank them in the order of priority. This is an ordinal scale that describes the favoured and unfavoured objects, but does not reveal the distance between the objects. For example, if you are interested in ranking the preference of some selected brands of cold drinks, you may use the following format for recording the responses.

Preference of cold drink brands using rank order scaling

Instructions: Rank the follwing brands of cold drinks in order of preferene. Began by picking out the one brand you like most and assign it a number 1. Then find the second most preferred brand and assign it a number 2. Continue this procedure until you have ranked all the brands of cold drinks in order of preference. The least preferred brand should be assigned a rank of 4. Also remember no two brands receive the same rank order.

Format:

Brand	Rank
(a) Coke	3
(b) Pepsi	1
(c) Limca	2
(d) Sprite	4

Like paired comparison, the rank order scale, is also comparative in nature. The resultant data in rank order is ordinal data. This method is more realistic in obtaining the responses and it yields better results when direct comparison is required between the given objects. The major disadvantage of this technique is that only ordinal data can be generated.

c. Constant Sum Scale: In this scale, the respondents are asked to allocate a constant sum of units such as points, rupees, or chips among a set of stimulus objects with respect to some criterion. For example, you may wish to determine how important the attributes of price, fragrance, packaging, cleaning power, and lather of a detergent are to consumers. Respondents might be asked to

divide a constant sum to indicate the relative importance of the attributes using the following format.

Importance of detergent attributes using a constant sum scale

Instructions: Between attributes of detergent please allocate 100 points among the attributes os that your allocation reflects the relative importance you attach to each attribute. The more points an attributes receives, the more important the attributes is. If an attribute is not at all important, assign it zero points. If an attribute is twice as important as some other attribute, it should receive twice as many points.

Format:

Attribute	Number of Points
(a) Price	50
(b) Fragrance	05
(c) Packaging	10
(d) Cleaning Power	30
(e) Lather	05
Total Points	100

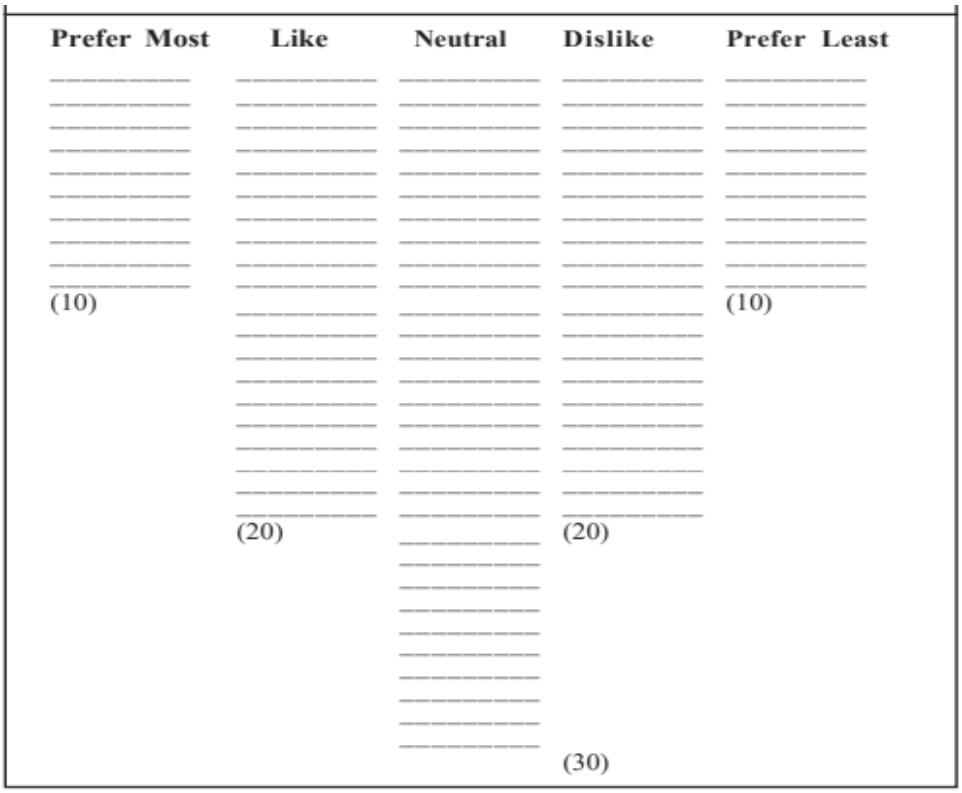
"If an attribute is assigned a higher number of points, it would indicate that the attribute is more important." From the above Table, the price of the detergent is the most important attribute for the consumers followed by cleaning power, packaging. Fragrance and lather are the two attributes that the consumers cared about the least but preferred equally." The advantage of this technique is saving time. However, there are two main disadvantages. The respondents may allocate more or fewer points than those specified. The second problem is rounding off error if too few attributes are used and the use of a large number of attributes may be too taxing on the respondent and cause confusion and fatigue.

d. Q-Sort Scale: This is a comparative scale that uses a rank order procedure to sort objects based on similarity with respect to some criterion. The important characteristic of this methodology is that it is more important to make comparisons among different responses of a respondent than the responses between different respondents. Therefore, it is a comparative method of scaling rather than an absolute rating scale. In this method the respondent is given statements in a large number for describing the characteristics of a product or a large number of brands of a product. For example, you may wish to determine the preference from among a large number of magazines. The following format shown in Table 5.5 may be given to a respondent to obtain the preferences.

Preference of Magazines Using Q-Sort Scale Procedure

Instructions: The bag given to you contain pictures of 90 magazines. Please choose 10 magazines you 'prefer most', 20 magazines you 'like', 30 magazines to which you are 'neutral (neither like nor dislike)', 20 magazines you 'dislike', and 10 magazines you 'prefer least'. Please list the sorted magazine names in the respective columns of the form provided to you. **Format:**

NOTES



Note that the number of responses to be sorted should not be less than 60 or not more than 140. A reasonable range is 60 to 90 responses that result in a normal or quasi-normal distribution. This method is faster and less tedious than paired comparison measures. It also forces the subject to conform to quotas at each point of scale so as to yield a quasi-normal distribution. The utility of Qsort in marketing research is to derive clusters of individuals who display similar preferences, thus representing unique market segments.

2. Non-Comparative Scales

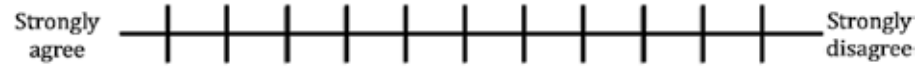
The non-comparative scaling techniques can be further divided into: (a) Continuous Rating Scale, and (b) Itemised Rating Scale.

a. Continuous Rating Scales

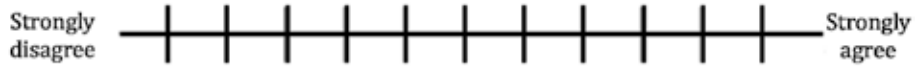
It is very simple and highly useful. In continuous rating scale, the respondent's rate the objects by placing a mark at the appropriate position on a continuous line that runs from one extreme of the criterion variable to the other. Examples of continuous rating scale are given below:

Question: How would you rate the TV advertisement as a guide for buying?

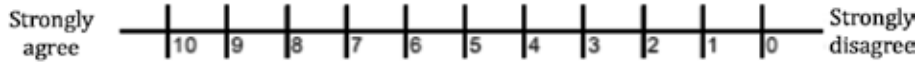
Scale Type A



Scale Type B



Scale Type C

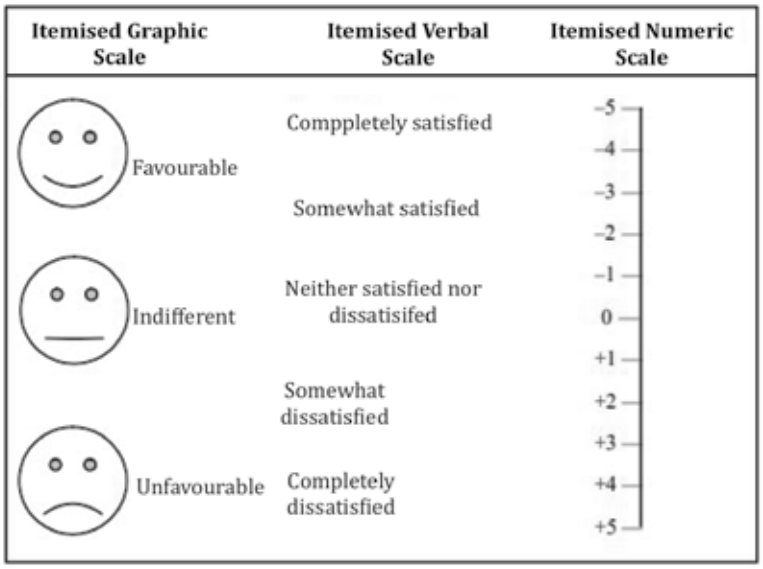


Scale Type D

When scale type A and B are used, the respondents score is determined either by dividing the line into as many categories as desired and assigning the respondent a score based on the category into which his/her mark falls, or by measuring distance, in millimeters, centimeters, or inches from either end of the scale. Whichever of the above continuous scale is used, the results are normally analysed as interval scaled.

b. Itemised Rating Scales

Itemised rating scale is a scale having numbers or brief descriptions associated with each category. The categories are ordered in terms of scale position and the respondents are required to select one of the limited number of categories that best describes the product, brand, company, or product attribute being rated. Itemised rating scales are widely used in marketing research. The itemised rating scales can be in the form of: (a) graphic, (b) verbal, or (c) numeric as shown below:



NOTES 

Some rating scales may have only two response categories such as: agree and disagree. Inclusion of more response categories provides the respondent more flexibility in the rating task. Consider the following questions:

1. How often do you visit the supermarket located in your area of residence?
   • Never, • Rarely, • Sometimes, • Often, • Very often
2. In your case how important is the price of brand X shoes when you buy them?
   • Very important, • Fairly important, • Neutral, • Not so important

Each of the above category scales is a more sensitive measure than a scale with only two responses since they provide more information. Wording is an extremely important factor in the usefulness of itemised scales. Table 5.6 shows some common wordings for categories used in itemised scales.

