### UNIT -1

**Introduction:** The word research is composed of two syllables "Re" and "Search". "Re" is the prefix meaning 'Again or over again or a new' and "Search" is the latter meaning 'to examine closely and carefully' or 'to test and try'. Together they form, a careful, systematic, patient study and investigation in some field of knowledge undertaken to establish principles / policies.

Meaning of Research: Research can be defined as

1. Search for knowledge

2. Systematic and scientific search for getting relevant answers on any taken up specific topic.

- 3. Scientific enquiry into a subject.
- 4. Research is a movement from the unknown to the known.
- 5. It is the voyage of discovery

**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 formulated hypothesis.

**D.** Slesinger and M. Stephenson in the Encyclopedia 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".

**Bulmer-** Research is primarily committed to establishing systematic, reliable and valid knowledgeabout the social world.

Thus research is the search for knowledge through objective and systematic method for finding solution to a problem.

**Objectives of Research:** The objective of research is to find answers to the questions by applyingscientific procedures. In other words, the main aim of research is to find out the truth which is hidden and has not yet been discovered. Although every research study has its own specific objectives, the research objectives may be broadly grouped as follows:

1. To gain familiarity with a phenomenon or to achieve new insights into it (exploratory or formulative research studies)

2. To portray accurately the characteristics of a particular individual, situation or a group(descriptive research studies)

3. To determine the frequency with which something occurs or with which it is associated with something else (diagnostic research studies)

4. To test a hypothesis of a causal relationship between variables (hypothesis-testing researchstudies)

**Motivation in Research:** The intention of doing research may be one or more of the following:

1. Get a research degree along with its consequential benefits

2. Face the challenges in solving the unsolved problems, i.e., concern over practical problemsinitiates research

- 3. Intellectual joy of doing some creative work
- 4. Service to society
- 5. Get respect.

Factors like directives of government, employment conditions, curiosity about new things, desire understand causal relationships, social thinking and awakening, and the like may as well motivate people to carry research.

#### **Types of Research:**

#### 1. Descriptive Vs Analytical Research

Descriptive research consists of surveys and fact-finding enquiries of different types. The main objective of descriptive research is describing the state of affairs as it prevails at the time of study. The term 'ex post facto research' is quite often used for descriptive research studies in social sciences and business research. The most distinguishing feature of this method is that the researcher has no control over the variables here. He/she has to only report what is happening or what has happened. Majority of the ex post facto research projects are used for descriptive studies in which the researcher attempts to examine phenomena, such as the consumers' preferences, frequency of purchases, shopping, etc. Despite the inability of the researchers to control the variables, ex post facto studies may also comprise attempts by them to discover the causes of the selected problem. The methods of research adopted in conducting descriptive research are survey methods of all kinds, including correlational and comparative methods. Meanwhile in the Analytical research, the researcher has to use the already available facts or information, and analyze them to make a critical evaluation of the subject.

#### Descriptive is fact finding to describe the state of affairs

# Analytical Research is using already available information and analyse to make a critical evaluation of the material.

#### 2. Applied Vs Fundamental Research

An attempt to find a solution to an immediate problem encountered by a firm, an

industry, a business organization, or the society is known as applied research. Researchers engaged in such researches aim at drawing certain conclusions confronting a concrete social or business problem

Fundamental research mainly concerns generalizations and formulation of a theory. In other words, —Gathering knowledge for knowledge's sake is termed pure or basic research (Young in Kothari, 1988). Researches relating to pure mathematics or concerning some natural phenomenonare instances of Fundamental Research. Likewise, studies focusing on human behaviour also fall under the category of fundamental research.

### Applied aims at finding a solution to the problem faced by the society/ organisation. Fundamental Research is concerned with generalization and formulation of a theory

#### 3. Quantitative Vs Qualitative Research

Quantitative research relates to aspects that can be quantified or can be expressed in terms of quantity. It involves the measurement of quantity or amount. Various available statistical and econometric methods are adopted for analysis in such research. Which includes correlation, regressions and time series analysis etc

Qualitative research is concerned with qualitative phenomena, or more specifically, the aspects related to or involving quality or kind. For example, an important type of qualitative research is Motivation Research<sup>4</sup>, which investigates into the reasons for certain human behavior. The main aim of this type of research is discovering the underlying motives and desires of in-depth interviews. The other techniques employed in such research are story completion tests, sentence completion tests, word association tests, and other similar projective methods. Qualitative research is particularly significant in the context of behavioral sciences, which aim at discovering the underlying motives of human behaviour.

#### Quantitative research is based on quantitative measurements of some characteristicsQualitative research is concerned with qualitative phemomenon 4. Conceptual vs. Empirical

The research related to some abstract idea or theory is known as Conceptual Research. Generally, philosophers and thinkers use it for developing new concepts or for reinterpreting the existing ones.

Empirical Research, on the other hand, exclusively relies on the observation or experience with hardly any regard for theory and system. Such research is data based, which often comes up with conclusions that can be verified through experiments or observation. Empirical research is also known as experimental type of research, in which it is important to first collect the facts and theirsources, and actively take steps to stimulate the production of desired information. In this type of research, the researcher first formulates a working hypothesis, and then gathers sufficient facts to prove or disprove the stated hypothesis. He/she formulates the experimental design, which according to him/her would manipulate the variables, so as to obtain the desired information.

The results obtained by using the experimental or empirical studies are considered to be the mostpowerful evidences for a given hypothesis.

#### Conceptual research is related to some abstract ideas

Empirical research is data based research which relies on observation or experience.

5. Some other types of research

• **one-time research or longitudinal research-** depends upon the time of doing research

• **field-setting research or laboratory research or simulation research** depends upon the environment in which research is carried on.

• **clinical or diagnostic research-** in-depth approaches or case study method may beemployed to analyse the basic causal relations

• **Exploratory research-** consist of substantial structure and specific hypotheses to beverified

• **Historical Research-** sources like historical documents, remains, etc. Are utilized to studypast events or ideas.

**Research Approaches:** The above description of the types of research shows that there are two basic approaches to research, viz., quantitative approach and the qualitative approach.

#### Quantitative approach can be further sub-classified into

Inferential research - inferential approach to research is to form a data base from which to infer characteristics or relationships of population. This usually means survey research where a sample of population is studied (questioned or observed) to determine its characteristics, and it is then inferred that the population has the same characteristics

Experimental research- characterized by much greater control over the research environment and in this case some variables are manipulated to observe their effect on other variables

Simulation research- Simulation approach involves the construction of an artificial environment within which relevant information and data can be generated. This permits an observation of the dynamic behavior of a system (or its sub-system) under controlled conditions

<u>Qualitative</u> 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.

#### Significance of Research:

 $\succ$  Research inculcates scientific and inductive thinking and promotes the development of logical habits of thinking.

Research provides the basis for all government policies in our economic

system.

Research has its special significance in solving various operational and planning problems of business and industry

Research is equally important for social scientists in studying social relationships and inseeking answers to various social problems.

(a) To those students who are to write a master's or Ph.D. thesis, research may mean a careerismor a way to attain a high position in the social structure

(b) To professionals in research methodology, research may mean a source of livelihood

(c) To philosophers and thinkers, research may mean the outlet for new ideas and insights

(d) To literary men and women, research may mean the development of new styles and creativework

(e) To analysts and intellectuals, research may mean the generalizations of new theories.