Some common words for categories used in Itemised Rating scales

Quality: Excellent Very Good	Good Good	Not decided Neither good nor bad	Poor Fair	Worst Poor
Importance: Very Important	Fairly important	Neatural	Not so important	Not at all important
Interest: Very interested	Somewhat interested	Neither interested nor disinterested	Somewhat uninterested	Not very interested
Satisfaction: Completely satisfied	Somewhat satisfied	Neither satisfied nor dissatisfied	Somewhat dissatisfied	Completely dissatisfied
Frequency: All of the time Very ofen	Very often Often	Often Sometimes	Sometimes Rarely	Hardly ever Never
Truth: Very true	Somewhat true	NOt very true	Not at all true	
Purchase Interest: Definitely will buy	Probably will buy	Probably will not buy	Definitely will not buy	
Level of Agreement: Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
Dependability: Completely dependable	Comewhat dependable	Not very dependable	Not at all dependable	

NOTES

Style: Very stylish	Somewhat stylish	Not very stylish	Completely unstylish	
Cost: Extremly expensive	Expensive	Neither expensive nor intexpensive	Slightly inexpensive	Very inexpensive
Ease of use: Very ease to use	Somewhat easy to use	Not very easy to use	Difficult to use	
Modernity: Very modern	Somewhat modern	Neither modern nor old-fashioned	Somewhat old fashioned	Very old fashioned
Alert: Very alert	Alert	Not alert	Not at all alert	

In this section we will discuss three itemised rating scales, namely (a) Likert scale, (b) Semantic Differential Scale, and (c) Stapel Scale

- **Likert Scale:** In business research, the Likert scale, developed by Rensis Likert, is extremely popular for measuring attitudes, because, the method is simple to administer. With the Likert scale, the respondents indicate their own attitudes by checking how strongly they agree or disagree with carefully worded statements that range from very positive to very negative towards the attitudinal object. Respondents generally choose from five alternatives (say strongly agree, agree, neither agree nor disagree, disagree, strongly disagree). Consider the following example of a study or measuring attitudes towards cricket.

	Strongly agree	Agree	Not sure	Disagree	Strongly disagree
It is more fun to play a tough, competitive cricket match than to play an easy one.	5	4	3	2	1

To measure the attitude, the researchers assign weights or scores to the alternative responses. In the above example the scores 5 to 1 are assigned to the responses. Strong agreement of the respondent indicates the most favourable attitudes on the statement, and the score 5 is assigned to it. On the other hand, strong disagreement of the respondent indicates the most unfavourable attitude on the statement, and the score 1 is assigned to it. If a negative statement towards the object is given, the corresponding scores would be reversed. In this case, the response 'strongly agree' will get a score of 1 and the response 'strongly disagree' will get a score of 5.

A Likert scale may include a number of items or statements. Each statement is assumed to represent an aspect of an attitudinal domain. For example, Table 5.7 shows the items in a Likert Scale to measure opinions on food products.

NOTES 

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
If the price of raw materials fall, firms too should reduce the price of the food products.	1	2	3	4	5
There should be uniform price though out the countrhy for food products	1	2	3	4	5
The food companies should concentrate more on keeping hygiene while manufacturing food products.	1	2	3	4	5
The expiry dates should be printed on the food products before they are delivered to consumers in the market.	1	2	3	4	5
There should be govbernment regulations on the firms in keeping acceptable quility and on the prices	1	2	3	4	5
Now-a-days must food companies are concerned only with profit making rather than taking care of quality.	1	2	3	4	5

Each respondent is asked to circle his opinion on a score against each statement. The final score for the respondent on the scale is the sum of their ratings for all the items. The very purpose of Likert's Scale is to ensure the final items evoke a wide response and discriminate among those with positive and negative attitudes. Items that are poor (because they lack clarity or elicit mixed response patterns) are detected from the final statement list. This will ensure us to discriminate between high positive scores and high negative scores. However, many business researchers do not follow this procedure and you may not be in a position to distinguish between high positive scores and high negative scores because all scores look alike. Hence a disadvantage of the Likert Scale is that it is difficult to

know what a single summated score means. Many patterns of response to the various statements can produce the same total score. The other disadvantage of Likert Scale is that it takes longer time to complete than other itemised rating scales because respondents have to read each statement. Despite the above disadvantages, this scale has several advantages. It is easy to construct, administer and use.

- **Semantic Differential Scale:** This is a seven-point rating scale with end points associated with bipolar labels (such as good and bad, complex and simple) that have semantic meaning. The Semantic Differential scale is used for a variety of purposes. It can be used to find whether a respondent has a positive or negative attitude towards an object. It has been widely used in comparing brands, products and company images. It has also been used to develop advertising and promotion strategies and in a new product development study. Look at the following Table, for examples of Semantic Differential Scale.

Examples of Semantic Differential Scale

Modern	-	-	-	-	-	-	-	Old-fashioned
Good	-	-	-	-	-	-	-	Bad
Clean	-	-	-	-	-	-	-	Dirty
Important	-	-	-	-	-	-	-	Unimportant
Expensive	-	-	-	-	-	-	-	Inexpensive
Useful	-	-	-	-	-	-	-	Useless
Stront	-	-	-	-	-	-	-	Weak
Quick	-	-	-	-	-	-	-	Slow

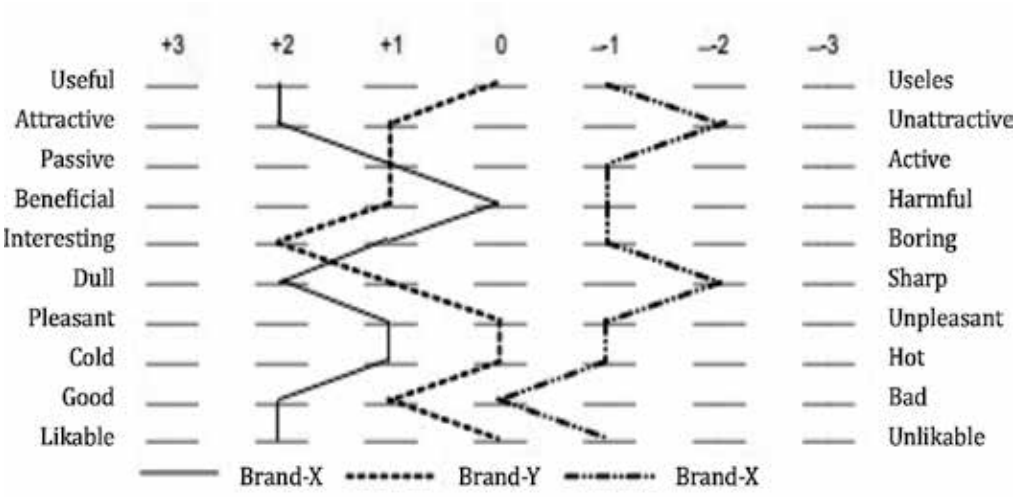
In the Semantic Differential scale only, extremes have names. The extreme points represent the bipolar adjectives with the central category representing the neutral position. The in between categories have blank spaces. A weight is assigned to each position on the scale. The weights can be such as +3, +2, +1, 0, –1, –2, –3 or 7,6,5,4,3,2,1. The following is an example of Semantic Differential Scale to study the experience of using a particular brand of body lotion.

In my experience, the use body lotion of Brand-X was:

	+3	+2	+1	0	-1	-2	-3	
Useful	-	-	-	-	-	-	-	Useless
Attractive	-	-	-	-	-	-	-	Unattractive
Passive	-	-	-	-	-	-	-	Active
Baneficial	-	-	-	-	-	-	-	Harmful
Interesting	-	-	-	-	-	-	-	Boring
Dull	-	-	-	-	-	-	-	Sharp
Pleasant	-	-	-	-	-	-	-	Unpleasant
Cold	-	-	-	-	-	-	-	Hot
Good	-	-	-	-	-	-	-	Bad
Likable	-	-	-	-	-	-	-	Unlikable

NOTES

In the semantic Differential scale, the phrases used to describe the object form a basis for attitude formation in the form of positive and negative phrases. The negative phrase is sometimes put on the left side of the scale and sometimes on the right side. This is done to prevent a respondent with a positive attitude from simply checking the left side and a respondent with a negative attitude checking on the right side without reading the description of the words. The respondents are asked to check the individual cells depending on the attitude. Then one could arrive at the average scores for comparisons of different objects. The following Figure shows the experiences of 100 consumers on 3 brands of body lotion.



In the above example, first the individual respondent scores for each dimension are obtained and then the average scores of all 100 respondents, for each dimension and for each brand were plotted graphically. The maximum score possible for each brand is + 30 and the minimum score possible for each brand is –30. Brand-X has score +14. Brand-Y has score +7, and Brand-Z has score –11. From the scale we can identify which phrase needs improvement for each Brand. For example, Brand-X needs to be improved upon benefits and Brand-Y on pleasantness, coldness and likeability. Brand Z needs to be improved on all the attributes.

- **Staple Scale:** The Stapel scale was originally developed to measure the direction and intensity of an attitude simultaneously. Modern versions of the Stapel scale place a single adjective as a substitute for the Semantic differential when it is difficult to create pairs of bipolar adjectives. The modified Stapel scale places a single adjective in the centre of an even number of numerical values (say, +3, +2, +1, 0, –1, –2, –3). This scale measures how close to or how distant from the adjective a given stimulus is perceived to be. The following is an example of a Staple scale.

NOTES

Instructions: Select a plus number for words that you think describe personnel banking of a bank accurately. The more accurately you think the word describes the bank, the larger the plus number you should choolse. Select a minus number for words you think do not describe the bak accurately. The less accurately you think the word describes the bank, the larger the minus number you should choose.

Format:

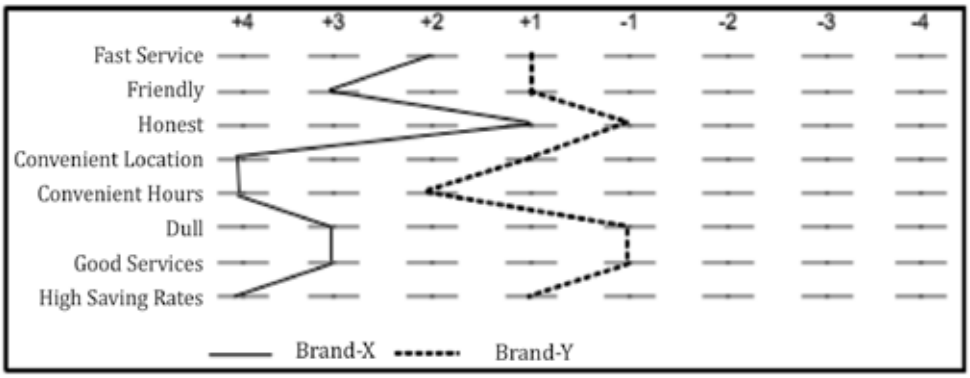
+5	+5
+4	+4
+3	+3
+2	+2
+1	+1
Friendly Personnel	Competitive Loan Rates
-1	-1
-2	-2
-3	-3
-4	-4
-5	-5

The following format shows an example of Stapel scale that illustrates respondents' description on personnel banking of a bank.

	+4	+3	+2	+1	-1	-2	-3	-4
Fast Services	-	-	-	-	-	-	-	-
Friendly	-	-	-	-	-	-	-	-
Honest	-	-	-	-	-	-	-	-
Convenient Location	-	-	-	-	-	-	-	-
Convenient Hours	-	-	-	-	-	-	-	-
Dull	-	-	-	-	-	-	-	-
Good Services	-	-	-	-	-	-	-	-
High Saving Rates	-	-	-	-	-	-	-	-

Each respondent is asked to circle his opinion on a score against each phrase that describes the object. The final score of the respondent on a scale is the sum of their ratings for all the items. Also, the average score for each phrase is obtained by totaling the final score of all the respondents for that phrase divided by the number of respondents of the phrase. The following Figure shows the opinions of 100 respondents on two banks.