#### **Research Methods versus Methodology:**

Research methods may be understood as all those methods/techniques that are used for conduction of research. Research methods or techniques thus, refer to the methods the researchers use in performing research operations.

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 behindthem

Research Method	Research methodology
Research methods refers to all those methods/techniques that are used for conduction of research. Research methods or techniques	Research methodology is a way to systematically solve the research problem. The scope of research methodology is wider thanthat of research methods
It involves conduction of experiments, tests,surveys etc	Techniques that can be used to conduct the experiments, tests etc
Aims at finding a solution to the researchproblem	Aims at employment of correct procedure to findthe solution

#### **Research and Scientific Method:**

The two terms, research and scientific method, are closely related. Research, as we have already stated, can be termed as "an inquiry into the nature of, the reasons for, and the consequences of any particular set of circumstances, whether these circumstances are experimentally controlled or recorded just as they occur. Further, research implies the researcher is interested in more than particular results; he is interested in the repeatability of the results and in their extension to more complicated and general situations."

The philosophy common to all research methods and techniques, although they may vary considerably from one science to another, is usually given the name of scientific method.

Karl Pearson writes, "The scientific method is one and same in the branches (of science) and that method is the method of all logically trained minds ... the unity of all sciences consists alone in its methods, not its material; the man who classifies facts of any kind whatever, who sees their mutual relation and describes their sequences, is applying the Scientific Method and is a man of science."

Scientific method is the pursuit of truth as determined by logical considerations. The ideal of science is to achieve a systematic interrelation of facts.

The scientific method is, based on the following basic postulates:

- 1. It relies on empirical evidence
- 2. It utilizes relevant concepts
- 3. It is committed to only objective considerations

4. It presupposes ethical neutrality, i.e., it aims at nothing but making only adequate and correctstatements about population objects

5. It results into probabilistic predictions

6. Its methodology is made known to all concerned for critical scrutiny are for use in testing the conclusions through replication

7. It aims at formulating most general axioms or what can be termed as scientific theories

Scientific method implies an objective, logical and systematic method, i.e., a method free from personal bias or prejudice, a method to ascertain demonstrable qualities of a phenomenon capable of being verified, a method wherein the researcher is guided by the rules of logical reasoning, a method wherein the investigation proceeds in an orderly manner and a method that implies internal consistency

#### **Importance of Knowing How Research is Done:**

The importance of knowing how to conduct research are listed below:

> The knowledge of research methodology provides training to new researchers and enables them to do research properly. It helps them to develop disciplined thinking or a 'bent of mind' to objectively observe the field

> The knowledge of doing research inculcates the ability to evaluate and utilize the researchfindings with confidence;

> The knowledge of research methodology equips the researcher with the tools that help him/her to make the observations objectively and

> The knowledge of methodology helps the research consumers to evaluate research and make rational decisions.

#### **Research Process:**

Research process consists of series of actions or steps necessary to effectively carry out researchand the desired sequencing of these steps. The following Figure well illustrates a research process.

The various steps in a research process are as follows:

- formulating the research problem
- extensive literature survey
- developing the hypothesis
- preparing the research design
- determining sample design
- collecting the data
- execution of the project
- analysis of data
- hypothesis testing
- generalisations and interpretation
- preparation of the report or presentation of the results, i.e., formal write-up of conclusions reached.

# **Research Process Flow Chart**



(1) Formulating a Research problem: In research process, the first step a researcher does is formulate a problem and define it properly. Research forms a circle. It starts with a problem and ends with a solution to the problem.

A research problem is anything that a researcher finds unsatisfactory or unsettling, a difficulty of some sort, a state of affairs that needs to be changed, anything that is not working well as it was expected (Creswell, 2009)

A problem statement consists of four parts:

- ➤ the ideal
- the reality or real situation
- the consequences or impacts,
- $\blacktriangleright$  what the study wants to address or the aim of the study.

(2) Extensive Literature Survey: abstracting and indexing Journals, conference proceedings, government reports, books etc.

(3) Development of Working Hypothesis: A tentative assumption made to test its logical or empirical consequences. The role of hypothesis is to guide the researcher by delimiting the area and keep him on right track

Steps to develop a working hypothesis:

Discussions with colleagues and experts about the problem, its origin and the objectives inseeking a solution

Examination of data and records,

Review of similar studies in the area or of the studies on similar problems

 $\succ$  Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals.

(4) **Prepare the Research Design:** state the conceptual structure within which Research will beconducted

Several research designs- Experimental and Non-Experimental Hypothesis testing. Experimental design can be either informal or formal

- (5) Determining Sample Design: Census Survey, Sample Survey Types of Sampling
- Deliberate sampling
- Simple random sampling
- Systematic sampling
- Stratified sampling
- Quota sampling
- Cluster sampling and area sampling
- Multi-stage sampling
- Sequential sampling

#### (6) Data Collection:

- > Observation
- Personal Interview
- Telephone Interview
- > Questionaires
- > Schedules

(7) **Execution of the Project:** In a systematic manner and time

(8) Analysis of Data: 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.

(9) Hypothesis Testing: Various tests, such as Chi square test, t-test, F-test, have been developed by statisticians for testing the hypothesis. 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.

(10) Generalization and Interpretations: If a hypothesis is tested and upheld several times, researcher may 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.

#### (11) Report or thesis writing:

(a) layout

- Preliminary pages
- Main text
- End matter

(b) concise and objective style in simple language avoiding vague expressions such as 'it seems,' 'there may be', and the like.

(c) Charts and illustrations should be used only if they present the information more clearly and forcibly.

(d) various constraints experienced in conducting research operations must be mentioned.

**Criteria of Good Research:** Whatever may be the types of research works and studies, one thing that is important is that they all meet on the common ground of 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 be used.

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 a good research as under:

1. <u>Good research is systematic</u>: 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. Systematic characteristic 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. <u>Good research is logical</u>: 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 which follows from that very premise. In fact, logical reasoning makes research more

meaningful in the context of decision making.

3. <u>Good research is empirical</u>: 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. <u>Good research is replicable</u>: This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

#### Problems Encountered by Researchers in India:

Researchers in India, particularly those engaged in empirical research, are facing several problems.Some of the important problems are as follows:

1. The lack of a scientific training in the methodology of research is a great impediment for researchers in our country. There is paucity of competent researchers. Many researchers take a leap in the dark without knowing research methods. Most of the work, which goes in the name of research is not methodologically sound. Research to many researchers and even to their guides, is mostly a scissor and paste job without any insight shed on the collated materials. The consequence is obvious, viz., the research results, quite often, do not reflect the reality or realities. Thus, a systematic study of research methodology is an urgent necessity. Before undertaking research projects, researchers should be well equipped with all the methodological aspects. As such, efforts should be made to provide short duration intensive courses for meeting this requirement.

2. There is insufficient interaction between the university research departments on one side and business establishments, government departments and research institutions on the other side. A great deal of primary data of non-confidential nature remains untouched/untreated by the researchers for want of proper contacts. Efforts should be made to develop satisfactory liaison among all concerned for better and realistic researches. There is need for developing some mechanisms of a university—industry interaction program so that academics can get ideas from practitioners on what needs to be researched and practitioners can apply the research done by the academics.

3. Most of the business units in our country do not have the confidence that the material supplied by them to researchers will not be misused and as such they are often reluctant in supplying the needed information to researchers. The concept of secrecy seems to be sacrosanct to business organizations in the country so much so that it proves an impermeable barrier to researchers. Thus, there is the need for generating the confidence that thein formation/data obtained from a businessunit will not be misused.

4. Research studies overlapping one another are undertaken quite often for want of adequate information. This results in duplication and fritters away resources. This problem can be solved by proper compilation and revision, at regular intervals, of a list of subjects on which and the places where the research is going on. Due attention should be given toward identification of research problems in various disciplines of applied science which are of immediate concern to theindustries.

5. There does not exist a code of conduct for researchers and inter-university and interdepartmental rivalries are also quite common. Hence, there is need for developing a code of conduct for researchers which, if adhered sincerely, can win over this problem.

6. Many researchers in our country also face the difficulty of adequate and timely secretarial assistance, including computerial assistance. This causes unnecessary delays in the completion of research studies. All possible efforts be made in this direction so that efficient secretarial assistance is made available to researchers and that too well in time. University Grants Commission must play a dynamic role in solving this difficulty.

7. Library management and functioning is not satisfactory at many places and much of the time and energy of researchers are spent in tracing out the books, journals, reports, etc., rather than in tracing out relevant material from them.

8. There is also the problem that many of our libraries are not able to get copies of old and new Acts/Rules, reports and other government publications in time. This problem is felt more in libraries which are away in places from Delhi and/or the state capitals. Thus efforts should be made for the regular and speedy supply of all governmental publications to reach our libraries.

9. There is also the difficulty of timely availability of published data from various government and other agencies doing this job in our country. Researcher also faces the problem on account of the fact that the published data vary quite significantly because of differences in coverage by the concerning agencies.

10. There may, at times, take place the problem of conceptualization and also problems relating to the process of data collection and related things

#### UNIT 2

#### **Defining the Research Problem:**

Research Problem:

The first and foremost stage in the research process is to select and properly define the research problem. A researcher should first identify a problem and formulate it, so as to make it amenable or susceptible to research. In general, a research problem refers to an unanswered question that a researcher might encounter in the context of either a theoretical or practical situation, which he/shewould like to answer or find a solution to. A research problem is generally said to exist if the following conditions emerge (Kothari, 1988):

i. There should be an individual or an organization, say X, to whom the Problem can be attributed. The individual or the organization is situated in an environment Y, which is governed by certain uncontrolled variables Z

ii. There should be at least two courses of action to be pursued, say A1 and A2. These courses of action are defined by one or more values of the controlled variables. For example, the number of items purchased at a specified time is said to be one course of action.

iii. There should be at least two alternative possible outcomes of the said courses of action, say B1 and B2. Of them, one alternative should be preferable to the other. That is, at least one outcome should be what the researcher wants, which becomes an objective.

iv. The courses of possible action available must offer a chance to the researcher to achieve the objective, but not the equal chance. Therefore, if P(Bj / X, A, Y) = 10 represents the probability of the occurrence of an outcome Bj when X selects Aj in Y, then  $P(B1 / X, A1, Y) \neq P(B1 / X, A2, Y)$ . Putting it in simple words, it means that the choices must not have equal efficiencies for the desired outcome.

Above all these conditions, the individual or organization may be said to have arrived at the research problem only if X does not know what course of action to be taken is the best. In other words, X should have a doubt about the solution. Thus, an individual or a group of persons can be said to have a problem if they have more than one desired outcome. They should have two or morealternative courses of action, which have some but not equal efficiency. This is required for probing the desired objectives, such that they have doubts about the best course of action to be taken. Thus, the components of a research problem may be summarized as:

a) There should be an individual or a group who have some difficulty or problem.

b) There should be some objective(s) to be pursued. A person or an organization who wants nothing cannot have a problem.

c) There should be alternative ways of pursuing the objective the researcher wants to pursue. This implies that there should be more than one alternative means available to the researcher. This is because if the researcher has no choice of alternative means, he/she would not have a problem.

d) There should be some doubt in the mind of the researcher about the choice of alternative means. This implies that research should answer the question relating to the relative efficiency or suitability of the possible alternatives.

#### Selecting the Problem:

Criteria for selection of research problem depend on the following characteristics.

- Personal Inclination.
- Resources Availability.
- Relative Importance.
- Researcher Knowledge
- Practicality: Practicality is also responsible for the selection.
- Time-lines of the Problem.
- Urgency.

Personal Inclination: The chief motivation in the way of selecting research problem is the personal inclination of the researcher. If a researcher has personal interest in the topic, he would select that problem for his research work

Resources Availability: During the selection, a researcher will see to the resources available. If these resources like money, time, accommodation and transport are available to the selection place, then the selection of the problem is easy.

Data Availability: If the desired data is available to the researcher, then the problem would be selected.

Urgency: Urgency is a pinpoint in the way of the selection of research problem. Urgent problem must be given priority because the immediate solution can benefit the people.

Feasibility: Feasibility is also an important factor for the selection of the research problem. The researcher qualification, training and experience should match the problem.

Area Culture: The culture of the area for which a researcher conducts his research is also responsible for the selection of research problem.

**Necessity of Defining the Problem:** A problem clearly stated is a problem half solved. This statement signifies the need for defining a research problem. The problem to be investigated must be defined unambiguously for that will help to discriminate relevant data from the irrelevant ones. A proper definition of research problem will enable the researcher to be on the track whereas an ill-defined problem may create hurdles. Questions like: What data are to be collected? What characteristics of data are relevant and need to be studied? What relations are to be explored. Whattechniques are to be used for the purpose? and similar other questions crop up in the mind of the researcher who can well plan his strategy and find answers to all such questions only when the research problem has been well defined. Thus, defining a research problem properly is a

prerequisite for any study and is a step of the highest importance.

#### **Technique Involved in Defining a Problem:**

The technique for the purpose involves the undertaking of the following steps generally one after the other:

- (i) statement of the problem in a general way
- (ii) understanding the nature of the problem
- (iii) surveying the available literature
- (iv) developing the ideas through discussions and
- (v) rephrasing the research problem into a working proposition.

#### **Reviewing the literature**:

A literature review is a survey of scholarly sources on a specific topic. It provides an overview ofcurrent knowledge, allowing you to identify relevant theories, methods, and gaps in the existing research.

Writing a literature review involves finding relevant publications (such as books and journal articles), critically analyzing them, and explaining what you found. There are five key steps:

- Search for relevant literature
- Evaluate sources
- Identify themes, debates and gaps
- Outline the structure
- Write your literature review

A good literature review doesn't just summarize sources – it analyzes, synthesizes, and criticallyevaluates to give a clear picture of the state of knowledge on the subject.

**Reviewing the literature**: -Essential preliminary task in order to acquaint yourself with the available body of knowledge in your area of interest.

Literature review is integral part of entire research process and makes valuable contribution to every operational step. -Reviewing literature can be time-consuming, daunting and frustrating, butis also rewarding. Its functions are:

a. **Bring clarity and focus to your research problem** - The process of reviewing the literature helps you to understand the subject area better and thus helps you to conceptualise your research problem clearly and precisely. It also helps you to understand the relationship between your research problem and the body of knowledge in the area

b. **Improve your methodology** - A literature review tells you if others have used procedures and methods similar to the ones that you are proposing, which procedures

and methods have worked well for them, and what problems they have faced with them. Thus you will be better positioned to select a methodology that is capable of providing valid answer to your research questions

c. **Broaden your knowledge-** It ensures you to read widely around the subject area in which youintend to conduct your research study. As you are expected to be an expert in your area of study, it helps fulfill this expectation. It also helps you to understand how the findings of your study fit into the existing body of knowledge.

d. **Contextualise your findings-** How do answers to your research questions compare with what others have found? What contribution have you been able to make in to the existing body of knowledge? How are your findings different from those of others? For you to be able to answer these questions, you need to go back to your literature review. It is important to place your findings in the context of what is already known in your field of enquiry.