NOTES



In the above example first the individual respondent's scores for each phrase that describes the selected bank are obtained and then the average scores of all 100 respondents for each phrase are plotted graphically. The maximum score possible for each bank is +32 and the minimum possible score for each brand is –32. In the example, Bank-X has score +24, and Bank-Y has score +3. From the scale we can identify which phrase needs improvement for each Bank. The advantages and disadvantages of the Stapel scale are very similar to those for the Semantic differential scale. However, the Stapel scale tends to be easier to construct and administer, especially over telephone, since the Stapel scale does not call for the bipolar adjectives as does the Semantic differential scale. However, research on comparing the Stapel scale with Semantic differential scale suggests that the results of both the scales are largely the same.

4.7 SELECTION OF AN APPROPRIATE SCALING TECHNIQUE

In this unit, so far, you have learnt some of the important scaling techniques that are frequently used in attitudinal research for the measurement of attitudes. Each of these techniques has some advantages and disadvantages. Now you may ask which technique is more appropriate to use to measure attitudes. Virtually any technique can be used to measure the attitudes. But at the same time all techniques are not suitable for all purposes. As a general rule, you should use a scaling technique that will yield the highest level of information feasible in a given situation. Also, if possible, the technique should permit you the use of a variety of statistical analysis. A number of issues decide the choice of scaling technique. Some significant issues are:

- **Problem Definition and Statistical Analysis:** The Choice between ranking, sorting, or rating techniques is determined by the problem definition and the type of statistical analysis likely to be performed. For example, ranking provides only ordinal data that limits the use of statistical techniques.
- **The Choice between Comparative and Non-comparative Scales:** Sometimes it is better to use a comparative scale rather than a non-comparative scale. Consider the following example: How satisfied you are with the brand- X detergent that you are presently using?

Completely satisfied	*Somewhat satisfied*	*Neither satisfied nor dissatisifed*	*Somewhat dissatisfied*	*Completely dissatisfied*

This is a non-comparative scale since it deals with a single concept (the brand of a detergent). On the other hand, a comparative scale asks a respondent to rate a concept. For example, you may ask: *Which one of the following brands of detergent you prefer?*

Brand-X Brand-Y

In this example you are comparing one brand of detergent with another brand. Therefore, in many situations, comparative scaling presents 'the ideal situation' as a reference for comparison with actual situation.

- **Type of Category Labels:** We have discussed different types of category labels used in constructing measurement scales such as verbal categories and numeric categories. Many researchers use verbal categories since they believe that these categories are understood well by the respondents. The maturity and the education level of the respondents influences this decision.
- **Number of Categories:** While there is no single, optimal number of categories, traditional guidelines suggest that there should be between five and nine categories. Also, if a neutral or indifferent scale response is possible for at least some of the respondents, an odd number of categories should be used. However, the researcher must determine the number of meaningful positions that are best suited for a specific problem.
- **Balanced versus Unbalanced Scale:** In general, the scale should be balanced to obtain objective data.
- **Forced versus Nonforced Categories:** In situations where the respondents are expected to have no opinion, the accuracy of data may be improved by a non-forced scale that provides a 'no opinion' category.

4.8 ORGANIZING DATA: FREQUENCY DISTRIBUTION

A frequency distribution shows the frequency of repeated items in a graphical form or tabular form. It gives a visual display of the frequency of items or shows the number of times they occurred. Let's learn about frequency distribution in this article in detail.

What is Frequency Distribution?

Frequency distribution is used to organize the collected data in table form. The data could be marks scored by students, temperatures of different towns, points scored in a volleyball match, etc. After data collection, we have to show data in a meaningful manner for better understanding. Organize the data in such a way that all its features are summarized in a table. This is known as frequency distribution.

Let's consider an example to understand this better. The following are the scores of 10 students in the G.K. quiz released by Mr. Chris 15, 17, 20, 15, 20, 17, 17, 14, 14, 20. Let's represent this data in frequency distribution and find out the number of students who got the same marks.

NOTES 

Quiz Marks	No. of Students
15	2
17	3
20	3
14	2

We can see that all the collected data is organized under the column quiz marks and the number of students. This makes it easier to understand the given information and we can see that the number of students who obtained the same marks. Thus, frequency distribution in statistics helps us to organize the data in an easy way to understand its features at a glance.

FREQUENCY DISTRIBUTION GRAPHS

There is another way to show data that is in the form of graphs and it can be done by using a frequency distribution graph. The graphs help us to understand the collected data in an easy way. The graphical representation of a frequency distribution can be shown using the following:

- **Bar Graphs:** Bar graphs represent data using rectangular bars of uniform width along with equal spacing between the rectangular bars.
- **Histograms:** A histogram is a graphical presentation of data using rectangular bars of different heights. In a histogram, there is no space between the rectangular bars.
- **Pie Chart:** A pie chart is a type of graph that visually displays data in a circular chart. It records data in a circular manner and then it is further divided into sectors that show a particular part of data out of the whole part.
- **Frequency Polygon**: A frequency polygon is drawn by joining the mid-points of the bars in a histogram.

TYPES OF FREQUENCY DISTRIBUTION

There are four types of frequency distribution under statistics which are explained below:

- **Ungrouped frequency distribution:** It shows the frequency of an item in each separate data value rather than groups of data values.
- **Grouped frequency distribution:** In this type, the data is arranged and separated into groups called class intervals. The frequency of data belonging to each class interval is noted in a frequency distribution table. The grouped frequency table shows the distribution of frequencies in class intervals.
- **Relative frequency distribution:** It tells the proportion of the total number of observations associated with each category.
- **Cumulative frequency distribution:** It is the sum of the first frequency and all frequencies below it in a frequency distribution. You have to add a value with the next value then add the sum with the next value again and so on till the last. The last cumulative frequency will be the total sum of all frequencies.

NOTES

FREQUENCY DISTRIBUTION TABLE

A frequency distribution table is a chart that shows the frequency of each of the items in a data set. Let's consider an example to understand how to make a frequency distribution table using tally marks. A jar containing beads of different colors- red, green, blue, black, red, green, blue, yellow, red, red, green, green, green, yellow, red, green, yellow. To know the exact number of beads of each particular color, we need to classify the beads into categories. An easy way to find the number of beads of each color is to use tally marks. Pick the beads one by one and enter the tally marks in the respective row and column. Then, indicate the frequency for each item in the table.

Category	Telly Marks	frequency
Red	IIII	5
Green	IIII I	6
Blue	II	2
Black	I	1
yellow	III	3

Thus, the table so obtained is called a frequency distribution table.

TYPES OF FREQUENCY DISTRIBUTION TABLE

There are two types of frequency distribution tables: Grouped and ungrouped frequency distribution tables.

Grouped Frequency Distribution Table: To arrange a large number of observations or data, we use grouped frequency distribution table. In this, we form class intervals to tally the frequency for the data that belongs to that particular class interval.

For example, Marks obtained by 20 students in the test are as follows. 5, 10, 20, 15, 5, 20, 20, 15, 15, 15, 10, 10, 10, 20, 15, 5, 18, 18, 18, 18. To arrange the data in grouped table we have to make class intervals. Thus, we will make class intervals of marks like 0 – 5, 6 – 10, and so on. Given below table shows two columns one is of class intervals (marks obtained in test) and the second is of frequency (no. of students). In this, we have not used tally marks as we counted the marks directly.

Marks obtained in Test (class intervals)	No. of Students (Frequency)
0 – 5	3
6 – 10	4
11 – 15	5
16 – 20	8
Total	20

Ungrouped Frequency Distribution Table: In the ungrouped frequency distribution table, we don't make class intervals, we write the accurate frequency of individual data. Considering the above example, the ungrouped table will be like this. Given below table shows two columns: one is of marks obtained in the test and the second is of frequency (no. of students).

NOTES

Marks obtained in Test	No. of Students
5	3
10	4
15	5
18	4
20	4

Important Notes:

Following are the important points related to frequency distribution.

- Figures or numbers collected for some definite purpose is called data.
- Frequency is the value in numbers that shows how often a particular item occurs in the given data set.
- There are two types of frequency table - Grouped Frequency Distribution and Ungrouped Frequency Distribution.
- Data can be shown using graphs like histograms, bar graphs, frequency polygons, and so on.

4.9 DESCRIPTIVE STATISTICS: MEASURES OF CENTRAL TENDENCY

A measure of central tendency (also referred to as measures of centre or central location) is a summary measure that attempts to describe a whole set of data with a single value that represents the middle or centre of its distribution. There are three main measures of central tendency: the mode, the median and the mean. Each of these measures describes a different indication of the typical or central value in the distribution.

What is the mode?

The mode is the most commonly occurring value in a distribution. Consider this dataset showing the retirement age of 11 people, in whole years:

54, 54, 54, 55, 56, 57, 57, 58, 58, 60, 60

This table shows a simple frequency distribution of the retirement age data:

Age	Frequency
54	3
55	1
56	1
57	2
58	2
59	2

The most commonly occurring value is 54, therefore the mode of this distribution is 54 years.

Advantage of the mode:

The mode has an advantage over the median and the mean as it can be found for both numerical and categorical (non-numerical) data.

Limitations of the mode:

The are some limitations to using the mode. In some distributions, the mode may not reflect the centre of the distribution very well. When the distribution of retirement age is ordered from lowest to highest value, it is easy to see that the centre of the distribution is 57 years, but the mode is lower, at 54 years.

54, 54, 54, 55, 56, 57, 57, 58, 58, 60, 60

It is also possible for there to be more than one mode for the same distribution of data, (bi-modal, or multi-modal). The presence of more than one mode can limit the ability of the mode in describing the centre or typical value of the distribution because a single value to describe the centre cannot be identified.In some cases, particularly where the data are continuous, the distribution may have no mode at all (i.e. if all values are different).In cases such as these, it may be better to consider using the median or mean, or group the data in to appropriate intervals, and find the modal class.

What is the median?

The median is the middle value in distribution when the values are arranged in ascending or descending order. The median divides the distribution in half (there are 50% of observations on either side of the median value). In a distribution with an odd number of observations, the median value is the middle value. Looking at the retirement age distribution (which has 11 observations), the median is the middle value, which is 57 years:

54, 54, 54, 55, 56, 57, 57, 58, 58, 60, 60

When the distribution has an even number of observations, the median value is the mean of the two middle values. In the following distribution, the two middle values are 56 and 57, therefore the median equals 56.5 years:

52, 54, 54, 54, 55, 56, 57, 57, 58, 58, 60, 60

Advantage of the median:

The median is less affected by outliers and skewed data than the mean, and is usually the preferred measure of central tendency when the distribution is not symmetrical.

Limitation of the median:

The median cannot be identified for categorical nominal data, as it cannot be logically ordered.

What is the mean?