#### **Procedure for reviewing the literature:**

i) **search for existing literature in your area of study-** To effectively search for literature in your field of enquiry, it is imperative that you have in mind at least some idea of broad subject area and of the problem you wish to investigate, in order to set parameters for your search. -Next compile a bibliography for this broad area. Sources are:

1. **books-** BOOKS comprise a central part of any bibliography. Advantage-material published generally is of good quality and the findings are integrated with other research to form a coherentbody of knowledge. Disadvantage-material is not completely up to date, as it can take a few yearsbetween the completion of a work and publication in the form of a book. Search for books in your area of interest, prepare a final list, locate these books in the libraries or borrow from other sources. Examine their content, if contents are not found to be relevant to your topic, delete it from your reading list.

2. **journals** -Journals provide you with the most up-to-date information, even though there is a gap of two to three years between the completion of a research project and the publication in a journal. As with books, you need to prepare a list of journals for identifying literature relevant to your study. This can be done as follows: -locate the hard copies of the journal that are appropriate to your study; - use the internet - look at the index of research abstracts in the relevant field to identify and read the articles. Whichever method you choose, first identify the journals you want to look at in more detail for your review of literature. Select the latest issue, examine its content page to see if there is an article of relevance to your research topic. If you feel a particular article is of relevance to you, read its abstract. If you think you are likely to use it, photocopy or prepare a summary andrecord it for reference for later use.

ii) **review the literature selected-** After identifying books and articles as useful, the next step is to start reading them critically to pull together themes and issues that are associated. If you do not have a theoretical framework of themes in mind to start with, use separate sheets of paper for each article or book. Once you develop a rough framework, slot the findings from the material so far reviewed into that framework,

using a separate sheet of paper for each theme of that framework. As you read further, go on slotting the information where it logically belongs under the theme so far developed. You may need to add more themes as you go. Read critically with particular reference to the following aspects:

• Note whether the knowledge relevant to your theoretical framework is confirmed beyond doubt.

• Note the theories put forward, the criticisms of these and their basis, the methodologies adopted and the criticisms of them.

• Examine to what extent the findings can be generalized to other situations. Ascertain the areas in which little or nothing is known-the gaps that exist in the body of knowledge.

iii) **develop a theoretical framework-** As you have limited time it is important to set parametersby reviewing the literature in relation to some main themes pertinent to your research topic. As you start reading the literature, you will realize that it deals with a number of aspects that have a direct `and indirect bearing on your research topic. Use these aspects as a basis for developing your theoretical framework. Until you go through the literature you cannot develop a theoretical framework and until you have developed a theoretical framework, you cannot effectively review the literature. Literature pertinent to your study may deal with two types of information: - universal; - more specifi c( i.e. local trends or specific program) In writing about such information you should start with the general information, gradually narrowing down to the specific.

#### iv) develop a conceptual framework.

Writing up the literature reviewed: In order to comply with the first function of literature review

i.e. to provide theoretical background to your study: -List the main themes that have emerged while reading literature. -Convert them into subheadings. These subheadings should be precise, descriptive of the theme in question, and follow a logical progression. -Now, under each subheading, record the main findings with respect to the theme in question, highlighting the reasons for and against an argument if they exist, and identify gaps and issues. In order to comply with the second function of literature review i.e. contextualising the findings of your study- requires you to very systematically compare your findings with those made by others. Quote from these studies to show how your findings contradict, confirm or add to them. It places your findings in the context of what others have found out. This function is undertaken when writing about yourfindings i.e. after analysis of your data.

#### **Research Design:**

**Meaning of Research Design:** The most important step after defining the research problem is preparing the design of the research project, which is popularly known as the research design. A research design helps to decide upon issues like what, when, where, how much, by what means etc. With regard to an enquiry or a research study. A research design is the arrangement of conditions for collection and analysis of data in a

manner that aims to combine relevance to the research purpose with economy in procedure. In fact, research design is the conceptual structure within which research is conducted; it constitutes the blueprint for the collection, measurement and analysis of data (Selltizetal, 1962). Thus, research design provides an outline of what the researcher is going to do in terms of framing the hypothesis, its operational implications and the final data analysis. Specifically, the research design highlights decisions which include:

1. The nature of the study2. The purpose of the study

3. The location where the study would be conducted 4. The nature of data required

5.From where the required data can be collected 6.What time period the study would cover 7.The type of sample design that would be used

8. The techniques of data collection that would be used 9. The methods of data analysis that would be adopted and

10. The manner in which the report would be prepared

In view of the stated research design decisions, the overall research design may be divided into the following (Kothari 1988):

a) The sampling design that deals with the method of selecting items to be observed for the selectedstudy b) The observational design that relates to the conditions under which the observations are to be made

c) The statistical design that concerns with the question of how many items are to be observed, andhow the information and data gathered are to be analyzed and

d) The operational design that deals with the techniques by which the procedures specified in thesampling, statistical and observational designs can be carried out

#### Features of a Good Design:

The important features of Research Design may be outlined as follows:

i. It constitutes a plan that identifies the types and sources of information required for theResearchproblem

ii. It constitutes a strategy that specifies the methods of data collection and analysis which wouldbe adopted and

iii. It also specifies the time period of research and monetary budget involved in conducting the study, which comprise the two major constraints of undertaking any research

#### Important Concepts Relating to Research Design:

1. Dependent and independent variables: A magnitude that varies is known as a variable. The concept may assume different quantitative values like height, weight, income etc. Qualitative variables are not quantifiable 17 in the strictest sense of the term. However, the qualitative phenomena may also be quantified in terms of the presence or absence of the attribute(s) considered. The phenomena that assume different values quantitatively even in decimal points are known as continuous variables. But all variables need not be continuous. Values that can be expressed only in integer values are called \_non-continuous variables. In statistical terms, they are also known as discrete variables. For example, age is a continuous variable, whereas the number of children is a non-continuous variable. When changes in one variable depend upon the changes in other variable or variables, it is known as a dependent or endogenous variable, and the

variables that cause the changes in the dependent variable are known as the independent or explanatory or exogenous variables. For example, if demand depends upon price, then demand is a dependent variable, while price is the independent variable. And, if more variables determine demand, like income and price of the substitute commodity, then demand also depends upon them in addition to the price of original commodity. In other words, demand is a dependent variable which is determined by the independent variables like price of the original commodity, income and price of substitutes.