The mean is the sum of the value of each observation in a dataset divided by the number of observations. This is also known as the arithmetic average. Looking at the retirement age distribution again:

54, 54, 54, 55, 56, 57, 57, 58, 58, 60, 60

NOTES

The mean is calculated by adding together all the values (54+54+54+55+56+57+57+58 +58+60+60 = 623) and dividing by the number of observations (11) which equals 56.6 years.

Advantage of the mean:

The mean can be used for both continuous and discrete numeric data.

Limitations of the mean:

The mean cannot be calculated for categorical data, as the values cannot be summed. As the mean includes every value in the distribution the mean is influenced by outliers and skewed distributions.

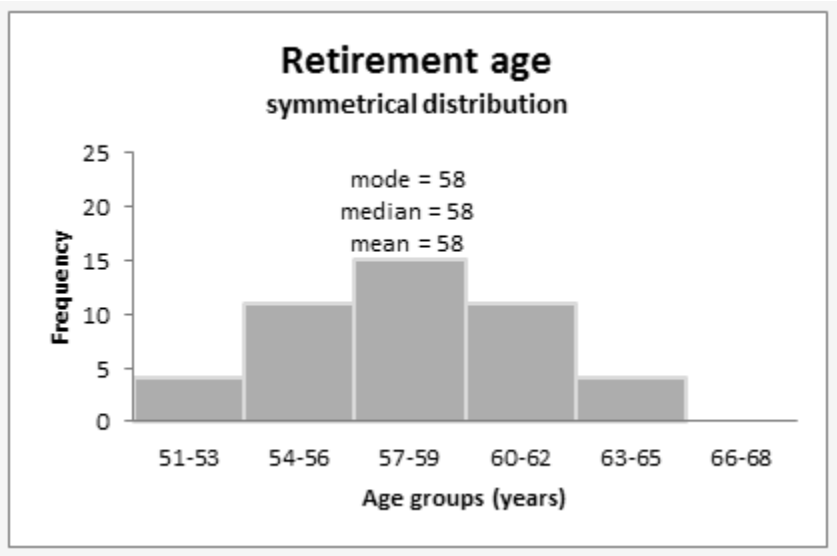
What else do I need to know about the mean?

The population mean is indicated by the Greek symbol μ (pronounced 'mu'). When the mean is calculated on a distribution from a sample it is indicated by the symbol $\bar{x}$ (pronounced X-bar).

How does the shape of a distribution influence the Measures of Central Tendency?

Symmetrical distributions:

When a distribution is symmetrical, the mode, median and mean are all in the middle of the distribution. The following graph shows a larger retirement age dataset with a distribution which is symmetrical. The mode, median and mean all equal 58 years.

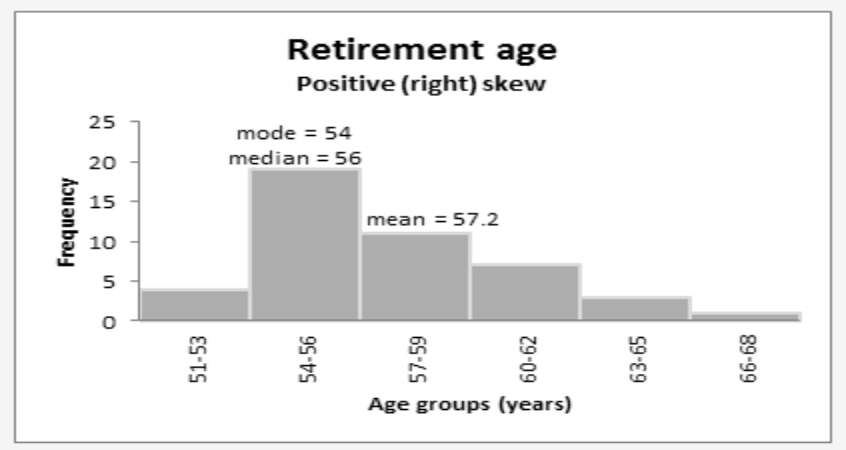


Skewed distributions:

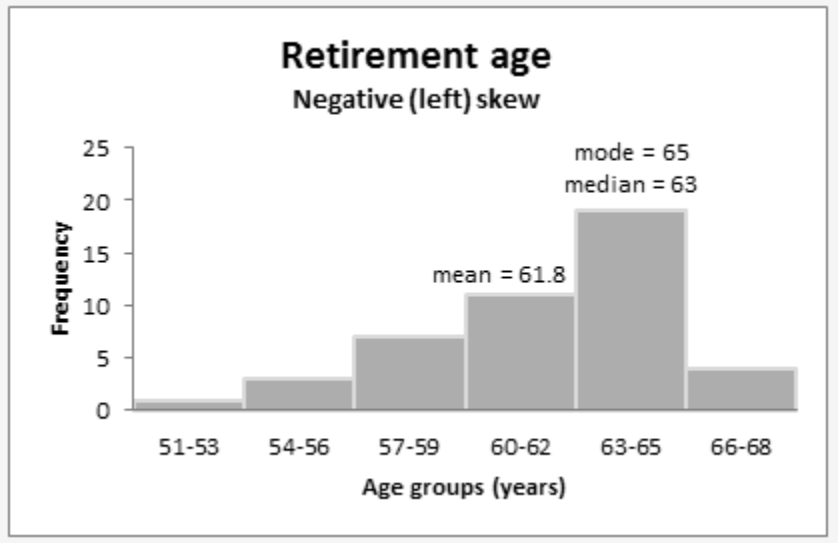
When a distribution is skewed the mode remains the most commonly occurring value, the median remains the middle value in the distribution, but the mean is generally 'pulled' in the direction of the tails. In a skewed distribution, the median is often a preferred measure of central tendency, as the mean is not usually in the middle of the distribution. A distribution is said to be positively or right skewed when the tail on the right side of the distribution is longer than the left side. In a positively skewed distribution it is common for the mean to be 'pulled' toward the right tail of the distribution. Although there are

exceptions to this rule, generally, most of the values, including the median value, tend to be less than the mean value.

The following graph shows a larger retirement age data set with a distribution which is right skewed. The data has been grouped into classes, as the variable being measured (retirement age) is continuous. The mode is 54 years, the modal class is 54-56 years, the median is 56 years and the mean is 57.2 years.



A distribution is said to be negatively or left skewed when the tail on the left side of the distribution is longer than the right side. In a negatively skewed distribution, it is common for the mean to be 'pulled' toward the left tail of the distribution. Although there are exceptions to this rule, generally, most of the values, including the median value, tend to be greater than the mean value. The following graph shows a larger retirement age dataset with a distribution which left skewed. The mode is 65 years, the modal class is 63-65 years, the median is 63 years and the mean is 61.8 years.



How do outliers influence the measures of central tendency?

Outliers are extreme, or atypical data value(s) that are notably different from the rest of the data. It is important to detect outliers within a distribution, because they can alter the

NOTES 

results of the data analysis. The mean is more sensitive to the existence of outliers than the median or mode. Consider the initial retirement age dataset again, with one difference; the last observation of 60 years has been replaced with a retirement age of 81 years. This value is much higher than the other values, and could be considered an outlier. However, it has not changed the middle of the distribution, and therefore the median value is still 57 years.

54, 54, 54, 55, 56, 57, 57, 58, 58, 60, 81

As the all values are included in the calculation of the mean, the outlier will influence the mean value. (54+54+54+55+56+57+57+58+58+60+81 = 644), divided by 11 = 58.5 years

In this distribution the outlier value has increased the mean value. Despite the existence of outliers in a distribution, the mean can still be an appropriate measure of central tendency, especially if the rest of the data is normally distributed. If the outlier is confirmed as a valid extreme value, it should not be removed from the dataset. Several common regression techniques can help reduce the influence of outliers on the mean value.

4.10 INFERENTIAL STATISTICS: Z TEST – T TEST

Z TEST

Z test is a statistical test that is conducted on data that approximately follows a normal distribution. The z test can be performed on one sample, two samples, or on proportions for hypothesis testing. It checks if the means of two large samples are different or not when the population variance is known. A z test can further be classified into left-tailed, right-tailed, and two-tailed hypothesis tests depending upon the parameters of the data. In this article, we will learn more about the z test, its formula, the z test statistic, and how to perform the test for different types of data using examples.

What is Z Test?

A z test is a test that is used to check if the means of two populations are different or not provided the data follows a normal distribution. For this purpose, the null hypothesis and the alternative hypothesis must be set up and the value of the z test statistic must be calculated. The decision criterion is based on the z critical value.

Z Test Definition

A z test is conducted on a population that follows a normal distribution with independent data points and has a sample size that is greater than or equal to 30. It is used to check whether the means of two populations are equal to each other when the population variance is known. The null hypothesis of a z test can be rejected if the z test statistic is statistically significant when compared with the critical value.

Z Test Formula

The z test formula compares the z statistic with the z critical value to test whether there is a difference in the means of two populations. In hypothesis testing, the z critical value divides the distribution graph into the acceptance and the rejection regions. If the test statistic falls in the rejection region, then the null hypothesis can be rejected otherwise

it cannot be rejected. The z test formula to set up the required hypothesis tests for a one sample and a two-sample z test are given below.

One-Sample Z Test

A one-sample z test is used to check if there is a difference between the sample mean and the population mean when the population standard deviation is known. The formula for the z test statistic is given as follows:

$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$. $\bar{x}$ is the sample mean, μ is the population mean, σ is the population standard deviation and n is the sample size? The algorithm to set a one sample z test based on the z test statistic is given as follows:

Left Tailed Test:

Null Hypothesis: H0: μ=μ0

Alternate Hypothesis: H1: μ<μ0

Decision Criteria: If the z statistic < z critical value then rejects the null hypothesis.

Right Tailed Test:

Null Hypothesis: H0: μ=μ0

Alternate Hypothesis: H1: μ>μ0

Decision Criteria: If the z statistic > z critical value then rejects the null hypothesis.

Two Tailed Test:

Null Hypothesis: H0: μ=μ0

Alternate Hypothesis: H1: μ ≠ μ0

Decision Criteria: If the z statistic > z critical value then rejects the null hypothesis.