2. Extraneous variables: The independent variables which are not directly related to the purpose of the study but affect the dependent variables, are known as extraneous variables. For instance, assume that a researcher wants to test the hypothesis that there is a relationship between children's school performance and their self-confidence, in which case the latter is an independent variable and the former, a dependent variable. In this context, intelligence may also influence the school performance. However, since it is not directly related to the purpose of the study undertaken by the researcher, it would be known as an extraneous variable. The influence caused by the extraneous variable(s) on the dependent variable is technically called the 'experimental error'. Therefore, a research study should always be framed in such a manner that the influence of extraneous variables on the dependent variable/s is completely controlled, and the influence of independent variable/s is clearly evident. Control, One of the most important features of a good research design is to minimize the effect of extraneous variable(s). Technically, the term 'control' is used when a researcher designs the study in such a manner that it minimizes the effects of extraneous variables. The term 'control' is used in experimental research to reflect the restrain in experimental conditions.

3. <u>Confounded relationship</u>: The relationship between the dependent and independent variables is said to be confounded by an extraneous variable, when the dependent variable is not free from itseffects.

4. <u>Research hypothesis:</u> When a prediction or a hypothesized relationship is tested by adopting scientific methods, it is known as research hypothesis. The research hypothesis is a predictive statement which relates to a dependent variable and an independent variable. Generally, a research hypothesis must consist of at least one dependent variable and one independent variable. Whereas, the relationships that are assumed but not to be tested are predictive statements that are not to be objectively verified, thus are not classified as research hypotheses.

5. <u>Experimental and non-experimental hypothesis testing research</u>: When the objective of a research is to test a research hypothesis, it is known as hypothesis testing research. Such research may be in the nature of experimental design or nonexperimental design. The research in which the independent variable is manipulated is known as experimental hypothesis-testing research', whereas the research in which the independent Variable is not manipulated is termed as 'non- experimental hypothesis-testing research'.

6. <u>Experimental and control groups</u>: When a group is exposed to usual conditions in an experimental hypothesis-testing research, it is known as 'control group'. On the other

hand, when the group is exposed to certain new or special condition, it is known as an 'experimental group'. In the afore-mentioned example, Group A can be called as control group and Group B as experimental group. If both the groups, A and B are exposed to some special feature, then both the groups may be called as 'experimental groups'. A research design may include only the experimental group or both the experimental and control groups together.

7. <u>Treatments</u>: Treatments refer to the different conditions to which the experimental and controlgroups are subject to.

8. Experiment: Experiment refers to the process of verifying the truth of a statistical hypothesis relating to a given research problem. For instance, an experiment may be conducted to examine the yield of a certain new variety of rice crop developed. Further, Experiments may be categorized into two types, namely. 'absolute experiment' and 'comparative experiment'. If a researcher wishes to determine the impact of a chemical fertilizer on the yield of a particular variety of rice crop, then it is known as absolute experiment. Meanwhile, if the researcher wishes to determine the impact of chemical fertilizer as compared to the impact of bio-fertilizer, then the experiment isknown as a comparative experiment.

9. <u>Experimental unit(s)</u>: Experimental units refer to the pre-determined plots, characteristics or the blocks, to which different treatments are applied.

#### **Different Research Designs:**

Types of research design: There are different types of research designs.

- (1) Exploratory research design
- (2) Descriptive and diagnostic research design
- (3) Hypothesis-testing research design.

1. Exploratory research design: The Exploratory Research Design is known as formulative research design. The main objective of using such a research design is to formulate a research problem for an in-depth or more precise investigation, or for developing a working hypothesis from an operational aspect. The major purpose of such the discoverv studies is of ideas and insights. Therefore. sucharesearchdesignsuitableforsuchastudyshouldbeflexibleenoughto provide opportunity for considering different dimensions of the problem understudy. The inbuilt flexibility in research design is required as the initial research problem would be transformed into a more precise one in the exploratory study, which in turn may necessitate changes in the research procedure for collecting relevant data.

2. **Descriptive and diagnostic research design**: A Descriptive Research Design is concerned with describing the characteristics of a particular individual or a group. Meanwhile, a diagnostic research design determines the frequency with which a variable occurs or its relationship with another variable. In other words, the study analyzing whether a certain variable is associated with another comprises a diagnostic research study. On the other hand, a study that is concerned with specific predictions or

with the narration of facts and characteristics related to an individual, groupor situation, are instances of descriptive research studies. Generally, most of the social research design falls under this category. As a research design, both the descriptive and diagnostic studies share common requirements, hence they are grouped together. However, the procedure to be used and the research design need to plan carefully. The research design must also make appropriate provision for protection against bias and thus maximize reliability, with due regard to the completion of the research study in an economical manner. The research design in such studies should be rigid and not flexible

3. **Hypothesis-Testing research design:** Hypothesis-Testing Research Designs are those in which the researcher tests the hypothesis of causal relationship between two or more variables. These studies require procedures that would not only decrease bias and enhance reliability, but also facilitate deriving inferences about the causality. Generally, experiments satisfy such requirements. Hence, when research design is discussed in such studies, it often refers to the designof experiments.

**Basic Principles of Experimental Designs :** Professor Fisher has enumerated three principles of experimental designs:

(1) **the Principle of Replication**-the experiment should be repeated more than once. Thus, each treatment is applied in many experimental units instead of one. By doing so the statistical accuracy of the experiments is increased.

(2) **the Principle of Randomization-** The Principle of Randomization provides protection, when we conduct an experiment, against the effect of extraneous factors by randomization. In other words, this principle indicates that we should design or plan the experiment in such a way that the variations caused by extraneous factors can all be combined under the general heading of "chance."

(3) **Principle of Local Control**- the extraneous factor, the known source of variability, is made to vary deliberately over as wide a range as necessary and this needs to be done in such a way that the variability it causes can be measured and hence eliminated from the experimental error. This means that we should plan the experiment in a manner that we can perform a two-way analysis of variance, in which the total variability of the data is divided into three components attributed to, the extraneous factor and experimental error. In other words, according to the principle of local control, we first divide the field into several homogeneous parts, known as blocks, and then each such block is divided into parts equal to the number of treatments. Then the treatments are randomly assigned to these parts of a block. Dividing the field into several homogenous parts is known as 'blocking'. In general, blocks are the levels at which we hold an extraneous factor fixed, so that we can measure its contribution to the total variability of the data by means of a two-way analysis of variance. In brief, through the principle of local control we can eliminate the variabilitydue to extraneous factor(s) from the experimental error.

#### **Important Experimental Designs.**

#### (a) Informal experimental designs:

(i) <u>Before-and-after without control design</u>-A single test group or area is selected and the dependent variable is measured. the treatment is then introduced and then the dependent variable is measured again.

the effect of the treatment: the level of the phenomenon after the treatment-the level of the phenomenon before the treatment.



(ii) <u>After-only with control design-</u>Two groups or areas (test area and control area) are selected and the treatment is introduced into the test area only.

Test area:	Treatment introduced	Level of phenomenon after treatment (Y)
Control area:		Level of phenomenon without
	Treatment Effect = (Y) – (Z)	ueaunent (Z)

#### (iii) Before-and-after with control design-



#### (b) Formal experimental designs:

(i) <u>Completely randomized design (C.R. Design).</u>



#### (ii) Randomized block design (R.B. Design).

	Very low I.Q.	Low I.Q.	Average I.Q.	High I.Q.	Very high I.Q.
	Student A	Student B	Student C	Student D	Student E
Form 1	82	67	57	71	73
Form 2	90	68	54	70	81
Form 3	86	73	51	69	84
Form 4	93	77	60	65	71

#### (iii) Latin square design (L.S. Design).

	FERTILITY LEVEL					
		1	11	m	ĩ٧	٧
	х,	А	в	с	D	Е
	X2	в	с	D	Е	A
Seeds differences	×,	с	D	Е	A	в
	х,	D	Е	A	в	С
	X,	Е	А	в	с	D

#### (iv) Factorial designs

	Experimental Variable			
Control variables	Treatment A	Treatment B		
Level I	Cell 1	Cell 3		
Level II	Cell 2	Cell 4		

#### UNIT -3

#### **Design of Sampling:**

**Introduction:** The use of sampling in making inferences about a population is possible and has been in operation right from beginning. When one has to make an inference about a lot of large size and it is not practicable to examine each individual unit, then few units of the lot are examined and on the basis of the information of those units, one makes decisions about whole lot. For example, a person would like to purchase a bag of rice may examine a handful of rice from the bag and on the basis of that he/she makes his/her decision about the purchase of full bag.