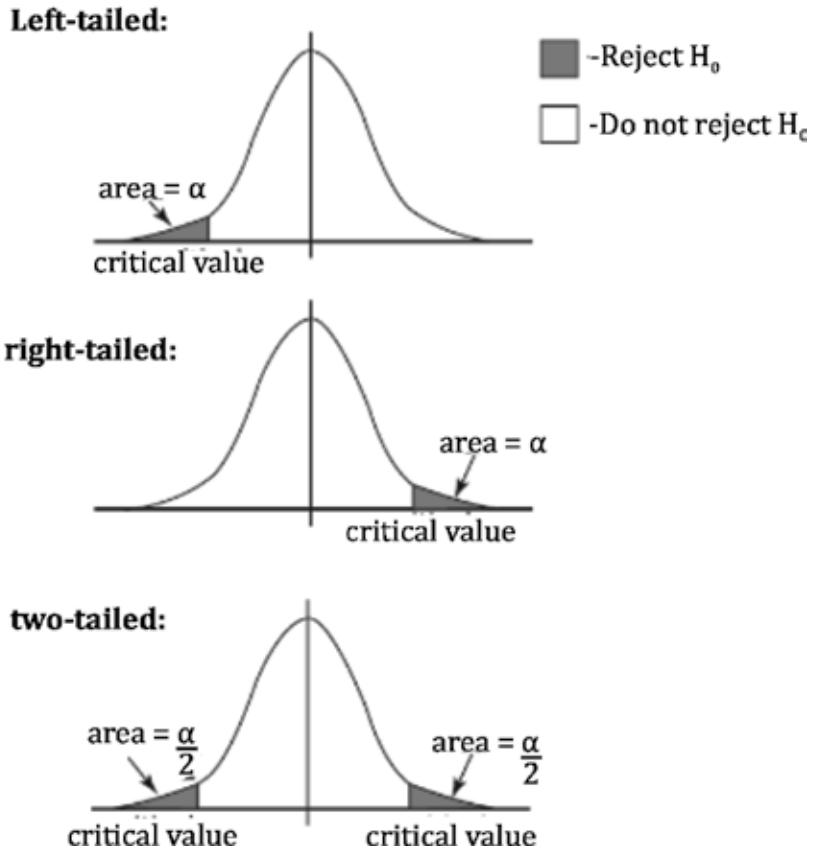
Two Sample Z Test

A two sample z test is used to check if there is a difference between the means of two samples. The z test statistic formula is given as follows:

$z = \frac{(\bar{x_1} - \bar{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$. $\bar{x_1}, \mu_1, \sigma_1^2$ are the sample mean, population mean and population variance respectively for the first sample. $\bar{x_2}, \mu_2, \sigma_2^2$ are the sample mean, population mean and population variance respectively for the second sample. The two-sample z test can be set up in the same way as the one-sample test. However, this test will be used to compare the means of the two samples. For example, the null hypothesis is given as $H_0 : \mu_1 = \mu_2$.

NOTES

Rejection Region for Null Hypothesis



Z Test for Proportions

A z test for proportions is used to check the difference in proportions. A z test can either be used for one proportion or two proportions. The formulas are given as follows.

One Proportion Z Test

A one proportion z test is used when there are two groups and compares the value of an observed proportion to a theoretical one. The z test statistic for a one proportion z test is given as follows:

$$z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}.$$

Here, p is the observed value of the proportion, p0 is the theoretical proportion value and n is the sample size. The null hypothesis is that the two proportions are the same while the alternative hypothesis is that they are not the same.

Two Proportion Z Test

A two-proportion z test is conducted on two proportions to check if they are the same or not. The test statistic formula is given as follows:

$$z = \frac{p_1 - p_2 - 0}{\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

$$\text{where } p = \frac{x_1 + x_2}{n_1 + n_2}$$

p1 is the proportion of sample 1 with sample size n1 and x1 number of trials. p2 is the proportion of sample 2 with sample size n2 and x2 number of trials.

How to Calculate Z Test Statistic?

The most important step in calculating the z test statistic is to interpret the problem correctly. It is necessary to determine which tailed test needs to be conducted and what type of test does the z statistic belong to. Suppose a teacher claims that his section's students will score higher than his colleague's section. The mean score is 22.1 for 60 students belonging to his section with a standard deviation of 4.8. For his colleague's section, the mean score is 18.8 for 40 students and the standard deviation is 8.1. Test his claim at $\alpha = 0.05$. The steps to calculate the z test statistic are as follows:

- Identify the type of test. In this example, the means of two populations have to be compared in one direction thus, the test is a right-tailed two-sample z test.
- Set up the hypotheses. H0: $\mu 1 = \mu 2$, H1: $\mu 1 > \mu 2$.
- Find the critical value at the given alpha level using the z table. The critical value is 1.645.
- Determine the z test statistic using the appropriate formula. This is given by $z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$.
- Substitute values in this equation. $\bar{x}1 = 22.1$, $\sigma 1 = 4.8$, $n1 = 60$, $\bar{x}2 = 18.8$, $\sigma 2 = 8.1$, $n2 = 40$ and $\mu 1 - \mu 2 = 0$. Thus $z = 2.32$
- Compare the critical value and test statistic to arrive at a conclusion. As $2.32 > 1.645$ thus, the null hypothesis can be rejected. It can be concluded that there is enough evidence to support the teacher's claim that the scores of students are better in his class.

Important Notes on Z Test:

- Z test is a statistical test that is conducted on normally distributed data to check if there is a difference in means of two data sets.
- The sample size should be greater than 30 and the population variance must be known to perform a z test.
- The one-sample z test checks if there is a difference in the sample and population mean,
- The two sample z test checks if the means of two different groups are equal.

NOTES

T-TEST FORMULA

The t-test formula helps us to compare the average values of two data sets and determine if they belong to the same population or are they different. The t-score is compared with the critical value obtained from the t-table. The large t-score indicates that the groups are different and a small t-score indicates that the groups are similar.

What Is the T-test Formula?

The t-test formula is applied to the sample population. The t-test formula depends on the mean, variance, and standard deviation of the data being compared. There are 3 types of t-tests that could be performed on the n number of samples collected.

- One-sample test,
- Independent sample t-test and
- Paired samples t-test

The critical value is obtained from the t-table looking for the degree of freedom (df = n-1) and the corresponding α value (usually 0.05 or 0.1). If the t-test obtained statistically > CV then the initial hypothesis is wrong and we conclude that the results are significantly different.

One-Sample T-Test Formula

For comparing the mean of a population $\bar{x}$ from n samples, with a specified theoretical mean μ, we use a one-sample t-test.

$$t= \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

where$\sigma/\sqrt{n}$ is the standard error

t-Test Formula

$$t= \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

where,
'x' bar is the mean of the sample,
μ is the assumed mean,
σ is the standard deviation
and n is the number of observations

Independent Sample T-Test

Students t-test is used to compare the mean of two groups of samples. It helps evaluate if the means of the two sets of data are statistically significantly different from each other.

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

T-test Formula

$$d = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

$$df = n_1 + n_2 - 1$$

$$s_1^2 = \frac{\sum_{i=1}^{n_1} (x_i - x_1)^2}{n_2 - 1}$$

$$s_2^2 = \frac{\sum_{i=1}^{n_2} (x_i - \overline{x_2})^2}{n_2 - 1}$$

where

- t = Student's t-test
- x1 = mean of first group
- x2= mean of second group
- s1 = standard deviation of group 1
- s2 = standard deviation of group 1
- n1= number of observations in group 1
- n2= number of observations in group 2

Paired Samples T-Test

Whenever two distributions of the variables are highly correlated, they could be pre and post test results from the same people. In such cases, we use the paired samples t-test.

$$t = \frac{\Sigma(x_1 - x_2)}{\frac{s}{\sqrt{n}}}$$

where

t = Student's t-test

$x_1 - x_2$ = Difference mean of the pairs

s= standard deviation

n = sample size

NOTES

Examples Using t-test Formula

Example: Calculate a t-test for the following data of the number of times people prefer coffee or tea in five-time intervals.

Coffee	Tea
4	3
5	8
7	6
6	4
9	7

Solution: let x1 be the sample of data that prefers coffee and x2 be the sample of data that prefers tea. let us find the mean, variance and the SD.

x_1	$(x_1 - \bar{x_1})$	$(x_1 - \bar{x_1})^2$	x_2	$(x_2 - \bar{x_2})$	$(x_2 - \bar{x_2})^2$
4	-2.2	4.84	3	-2.6	6.76
5	-1.2	1.44	8	2.4	5.76
7	0.8	0.64	6	0.4	0.16
6	-0.2	0.04	4	-1.6	2.56
9	2.8	7.84	7	1.4	1.96
6.2		**14.8**	**5.6**		**17.20**

$\bar{x_1} = 31/5 = 6.2$

$\bar{x_2} = 28/5 = 5.6$

$\Sigma(x_1 - \bar{x_1})^2 = 14.8$

$\Sigma(x_2-\bar{x}_2)^2 = 17.2$

$S_1 = 14.8/4 = 3.7$

$S_2 = 17.2/4 = 4.3$

According to the t-test formula,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

Applying the known values in the t-test formula, we get

$$t = \frac{6.2 - 5.6}{\sqrt{\left(\frac{3.7}{5} + \frac{4.3}{5}\right)}}$$

$$= \frac{0.6}{\sqrt{1.6}} = 0.6/1.26 = 0.47$$

t = 0.47

4.11 ANALYSIS OF VARIANCE

ANOVA (Analysis of Variance) is a statistical tool to test the homogeneity of different groups based on their differences.

- ANOVA is the method of analyzing the variance in a set of data and dividing the variance into groups according to the sources of those variations.
- ANOVA is based on the principle that the total amount of differences in a set of data can be divided into two types, the amount that can be attributed to chance and the other that is caused due to specific causes.
- In a population, ANOVA is used to determine the difference between the means of the samples by analyzing the variation within each of the samples, and relative to the variation between the samples.
- While performing ANOVA, two assumptions are made where the first is that the samples are extracted from a normal population, and the second is that all factors other than those being tested are controlled.

ONE-WAY ANOVA

- One-way ANOVA is a short-cut method where a single factor is considered, and its effect on the samples is observed.
- It is a commonly used technique as it is a more convenient method.

NOTES

- This method is performed when the means of the samples and/or the mean of the sample means are non-integer values.
- One drawback of one-way ANOVA is that it cannot tell which specific groups are different from each other but can tell that at least two groups are different.
- In one-way ANOVA, at least three groups are analyzed as a t-test can be used to determine the difference between two groups.
- Thus, instead of the F-test, the t-test can be performed, which significantly reduces the time and effort. The relation between ANOVA and t-test can be explained as F=t2.

TWO-WAY ANOVA

- The two-way ANOVA technique is used in cases when the given set of data is classified under two different independent factors.
- Here, measurements are taken for each factor separately, and thus the measurements may or may not repeated values.
- The primary purpose of two-way ANOVA is to determine if there is a relation between the independent factors and the dependent factor.
- This technique helps us to determine if the effect of the independent factor on the dependent factor is influenced by the other independent factor or not.

The following is the ANOVA table for two-way ANOVA:

Sources of variation	Sum of squares (SS)	Degrees of freedom (d.f)	Mean sum of square (MS)	F-ratio
Between columns	$\sum \frac{(Tj^2)}{Nj} - \frac{(T^2)}{n}$	(c-1)	$\frac{SS\ between\ columns}{(c-1)}$	$\frac{MS\ between\ columns}{MS\ residual}$
Between rows	$\sum \frac{(Ti^2)}{Ni} - \frac{(T^2)}{n}$	(r-1)	$\frac{SS\ between\ rows}{(r-1)}$	$\frac{MS\ between\ rows}{MS\ residual}$
Residual error	Total SS-(SS between columns and SS between rows)	(c-1)(r-1)	$\frac{SS\ residual}{(c-1)(r-1)}$	
Total	$\sum Xij^2 - \frac{(T^2)}{n}$	(c.r -1)		

In the table,

c= number of columns

r= number of rows

T= the total of the values of individual items

Tj= the sum of the values in the column

Ti= the sum of the values in the row

- The ANOVA table shows the statistics used to test hypotheses about the population means.
- Here, the F-ratios for rows and columns are compared with their corresponding table values, for the given degree of freedom and given level of significance.
- If the calculated F-ratio is found to be equal or higher than its table value, the differences or variation among the columns are considered significant.
- A similar process is employed for rows to determine the significance of the variation.

ANOVA EXAMPLES

- One example of one-way ANOVA is the analysis of the exam performance of students based on their test anxiety.
- One example of the classification of agricultural products on the basis of different seeds and different fertilizers used. Here, the agricultural output is the dependent factor, whereas seeds and fertilizers are independent factors.

4.12 CORRELATIONAL RESEARCH

Correlational research is a type of non-experimental research method in which a researcher measures two variables, understands and assesses the statistical relationship between them with no influence from any extraneous variable. Our minds can do some brilliant things. For example, it can memorize the jingle of a pizza truck. The louder the jingle, the closer the pizza truck is to us. Who taught us that? Nobody! We relied on our understanding and came to a conclusion. We don't stop there, do we? If there are multiple pizza trucks in the area and each one has a different jingle, we would memorize it all and relate the jingle to its pizza truck.

This is what correlational research precisely is, establishing a relationship between two variables, "jingle" and "distance of the truck" in this particular example. The correlational study is looking for variables that seem to interact with each other. When you see one variable changing, you have a fair idea of how the other variable will change.

Example

The correlation coefficient shows the correlation between two variables (A correlation coefficient is a statistical measure that calculates the strength of the relationship between two variables), a value measured between -1 and +1. When the correlation coefficient is close to +1, there is a positive correlation between the two variables. If the value is close to -1, there is a negative correlation between the two variables. When the value is close to zero, then there is no relationship between the two variables. Let us take an example to understand correlational research.

NOTES

Consider hypothetically; a researcher is studying a correlation between cancer and marriage. In this study, there are two variables: disease and marriage. Let us say marriage has a negative association with cancer. This means that married people are less likely to develop cancer. However, this doesn't necessarily mean that marriage directly avoids cancer. In correlational research, it is not possible to establish the fact, what causes what. It is a misconception that a correlational study involves two quantitative variables. However, the reality is two variables are measured, but neither is changed. This is true independent of whether the variables are quantitative or categorical.

TYPES OF CORRELATIONAL RESEARCH

Mainly three types of correlational research have been identified:

1. **Positive correlation:** A positive relationship between two variables is when an increase in one variable leads to a rise in the other variable. A decrease in one variable will see a reduction in the other variable. For example, the amount of money a person has might positively correlate with the number of cars the person owns.
2. **Negative correlation:** A negative correlation is quite literally the opposite of a positive relationship. If there is an increase in one variable, the second variable will show a decrease and vice versa. For example, being educated might negatively correlate with the crime rate when an increase in one variable leads to a decrease in another and vice versa. If the level of education in a country is improved, it can lower crime rates. Please note that this doesn't mean that lack of education leads to crimes. It only means that a lack of education and crime is believed to have a common reason – poverty.
3. **No correlation:** In this third type, there is no correlation between the two variables. A change in one variable may not necessarily see a difference in the other variable. For example, being a millionaire and happiness is not correlated. An increase in money doesn't lead to happiness.

CHARACTERISTICS OF CORRELATIONAL RESEARCH

Correlational research has three main characteristics. They are:

- **Non-experimental:** Correlational study is non-experimental. It means that researchers need not manipulate variables with a scientific methodology to either agree or disagree with a hypothesis. The researcher only measures and observes the relationship between the variables, without altering them or subjecting them to external conditioning.
- **Backward-looking:** Correlational research only looks back at historical data and observes events in the past. Researchers use it to measure and spot historical patterns between two variables. A correlational study may show a positive relationship between two variables, but this can change in the future.
- **Dynamic:** The patterns between two variables from correlational research are never constant and are always changing. Two variables having a negative correlation research in the past can have a positive correlation relationship in the future due to various factors.

DATA COLLECTION

The distinctive feature of correlational research is that the researcher can't manipulate either of the variable involved. It doesn't matter how or where the variables are measured. A researcher could observe participants in a closed environment or a public setting. Researchers use two data collection methods to collect information in correlational research.

- **Naturalistic observation**
  Naturalistic observation is a way of data collection in which people's behavior is observed in their natural environment, in which they typically exist. This method is a type of field research. It could mean a researcher might be observing people in a grocery store, at the cinema, playground, or similar places. Researchers who are usually involved in this type of data collection make observations as unobtrusively as possible so that the participants involved in the study are not aware that they are being observed else they might deviate from being their natural self.

  Ethically this method is acceptable if the participants remain anonymous, and if the study is conducted in a public setting, a place where people would not normally expect complete privacy. As mentioned previously, taking an example of the grocery store where people can be observed while collecting an item from the aisle and putting in the shopping bags. This is ethically acceptable, and that is the reason most researchers choose public settings for recording their observations. This data collection method could be both qualitative or quantitative.

- **Archival data**
  Another approach to correlational data is the use of archival data. Archival information is the data that has been previously collected by doing similar kinds of research. Archival data is usually made available through primary research. In contrast to naturalistic observation, the information collected through archived data can be quite straightforward. For example, counting the number of people named Richard in the various states of America based on social security records is quite straightforward.

4.13 CHAPTER SUMMARY

As we discussed earlier, the data consists of quantitative variables like price, income, sales etc., and qualitative variables like knowledge, performance, character etc. The qualitative information must be converted into numerical form for further analysis. This is possible through measurement and scaling techniques. A common feature of survey-based research is to have respondent's feelings, attitudes, opinions, etc. in some measurable form. For example, a bank manager may be interested in knowing the opinion of the customers about the services provided by the bank. Similarly, a fast-food company having a network in a city may be interested in assessing the quality and service provided by them. Measurement is the process of observing and recording the observations that are collected as part of research. The recording of the observations may be in terms of numbers or other symbols to characteristics of objects according to certain prescribed rules. Scaling is the assignment of objects to numbers or semantics according to a rule. In scaling, the objects are text statements, usually statements of attitude, opinion, or feeling. For example, consider a scale locating customers of a bank according to the characteristic "agreement to the satisfactory quality of service provided by

NOTES 

the branch". The level of measurement refers to the relationship among the values that are assigned to the attributes, feelings or opinions for a variable. A frequency distribution shows the frequency of repeated items in a graphical form or tabular form. It gives a visual display of the frequency of items or shows the number of times they occurred. Let's learn about frequency distribution in this article in detail. A measure of central tendency (also referred to as measures of centre or central location) is a summary measure that attempts to describe a whole set of data with a single value that represents the middle or centre of its distribution.

4.14 REVIEW QUESTIONS

SHORT ANSWER TYPE QUESTIONS

1. **Nominal scale**
2. **Explain correlation**
3. **What do you understand by measure of central tendency?**
4. **Explain ratio scale.**
5. **Explain q-sort scale.**

LONG ANSWER TYPE QUESTIONS

1. **Enlist the type of distributions.**
2. **Elaborate the inferential statistics.**
3. **Write a brief note on continuous rating scales.**
4. **Explain Likert scale.**
5. **What do you understand by semantic differential scale**

4.15 MULTIPLE CHOICE QUESTIONS

1. **_____ is the crudest among all measurement scales but it is also the simplest scale. In this scale the different scores on a measurement simply indicate different categories.**
   a. Nominal Scale
   b. Ordinal Scale
   c. Interval Scale
   d. Ratio Scale

2. **_____ involves the ranking of items along the continuum of the characteristic being scaled.**
   a. Nominal Scale
   b. Ordinal Scale
   c. Interval Scale
   d. Ratio Scale

3. **_____ is a scale in which the numbers are used to rank attributes such that numerically equal distances on the scale represent equal distance in the characteristic being measured.**
   a. Nominal Scale
   b. Ordinal Scale

c. Interval Scale
d. Ratio Scale

4. _____ is the highest level of measurement scales. This has the properties of an interval scale together with a fixed (absolute) zero point.
a. Nominal Scale
b. Ordinal Scale
c. Interval Scale
d. Ratio Scale

5. _____ is a comparative scaling technique in which a respondent is presented with two objects at a time and asked to select one object according to some criterion.
a. Paired Comparison Scale
b. Rank Order Scale
c. Constant Sum Scale
d. Q-Sort Scale

6. _____ is another type of comparative scaling technique in which respondents are presented with several items simultaneously and asked to rank them in the order of priority.
a. Paired Comparison Scale
b. Rank Order Scale
c. Constant Sum Scale
d. Q-Sort Scale

7. _____ in this scale, the respondents are asked to allocate a constant sum of units such as points, rupees, or chips among a set of stimulus objects with respect to some criterion.
a. Paired Comparison Scale
b. Rank Order Scale
c. Constant Sum Scale
d. Q-Sort Scale

8. _____ is a comparative scale that uses a rank order procedure to sort objects based on similarity with respect to some criterion.
a. Paired Comparison Scale
b. Rank Order Scale
c. Constant Sum Scale
d. Q-Sort Scale

9. _____ in business research, the Likert scale, developed by Rensis Likert, is extremely popular for measuring attitudes, because, the method is simple to administer.
a. Itemised Rating Scales
b. Likert Scale
c. Semantic Differential Scale
d. Stapel scale

NOTES 

10. The _____ was originally developed to measure the direction and intensity of an attitude simultaneously.

a. Itemised Rating Scales
b. Likert Scale
c. Semantic Differential Scale
d. Stapel scale

♦♦♦♦♦

UNIT V

REPORT WRITING

STRUCTURE

NOTES 

5.1 LEARNING OBJECTIVE

After learning this unit students will be able to:

- Understand the Report Writing.
- Understand the Technique of Interpretation.
- Understand the Mechanics of Writing a Research Report.

5.2 INTRODUCTION TO REPORT WRITING

A research report is considered a major component of any research study as the research remains incomplete till the report has been presented or written. No matter how good a research study, and how meticulously the research study has been conducted, the findings of the research are of little value unless they are effectively documented and communicated to others. The research results must invariably enter the general store of knowledge. Writing a report is the last step in a research study and requires a set of skills somewhat different from those called for in actually conducting research.

Significance / Objectives of A Report

A well written report helps in knowledge building in the concerned area but also helps in future research. While all the necessary information is presented in the appropriate manner so that the targeted readers may be able to understand and utilize the same. The objectives of a research report are:

1. Conveying of knowledge to the concerned people in the field of research
2. Proper presentation of the findings for further utilization of the recommendations.
3. Give impetus to research in the concerned knowledge area.
4. To re-examine the validity of generalizations drawn by the researcher after the report has been submitted.

5.3 TECHNIQUE OF INTERPRETATION

- After the data is collected and analysed using several data analysis methods, the next task is to draw Inferences from these data.
- In other words, Interpretation of data needs to be done, so as to derive certain conclusions, which is the whole purpose of the research study.