**Population** - A group of individuals having same characteristics in same surrounding is known aspopulation

**Census** - In census, we study about each and every unit of the population. Population means total units of investigation area. In census, whole group related to investigation is investigated and the information are collected, i.e. Census of population of a country, Census of import and export, etc

**Sample** - A finite subset of statistical individuals in a population is called a sample and the number of individuals in a sample is called the sample size.

**Sample Design:** Essentials of Sampling For obtaining the unbiased and real result by a samplingmethod, a sample should have the following factors (characteristics):

1. Homogeneity The nature of each and every unit of the population should not contain much difference. If two or more samples are selected then they should be similar in nature not in their response/output.

2. Representativeness The sample should represent all the characteristics of the population that canbe possible only when the selection of items or units has been done unbiased and each and every unit have an equal probability of chance to be selected in the sample.

3. Independency Each and every unit of the population should be independent. In other words, theselection of a unit in the sample should not be dependent on the selection of other units.

4. Adequacy The number of units or elements which are to be selected in the sample should be sufficient. If the sample size is not sufficient then results cannot be reliable. The more the sampleunits in the sample, more reliable results would occur.

**PRINCIPLE STEPS IN SAMPLE SURVEY** The main steps involved in the planning and execution of a sample survey are under the following heads:

1. Objectives The objective of the survey must be defined in clear and concrete terms. Generally, in survey a investigation team is not quite clear in mind as to what they want and how they are going to use the results. Some of the objectives may be immediate and some

far-reaching. The investigator should take care of these objectives with the available resources in terms of money, manpower and the time limit required for the availability of the survey.

2. Defining the Population The population from which sample is chosen should be defined in clear and unambiguous terms. The geographical, demographic and other boundaries of the population must be specified so that no ambiguity arises regarding the coverage of the survey.

3. Sampling Frame and Sampling Units The sampling unit is the ultimate unit to be sampled for the purpose of the survey. The sampling units must cover the entire population and they must be distinct, unambiguous and non-overlapping in the sense that every element of the population belongs to one and only one sampling unit. In a Socio economic survey, whether a family or a member of a family is to be the ultimate sampling unit. Once the sampling units are defined, one must see whether a sampling frame which is a list of all the units in the population, is available. The construction of the frame is often one of the major practical problem since it is the frame which determines the structure of the sample survey. The list of units have to be carefully scrutinized and examined to ensures that it is free from duplicity or incompleteness and are up-to- date. A good frame is hard to come by and only good experience helps to construct a good frame.

4. Selection of Proper Sampling Design This is the most important step in planning a sample survey. There is a group of sampling designs (to be discussed later) and selection of the proper one is an important task. The design should take into account the available resources and the time- limit, if any, besides the degree of accuracy desired. The cost and precision should also be considered before the final selection of sampling design.

5. Method of Collecton of Data For collection of data, either the interview method or the mail questionnaire method is to be adopted. Although the later method is less costly but there is a large scope of non-response in it. In the cases, where the information is to be collected by observation they must decide upon the method of measurement.

6. Data to be Collected Collection of data must be done in conformity with the objectives of the survey and the nature of the data. After it is decided upon, one must prepare a questionnaire or a schedule of enquiry. A schedule or a questionnaire contains a list of items of which information issought, but the exact form of the questions to be asked is not standardized but left to the judgment of the investigators. A questionnaire should be in a specified order. The questions should be clear, brief, collaborative, non offending and unambiguous and to the point so that not much scope of speculation is left on the part of the respondent or interviewer.

7. Field Work Organization Field work, itself has several stages and so it is to be well organized. The different stages include training the field workers, supervising the field workers, etc. It is absolutely essential that the personnel should be Introduction to Sample Surveys thoroughly trained in locating the sample units, the methods of collection of required data before starting the field word. The success of a survey to a great extent depends upon the reliable field work. Inspection after field work by the adequate supervisors should also be performed.

8. Summary and Analysis of Data This is the last step wherein inference is to be made on the basisof collected data. This step again consists of the following steps:

a) The filled in questionnaires should be carefully scrutinized to find out whether the data furnished re plausible and consistent;

b) Depending upon the quantity of data, a hand-tabulation or machine tabulation is to be drawn;

c) After the data has been properly scrutinized, edited and tabulated, a very careful statisticalanalysis is to be made; and

d) Finally a report incorporating detailed statement of the different stages of the survey should be prepared. In the presentation of the result, it is advisable to report technical aspects of the design

**Sampling and Non-sampling Errors:** The errors involved in the collection, processing and analysis of data may be broadly classified under the following two heads:

- 1. Sampling Error
- 2. Non-sampling Error

1. **Sampling Error** The error which arises only in sample survey is termed as sampling error. This error arises because in sample survey a part of the population is only studied. This is the reason why sampling error is absent in census. The main factors of sampling error are:

• Some of the bias is introduced by the use of defective sampling techniques for the selection f a sample

• Substitution of a non-selected a convenient unit of the population in place of a selected unitto which the investigation is difficult leads to some biases in the sample survey;

• Bias due to defective demarcation of sampling units, particularly in area/filed survey; and

• Constant errors due to improper choice of the statistics for estimating the population parameters.

2 **Non-Sampling Error** The non-sampling error arises at the stages of observation, ascertainment and processing of the data. This is the reason why the non-sampling error presents Statistical Techniques in both the census and the sample survey. Non-sampling error can occur at every stage of the planning or execution of census or sample survey. Non sampling errors arise due to the following factors:

• Data specification being inadequate and inconsistent with respect to the objective of thestudy;

- Error due to location of the units and actual measurement of the characteristics;
- Error due to ill designed questionnaire;
- Lack of trained and qualified investigators and lack of adequate supervisory staff;
- Errors due to lack of correct responses furnished by the respondents;
- Non-response biases occur if full information is not obtained on all the sampling units;

• If the objectives of the survey are not stated clearly, it may result in inclusion of the units which are not to be included and exclusion of the units which are to be included in the sample;

• Due to error in various operations of data processing such as editing and coding of the

responses, punching of cards, tabulation and summarizing the observation made in the survey; and

• The errors may be committed during presentation and printing the results of the survey.