Definition

"Interpretation refers to the process of making sense of numerical data that has been collected, analysed and presented".

Need of Interpretation

- Maintaining Continuity of Research
- Pointers for further Research
- Communicate Significance of Research
- Transition of Exploratory Research to Experimental Research

NOTES 

Techniques of Interpretation Steps

1. Explanations of the relations which the researcher has found during the research study
2. Interpretation of the lines of relationship in terms of the underlying processes
3. Finding thread of uniformity that lies beneath the diversified research findings and thereby, forming generalizations and theories.
4. Extraneous information, collected during the study, must also be considered while interpreting the final results of research study, as it may result in better understanding of the problem in hand.
5. Consulting an expert having an insight of the research study who points out omissions and errors in logical argumentation will result in correct interpretation and, add to usefulness of the research results.
6. All relevant factors affecting the problem must be considered before forming Interpretations or Generalizations. Otherwise, it may lead to incorrect conclusions.

Precautions in Interpretation

1. Ensure Proper Data Collection
2. Data Analysis
3. Errors can Arise
4. Results of Hypothesis Testing
5. Statistical Measures
6. Avoid Broad Generalizations
7. Quality of Interpretation

TYPES OF REPORT

Research reports vary greatly in length and type depending on the subject. For example, banks and other financial institutions prefer short balance sheet type of tabulations for their annual report. In mathematics, the report may consist of many algebraic notations, whereas a chemist's report may be in the form of symbols and formulae. Students of literature usually write a long report critically analysing a writer or book. The news items found in newspapers are also a form of report writing. Other examples of reports include book reviews, reports prepared by government bureaus, PhD theses, etc. Any research investigation may be presented in like a technical report, a popular report, an article, a monograph, or, at times, even in the form of an oral presentation. The technical report is prepared for specialists who have interest in understanding the technical procedure and terminology used in the research project. The report will be in technical language. In the technical report, the main emphasis is on: (i) the methods employed; (ii) assumptions made in the course of study and; (iii) the detailed presentation of the findings, including their limitations and supporting data. Popular data is intended for persons who have limited interest in the technical aspects of the research methodology and research findings. The audience will include laymen and even top executives who want summary reports. The popular report is one which gives emphasis on simplicity and attractiveness. The simplification should be sought through clear writing, minimizing of technical, particularly mathematical details, and liberal use of charts and diagrams. Attractive layouts along with

NOTES

large print and many subheadings is another feature of a popular report. In such a report, emphasis is given on practical aspects and policy implications.

WRITING THE RESEARCH REPORT

Once the data collection and analysis work is over, the researcher will start writing the research report. Social and development research reports need to have a logical, clear structure be to the point use simple language, and have a pleasant layout Just as an architect has to draw a layout plan for a house that is being designed, you first have to make an outline for your report. This outline will contain a head, a body, and a tail. The head consists of a description of your problem within its context (the country and research area), the objectives of the study and the methodology followed. This part should not comprise more than one quarter of the report, otherwise it becomes top-heavy. The body will form the bigger part of your report: it will contain the research findings. The tail, finally, consists of the discussion of your data, conclusions, and recommendations. Before you start writing, it is essential to group and review the data you have analysed by objective. Check whether all data has indeed been processed and analysed as you planned in the research protocol/proposal which is duly approved. Draw major conclusions and relate these to the research literature. Again, you may be inspired to go back to your raw data and refine your analysis, or to search for additional literature to answer questions that the analysis of your data may evoke. Compile the major conclusions and tables or quotes from qualitative data related to each specific objective. You are now ready to draft the report.

The research report will have, broadly, three parts.

Part I: The Preliminary Pages

Part II: The Main Text of the Research Report

Part III: The End Matter

THE PRELIMINARY PAGES OF THE RESEARCH REPORT

The preliminary pages of the research report should have the following main constituents.

• Title and cover page • A foreword • Preface • Acknowledgements • Table of contents • List of tables • List of figures • List of appendices • List of abbreviations • Executive Summary

i. Title and Cover page

The cover page should contain the title, the names of the authors with their designations, the institution that is publishing the report with its logo, (e.g., Health Systems Research Unit, Ministry of Health), the month, and the year of publication. The title could consist of a challenging statement or question, followed by an informative subtitle covering the content of the study and indicating the area where the study was implemented. However, this is suggestive in nature and should not be considered standard. It would be appropriate if the cover page is designed by an expert in computer graphics who may be suggested to include some important photograph related to the identity of the organization or problem under study or from the field within the background. Design software may be used. An example of a title of a research report is given in the box below.

Title of the research report

Labour Migration and its Implication on Rural Economy of Indo-Gangetic Plains of India

ii. Foreword

A foreword is usually a short piece of writing found at the beginning of a book or other piece of literature, before the introduction. This may or may not be written by the primary author of the work. Often, a foreword will tell of some interaction between the writer of the foreword and the story, or, the writer of the story. A foreword to later editions of a work often explains how the new edition differs from previous ones. Unlike a preface, a foreword is always signed. An example of a foreword is given in the box below.

Foreword

Migration of all kinds, particularly income seeking migration across state boundaries, has attracted much attention in recent scholarly and policy literature. This study provides sufficient evidences of the effect of labour migration, more specifically, male outmigration on the rural economy of the Indo-Gangetic region. The number of districts of high and moderately high male outmigration has increased. The findings reveal the holistic scenario of migration led changes in agricultural and household domains. I am sure that this volume would be of great interest to researchers, policy makers, and development agencies while framing strategies for agricultural and rural development.

iii. Preface

A preface, by contrast, is written by the author of the book. A preface generally covers the story of how the book came into being, or how the idea for the book was developed; this is often followed by thanks and acknowledgments to people who were helpful to the author during the time of writing. A preface is an introduction to a book or other literary work written by the work's author. An example of the preface is given in the box below.

Preface

The present study was conducted in three states of Bihar, Uttar Pradesh, and Punjab to study various aspects of labor migration, and its impact on the rural economy in the Indo-Gangetic plains in India. The study focused on labor outmigration across two states of the Indo-Gangetic Region and immigration in Punjab. The results of this study would help researchers, policymakers, and planners as well as development agencies in addressing various issues of labor migration and its implication in India.

iv. Acknowledgements

It is good practice to thank those who supported you technically or financially in the design and implementation of your study. You should not forget to thank your research guide and your employer, too, who has allowed you to invest time in the study; and, the respondents may be acknowledged. You should not forget to acknowledge the contribution of computer professionals, library staff, local officials, and the community at large that provided the information.

NOTES

Acknowledgments are usually placed right after the title page or at the end of the report, before the references. An example of acknowledgment is given in the box below.

Acknowledgments

I take this opportunity to thank the Indian Council of Agricultural Research for providing funds and facilities for the project. I offer my sincere thanks to the Director of, the Indian Agricultural Research Institute for his encouragement and support in pursuing this study. I am also grateful to the head, of the Division of Agricultural Economics, IARI for providing all needed support, encouragement, and technical guidance. All the Research Associates, Senior Research Fellows, and technical assistants working on the project deserve special appreciation for their hard work and sincere efforts in completing this project.

v. Table of Contents

A table of contents is essential. It provides the reader a quick overview of the chapters with major sections and sub-sections of your report, and page references so that the reader can go through the report in a different order, or skip certain sections. The sections and sub-sections within each chapter may be given numbers that are specific to the chapter. For example, a section in chapter III may be given no as 3.1; and, a subsection as 3.1.1. An example of a table of contents is given below.

Contents

S. No.	Contents	Pages
1	Introduction	
2	Review of Literatute	
3	Methodology	
3.1	Data	
3.2	Analytical Tools	
3.3	Profile of Are Under Study	
4	Research Findings	
4.1	Macro Level Evidences	
4.2	Evidences from filed Survey	
5	Discussion	
6	Conclusions and Policy Implications	
7	References	
	Appendix	

vi. List of Tables

If you have many tables or figures, it is essential to list these also in a table of contents with formatted with page numbers. The initial letters of the keywords in the title are capitalized and no terminal punctuation is used. An example is given below.

NOTES 

List of Tables

S. No.	Name of the Table	Pages
2.1	Simpling Pattern of Households in the Study Area	
3.1	Magrants by Last Residence in India	
3.2	Total Inter-State Migrants by Place of Birth in Major States	
3.3	Social Characteristics of Households in the Study Area	
-	-	
-	-	
-	-	

vii. List of Figures

The list of figures appears in the same format as the list of tables, titled List of Figures.

viii. List of Appendices

The appendices will contain any additional information that the researcher have collected while carrying out the study. It may be a questionnaire, a letter of appreciation, a government notification, etc. The list of appendices appears in the same format as the list of tables.

ix. List of Abbreviations (optional)

If abbreviations or acronyms are used in the report, these should be stated in full in the text the first time that they are mentioned. If there are many, they should be listed in alphabetical order as well. The list can be placed before the first chapter of the report. The table of contents and lists of tables, figures, and abbreviations should be prepared last, as only then can you include the page numbers of all chapters and sections, and sub-sections in the table of contents. Then, you can also finalize the numbering of figures and tables and include all abbreviations. An example of a List of Abbreviations follows.

x. List of Abbreviations

List of Abbreviations
AI: Agreement Index
CMIE: Centre for Monitoring of Indian Economy
CV: Coefficient of Variation
DEA: Data Envelopment Analysis

xi. Executive Summary

The summary should be written only after the first or even the second draft of the report has been completed. It should contain:

- a very brief description of the problem (Why this study was needed)— the main
- objectives (What has been studied)
- the place of study (Where)
- the type of study and methods used (How)
- the major findings and conclusions

NOTES

- the major (or all) recommendations.

The summary will be the first (and for busy program managers/decision makers most likely the only) part of your study that will be read. Therefore, it demands thorough reflection and is time-consuming. Several drafts may have to be made, each discussed by the research team as a whole. As you may have collaborated with various groups during the drafting and implementation of your research proposal, you may consider writing different summaries for each of these groups. For example, you may prepare different summaries for policymakers and program managers, for implementing staff of lower levels, for community members, or for the public at large (newspaper, TV). At a later stage, you may write articles in scientific journals. In this section, we discussed the types of reports and the contents to be included in the preliminary pages of the research report. Now answer the following questions.

5.4 THE MECHANICS OF WRITING A RESEARCH REPORT

Beginning researchers may find the writing style used for research reports awkward or unesthetic, but there is a definite purpose behind the rules governing scientific writing: clarity. Every effort must be made to avoid ambiguity. Given the wide variety of approaches to research, it stands to reason that the approaches to writing a research report are equally varied. Most research reports, however, include only five basic sections or chapters: introduction, literature review, methods, results, and discussion.