Sample Survey versus Census Survey: The advantages of sampling over complete census maybe outlined as follows:

1. Sampling requires less time and labor than census because only a part of the population has tobe examined. The sampling results also can be analysed much faster;

2. Sampling usually results in reduction in cost in terms of money and man powers. The total cost of the sample survey is expected to be much smaller than a complete census;

3. There is generally a greater scope in a sample survey than in census. Some inquiries may requirehighly trained personnel or specialized equipment for collection of data, then the census may be inconceivable;

4. In some cases a complete census is ruled out by the nature of the population. If there is a population which is infinite and/or hypothetical, then sampling is the only option;

5. A sample survey gives data of better quality than a complete census, because in a sample surveyit may be possible to use better resources than complete census;

6. If the population is too large, as for example, trees in a jungle, leaves in a tree i.e. we are left with no option but to resort to sampling; and

7. If testing is destructive, then complete enumeration is impracticable and sampling design is the only method to be used in such cases. For example, testing the breaking strength of a chalk, testing of lifetime of an electrical bulb, etc.

**Types of Sampling Designs** : According to the method of selection of sample, the sampling schemes can be categorised as follows:

- 1. Non-probability sampling
- 2. Probability or random sampling and
- 3. Mixed sampling.

1 **Non-Probability Sampling-** In this method, the sample is selected with a definite purpose in view and the choice of the sampling units depends entirely on the discretion and judgment of the investigator. While selecting a sample, investigator tries to include each and every characteristics of population in sample. Non-Probability Sampling scheme can be classified as:

• Purposive Sampling In this sampling the sample is selected with definite purpose in view and the choice of sampling units depends entirely on the discretion of the surveyor. This sampling suffers from drawback of favoritism and nepotism of the surveyor.

• Judgment Sampling In judgment sampling respondents are selected on the judgment of the surveyor with the hope that they will meet requirements of the study. The underling assumptions are that the respondent truly represents the entire population. To find out the potential guide for the food and catering technology a researcher go to the teachers of Hotel Management Department may be the example of judgment sampling.

• Deliberate Sampling In deliberate sampling, deliberate selection of sample is made so thatany important unit could not be leftout.

• Convenience Sampling In convenience sampling method, a surveyor selects the sample at his/her own convenience, often as the study is being conducted. Convenience Statistical Techniques sampling is based on assumption that the target population is homogeneous and the individuals selected and interviewed yields similar information with regard to the characteristics under study. If persons selected from restaurants to collect the information about quality of the food, service, etc. are supposed to represent the population of food takers. Such a sampling is known as convenience sampling

• Quota Sampling If the cost of selected random samples in each stratum is very high in stratified sampling then the sampling units are assigned in a quota (fixed number of units) in each stratum and the actual selection of units is left at the decision of the surveyor.

Merits of Non-Probability Sampling

1. This method of sampling is very simple;

2. After sample size determination with the help of planning, a suitable sample may possibly beobtained; and

3. Important units or members may be included in the sample.

Demerits of Non-Probability Sampling

1. Predetermined view of selector effects the selection of sample which impure the result. This effect does directly or indirectly on the process;

2. There is no place for probability in selection of units therefore sampling error cannot be obtained;

3. There is no guaranty of validity of the results from the sample selected by this method; and

4. The attitude and biasedness of the investigator also affect the selection of sample that's why the results obtained by this method are not reliable scientifically.

2 **Random or Probability Sampling** The technique of random sampling is of fundamental importance in the application of Statistics. Probability sampling is the scientific method of selecting samples accordingly to some laws of chance in which each unit in the population has some definite pre-assigned probability of being selected in the sample.

Merits of Random or Probability Sampling

1. No Plan for Selection There is no need to make any detailed plan for the selection of units.

2. Less Expensive In this method, money, time and hard work are very less.

3. Unbiased In this method there is no space for any biasedness. Every unit has same chance of selection

4. Inspection of Purity Inspection of purity of one sample can be done by other sample. In this method measure of statistical error can also be done.

5. Random Selector has not to use his mind. He selects units at random.

6. True Representation of Population In this method real characteristics can be represented through sample because it is based on the law of statistical regularity and law of inertia. In real, itbecomes a small part of population.

Demerits of Random or Probability Sampling

1. Inappropriate This method is not appropriate where some units are so important to be includednecessarily in the sample.

2. Less Representative It may be possible that sample could not represent the population if sampleis not sufficiently large.

3. Less Independency This method is useless if the units of the population are dependent.

#### Difference between Probability and Non-probability Sampling

1. In non-probability sampling the selection of units are pre-decided whether in probability sampling is based on chances.

2. Non-probability sampling is biased but probability sampling is unbiased.

3. In non-probability sampling, the errors are of cumulative in nature whereas in probability sampling errors are less Statistical Techniques

4. If a sample, from a population with homogeneous and important units, is to be selected then non-probability sampling is appropriate where as probability sampling is used in various kinds of population.

3. Mixed Sampling If the samples are selected partly according to some laws of chance and partly according to a fixed rule, they are called mixed samples and the method of selecting such samples is known as mixed sampling. The merits of this sampling are the mixture of the merits of both sampling. Selection of units is more reliable in this method because that is the representation of various stages of population. In mixed sample no important characteristics is left which is to be selected in the sample.

Measurement and Scaling: Qualitative and Quantitative Data

	Qualitative	Quantitative
Conceptual	Concerned with understanding human behaviour from the informant's perspective	Concerned with discovering facts about social phenomena
	Assumes a dynamic and negotiated reality	Assumes a fixed and measurable reality
Methodological	Data are collected through participant observation and interviews	Data are collected through measuring things
	Data are analysed by themes from descriptions by informants	Data are analysed through numerical comparisons and statistical inferences
	Data are reported in the language of the informant	Data are reported through statistical analyses

#### SAMPLING DESIGN

research does not exist without sampling. Every research study requires the selection of some kind of sample. It is the life blood of research.

Any research study aims to obtain information about the characteristics or parameters of a population. A population is the aggregate of all the elements that share some common set of characteristics and that comprise the universe for the purpose of the research problem. In other words, population is defined as the totality of all cases that conform to some designated specifications. The specification helps the researcher to define the elements that ought to be included and to be excluded. Sometimes, groups that are of, interest to the researcher may be significantly smaller allowing the researcher to collect data from all the elements of population. Collection of data from the entire population is referred to as census study. A census involves a complete enumeration of the elements of a population.

Collecting data from the aggregate of all the elements (population) in case of, the number of elements being larger, would sometimes render the researcher incur huge costs and time. It may sometimes be a remote possibility. An alternative way would be to collect information from a portion of the population, by taking a sample of elements from the population and the on the basis of information collected from the sample elements, the characteristics of the population is inferred. Hence, Sampling is the process of selecting units (e.g., people, organizations) from a population of interest so that by studying the sample we may fairly generalize our results back to the population from which they were chosen.

While deciding on the sampling, the researcher should clearly define the target population without allowing any kind of ambiguity and inconsistency on the boundary of the aggregate setof respondents. To do so, the researcher may have to use his wisdom, logic and judgment to define the boundary of the population keeping with the objectives of the study.

#### **TYPES OF SAMPLING PLANS**

Sampling techniques are classified into two broad categories of probability samples or non-probability samples.

#### **Probability Sampling Techniques**

Probability samples are characterised by the fact that, the sampling units are selected by chance. In such case, each member of the population has a known, non-zero probability of being selected. However, it may not be true that all sample would have the same probability of selection, but it is possible to say the probability of selecting any particular sample of a given size. It is possible that one can calculate the probability that any given population element would be included in the sample. This requires a precise definition of the target population as well as the sampling frame.