Introduction

The introduction should alert the reader to what is to follow. Most introductions usually contain the following:

1. **Statement of the problem.** The first job of the report writer is to provide some information about the background and the nature of the problem under investigation. If the research topic has a long history, then a short summary is in order. This section should also discuss any relevant theoretical background that pertains to the research topic.
2. **Justification.** Another important area to be covered in this initial section is the rationale and justification for the project. This section should address the question of why it is important for us to spend time and energy researching this particular problem. Research can be important because it deals with a crucial theoretical issue, because it has practical value, or because it has methodological value.
3. **Aims of the current study.** Most introductory sections conclude with an unequivocal statement of the hypothesis or research question to be answered by the study.

LITERATURE REVIEW

The second major section is the review of the literature. In some formats, the literature review is incorporated into the introduction. As the name suggests, the literature review section briefly recapitulates the work done in the field. This review need not be exhaustive; the writer should summarize only those studies most relevant to the current project. All literature reviews should be accurate and relevant.

NOTES

1. **Accuracy**
   A concise and accurate distillation of each study in your review is a prerequisite for any literature review. The main points of each study—hypotheses that were tested, sample, method, findings, and implications— should be briefly summarized. The review should be selective but thorough.

2. **Relevance**
   A literature review should be more than a rote recitation of research studies. It must also contain analysis and synthesis. The writer is obligated to discuss the relevance of the past work to the current study. What theoretic development can be seen in past work? What major conclusions have recurred? What were some common problems? How do the answers to these questions relate to the current study? The ultimate aim of the review is to show how your study evolved out of past efforts and how the prior research provides a justification for your study.

METHODS

The methods section describes the approach used to confront the research problem. Some of the topics that are usually mentioned in this section are as follows.

1. **Variables used in the analysis**
   This includes a description of both independent and dependent variables, explaining how the variables were selected for the study, what marker variables, if any, were included, and how extraneous variables were controlled. Each variable also requires some justification for its use — variables cannot be added without reason. The mean and the standard deviation for each variable should be reported when necessary.

2. **Sample size**
   The researcher should state the number of subjects or units of study and also explain how these entities were selected. Additionally, any departure from normal randomization must be described in detail.

3. **Sample characteristics**
   The sample should also be described in terms of its demographic, lifestyle, or other descriptor characteristics. When human subjects are used, at least their age and sex should be indicated.

4. **Methodology**
   Every research report requires a description of the methods used to collect and analyze data. The amount of methodological description to be included depends on the audience; articles written for journals, for instance, must contain more detailed information than reports prepared in private sector research.

5. **Data manipulation**
   Often the collected data are not normally distributed, and researchers must use data transformation to achieve an approximation of normality. If such a procedure is used, a full explanation should be given.

RESULTS

The results section contains the findings of the research. It typically contains the following:

1. **Description of the analysis**
   The statistical techniques used to analyze the data should be mentioned. If the

NOTES 

analysis used common or easily recognized statistics, a one-sentence description might be all that is needed, such as "Chi-square analyses were performed on the data" or "Analysis of variance was performed….."If appropriate, the particular statistical program used by the researcher should be identified. Finally, this part should include an overview of what is to follow: "This section is divided into two parts. We will first report the results of the analysis of variance and then the results of the regression analysis."

2. **Description of findings**
   The findings should be tied to the statement of the hypotheses or research questions mentioned in the introduction. The author should clearly state whether the results supported the hypotheses or whether the research questions were answered. Next, any peripheral findings can be reported. Many researchers and journal editors suggest that interpretation and discussion of findings be omitted from this section and that the writer should stick solely to the bare facts. Others think that this section should contain more than numbers, suggesting the implications of the findings as well. In fact, for some short research articles, this section is sometimes called "Findings and Discussion." The choice of what model to follow depends upon the purpose of the report and the avenue of publication.

3. **Tables**
   Tables, charts, graphs, and other data displays should be presented parsimoniously and, if the article is being submitted to a journal, in the proper format. Remember that many readers turn first to the tables and may not read the accompanying text; consequently, tables should be explicit and easily understood by themselves.

DISCUSSION

The last section of a research report is the discussion. The contents of this section are highly variable but the following elements are common.

1. **Summary**
   A synopsis of the main findings of the study often leads off this section.

2. **Implications/discussion/interpretations**
   This is the part of the report that discusses the meaning of the findings. If the findings are in line with current theory and research, the writer should include a statement of how they correspond with what was done in the past. If the findings contradict or do not support current theory, then some explanation for the current pattern of results is provided.

3. **Limitations**
   The conclusions of the study should be tempered by a report of some of its constraints. Perhaps the sample was limited or the response rate was low or the experimental manipulation was not as clean as it could have been. In any case, the researcher should list some of the potential weaknesses of the research.

4. **Suggestions for future research**
   In addition to answering questions, most research projects uncover new questions to be investigated. The suggestions for research should be relevant and practical.

5.5 CHAPTER SUMMARY

A research report is considered a major component of any research study as the research remains incomplete till the report has been presented or written. Writing a report is

the last step in a research study and requires a set of skills somewhat different from those called for in actually conducting research reports vary greatly in length and type depending on the subject. For example, banks and other financial institutions prefer short balance sheet type of tabulations for their annual report. In mathematics, the report may consist of many algebraic notations, whereas a chemist's report may be in the form of symbols and formulae. Once the data collection and analysis work is over, the researcher will start writing the research report. Social and development research reports need to have a logical, clear structure be to the point use simple language, and have a pleasant layout Just as an architect has to draw a layout plan for a house that is being designed, you first have to make an outline for your report. This outline will contain a head, a body, and a tail. The head consists of a description of your problem within its context (the country and research area), the objectives of the study and the methodology followed. Beginning researchers may find the writing style used for research reports awkward or unesthetic, but there is a definite purpose behind the rules governing scientific writing: clarity. Every effort must be made to avoid ambiguity.

5.6 REVIEW QUESTIONS

SHORT ANSWER TYPE QUESTIONS

1. **What do you understand by report writing?**
2. **Write the significance of a report.**
3. **Explain the technique of interpretation.**
4. **Give techniques of interpretation steps.**
5. **What are the precautions in interpretation?**

LONG ANSWER TYPE QUESTIONS

1. **Explain the types of report.**
2. **Elaborate the preliminary pages of the research report.**
3. **Briefly explain the mechanics of writing a research report.**
4. **What kind of methods involve in mechanics of writing a research report?**
5. **What are results mechanics of writing a research report?**

5.7 MULTIPLE CHOICE QUESTIONS

1. **_____ refers to the process of making sense of numerical data that has been collected, analysed and presented.**
   a. Report
   b. Interpretation
   c. Foreword
   d. Preface
2. **The _____ should contain the title, the names of the authors with their designations, the institution that is publishing the report with its logo, the month, and the year of publication.**
   a. cover page

NOTES

b. Foreword
c. Preface
d. Acknowledgements

3. _____ provides the reader a quick overview of the chapters with major sections and sub-sections of your report, and page references so that the reader can go through the report in a different order, or skip certain sections.

a. List of Tables
b. List of Figures
c. Table Of Contents
d. List of Appendices

4. The _____ will contain any additional information that the researcher has collected while carrying out the study.

a. List of Tables
b. List of Figures
c. Table Of Contents
d. List of Appendices

5. The _____ for research should be relevant and practical.

a. Summary
b. Implications
c. Limitations
d. suggestions

6. A _____is usually a short piece of writing found at the beginning of a book or other piece of literature, before the introduction.

a. foreword
b. Preface
c. List of Figures
d. List of Appendices

7. _____ this is often followed by thanks and acknowledgments to people who were helpful to the author during the time of writing.

a. foreword
b. Preface
c. List of Figures
d. List of Appendices

8. The _____ will contain any additional information that the researcher has collected while carrying out the study.

a. Preface
b. List of Figures
c. List of Appendices
d. List of Abbreviations

NOTES 

9. If _____ or acronyms are used in the report, these should be stated in full in the text the first time that they are mentioned.

a. Preface
b. List of Figures
c. List of Appendices
d. List of Abbreviations

10. The _____ section describes the approach used to confront the research problem.

a. Methods
b. Results
c. Discussion
d. Literature Review

♦♦♦♦♦

ANSWER KEY

UNIT I

Question	Answer	Question	Answer
1.	d	6.	a
2.	a	7.	d
3.	b	8.	b
4.	c	9.	c
5.	b	10.	d

UNIT II

Question	Answer	Question	Answer
1.	a	6.	b
2.	b	7.	c
3.	c	8.	b
4.	d	9.	d
5.	a	10.	a

UNIT III

Question	Answer	Question	Answer
1.	a	6.	a
2.	c	7.	d
3.	b	8.	b
4.	d	9.	b
5.	c	10.	c

UNIT IV

Question	Answer	Question	Answer
1.	a	6.	b
2.	b	7.	c
3.	c	8.	d
4.	d	9.	b
5.	a	10.	d

UNIT V

Question	Answer	Question	Answer
1.	b	6.	a
2.	a	7.	b
3.	c	8.	c
4.	d	9.	d
5.	d	10.	a

Research Methodology

Book References

- Ader, H.J., Mellenbergh, G.J., & Hand, D.J. (2008). Advising on Research Methods: A Consultant's Companion. Huizen, The Netherlands: Johannes van kessel Publishing.
- Blalock, HM (1960), Social Statistics, NY: McGraw-Hill.
- Goode, WJ & Hatt, PK (1981). Methods in Social Research. Tokyo: McGraw hill Book Company.
- Kerlinger, FN (1973). Foundations of Behavioural Research. New York: Rinehart and Winston.
- McGuigan, FJ (1990). Experimental Psychology: A Methodological Approach, Englewood Cliffs, New Jersey: Prentice-hall.
- Reichenbach, H (1938). Experience and Prediction. Chicago: University of Chicago press.
- Reichenbach, H (1947). Elements of Symbolic Logic, New York: Macmillan.
- Young, PV (1992). Scientific Social Survey and Research,New Delhi: Prenticehall of India.
- Grinnell, Richard Jr (ed.) 1988, Social Work Research and Evaluation (3rd edition) Itasca, Illinois, F.E. Peacock Publishers.
- Kerlinger, F. N. (1979) Foundation of Behavioural Research, New York, : H 107, Rinehart and Winstem Inc.
- Kuhn, T.S. (1970). The Structure of Scientific Revolutions (2nd edition) Chicago : University of Chicago Press.
- Kumar. R (2006) Research Methodology. New Delhi: Dorling Kingsley
- McGuigan, F.J. (1990): Experimental Psychology: A Methodological Approach. New York : Printice Hall.
- Shaywitz, S.E. (1996): Dyslexia. Scientific American, 275 (5) 98-104. Townsend, J.C. (1953): Introduction to Experimental Method. Tokyo : McGraw Hil
- Aylward, E.H., Richards, T.L., Berwinger, V.W., Nagy, W.E., Field, K.M., Grimme, A.C., Richard, A.L., Thomson, J.B. & Cramer, S.C. (2003). Instructional Treatment Associated with Changes in Brain Activities in Children with Dyslexia. Neurology, 61, 212-219.
- D; Amato, M.R. (1970): Experimental Psychology. Tokoya : McGraw – Hill.