Probability sampling techniques differ in terms of sampling efficiency which is a concept that refers to trade off between sampling cost and precision. Precision refers to the level of uncertainty about the characteristics being measured. Precision is inversely related to sampling errors but directly related to cost. The greater the precision, the greater the cost and there should be a tradeoff between sampling cost and precision. The researcher is required to design the mostefficient sampling design in order to increase the efficiency of the sampling.

Probability sampling techniques are broadly classified as simple random sampling, systematic sampling, and stratified sampling.

#### Simple Random Sampling

This is the most important and widely used probability sampling technique. They gain much significance because of their characteristic of being used to frame the concepts and arguments in statistics. Another important feature is that it allows each element in the population to have a known and equal probability of selection. This means that every element is selected independently of every other element. This method resembles lottery method where a in a systemnames are placed in a box, the box is shuffled, and the names of the winners are then drawn outin an unbiased manner.

Simple random sampling has a definite process, though not, so rigid. It involves compilation of a sampling frame in which each element is assigned a unique identification number. Random numbers are generated either using random number table or a computer to determine which elements to include in the sample. For example, a researcher is interested in investigating the behavioural pattern of customers while making a decision on purchasing a computer. Accordingly, the researcher is interested in taking 5 samples from a sampling frame containing 100 elements. The required sample may be chosen using simple random sampling technique by arranging the 100 elements in an order and starting with row 1 and column 1 of random table, and going down the column until 5 numbers between 1 and 100 are selected. Numbers outside this range are ignored. Random number tables are found in every statistics book. It consists of a randomly generated series of digits from 0 - 9. To enhance the readability of the numbers, row is given. The researcher may a the space between every 4 digit and between every 10 begin reading from anywhere in the random number table, however, once started the researcher should continue to read across the row or down a column. The most important feature of simple random sampling is that it facilitates representation of the population by the sample ensuring

that the statistical conclusions are valid.

#### Systematic Sampling

This is also another widely used type of sampling technique. This is used because of its ease and convenience. As in the case of simple random sampling, it is conducted choosing a random starting point and then picking every element in succession from the sampling frame. The sample interval, i, is determined by dividing the population size N by the sample size n and rounding to the nearest integer.

Consider a situation where the researcher intends to choose 10 elements from a population of 100. In order to choose these 10 elements, number the elements from one to 100. Within 20 population elements and a sample of size 10, the number is 10/100 = 1/10, meaning that one element in 10 will be selected. The sample interval will, therefore, be 10. This means that after a the random start from any point in the random table, the researcher has to choose every 10 element.

Systematic sampling is almost similar to simple random sampling in that each population element has a known and equal probability of selection. However, the difference lies in that simple random sampling allows only the permissible samples of size n drawn have a known and equal probability of selection. The remaining samples of size n have a zero probability of being selected

#### Stratified sampling

Stratified sampling is a two-way process. It is distinguished from the simple random sampling and systematic sampling, in that:

a) It requires division of the parent population into mutually exclusively and exhaustive subsets;

b) A simple random sample of elements is chosen independently from each group or subset.

Therefore, it characterises that, every population element should be assigned to one and only stratum and no population elements should be omitted. Next, elements are selected from each stratum by simple random sampling technique. Stratified sampling differs from quota sampling in that the sample elements are selected probabilistically rather than based on convenience or on judgemental basis.

Strata are created by a divider called the stratification variable. This variable divides the population into strata based on homogeneity, heterogeneity, relatedness or cost. Sometimes, more than one variable is used for stratification purpose. This type of sampling is done in orderto get homogenous elements within each strata and, the elements between each strata should have a higher degree of heterogeneity. The number of strata to be formed for the research is left to the discretion of the researcher, though, researchers agree that the optimum number of strata may be 6.

The reasons for using stratified sampling are as follows:

a) it ensures representation of all important sub-populations in the sample;

- b) the cost per observation in the survey may be reduced;
- c) it combines the use of simple random sampling with potential gains in precision;
- d) estimates of the population parameters may be wanted for each sub-population and;
- e) increased accuracy at given cost.

#### **Non-probability Sampling Methods**

Non-probability sampling does not involve random selection. It involves personal judgement of the researcher rather than chance to select sample elements. Sometimes this judgement is imposed by the researcher, while in other cases the selection of population elements to be includes is left to the individual field workers. The decision maker may also contribute to including a particular individual in the sampling frame. Evidently, non probability sampling does not include elements selected probabilistically and hence, leaves an degree of "sampling error" associated with the sample.

Sampling error is the degree to which a sample might differ from the population. Therefore, while inferring to the population, results could not be reported plus or minus the sampling error. In non-probability sampling, the degree to which the sample differs from the population remains unknown However, we canno come to a conclusion that sampling error is an inherent of non probability sample.

Non-probability samples also yield good estimates of the population characteristics. Since, inclusion of the elements in the sample are not determined in a probabilistic way, the estimates obtained are not statistically projectable to the population.

The most commonly used non-probability sampling methods are convenience sampling, judgment sampling, quota sampling, and snowball sampling.

#### **Convenience Sampling**

Convenience samples are sometimes called accidental samples because the elements included in the sample enter by "accident". It is a sampling technique where samples are obtained from convenient elements. This refers to happening of the element at the right place at the right time, that is, where and when the information for the study is being collected. The selection of the respondents is left to the discretion of the interviewer. The popular examples of convenience sampling include (a) respondents who gather in a church (b) students in a class room (c) mall intercept interviews

without qualifying the respondents for the study (d) tear-out questionnaire included in magazinesand (e) people on the street. In the above examples, the people may not be qualified respondents, however, form part of the sample by virtue of assembling in the place where the researcher is conveniently placed.

Convenience sampling is the least expensive and least time consuming of all sampling techniques. The disadvantage with convenience sampling is that the researcher would have no way of knowing if the sample chosen is representative of the target population.

Judgement Sampling This is a form of convenience sampling otherwise called as purposive

sampling because the sample elements are chosen since it is expected that they can serve the research purpose. The sample elements are chosen based on the judgement that prevails in the researcher"s mind about the prospective individual. The researcher may use his wisdom to conclude that a particular individual may be a representative of the population in which one is interested.

The distinguishing feature of judgment sampling is that the population elements are purposively selected. Again, the selection is not based on that they are representative, but rather because they

can offer the contributions sought. In judgement sampling, the researcher may be well aware of the characteristics of the prospective respondents, in order that, he includes the individual in the sample. It may be possible that the researcher has ideas and insights about the respondent"s requisite experience and knowledge to offer some perspective on the research question.

#### Quota Sampling

Quota sampling is another non-probability sampling. It attempts to ensure that the sample chosen by the researcher is a representative by selecting elements in such a way that the proportion of the sample elements possessing a certain characteristic is approximately the same as the proportion of the elements with the characteristic in the population.

Quota sampling is viewed as two-staged restricted judgemental sampling technique. The first stage consists of developing control categories, or quotas, of population elements. Control characteristics involve age, sex, and race identified on the basis of judgement. Then the distribution of these characteristics in the target population is determined. For example, the researcher may use control categories in that, he/she intends to study 40% of men and 60% of women in a population. Sex is the control group and the percentages fixed are the quotas.

In the second stage, sample elements are selected based on convenience or judgement. Once the quotas have been determined, there is considerable freedom to select the elements to be included in the sample. For example, the researcher may not choose more than 40% of men and 60% of women in the study. Even if the researcher comes across qualified men after reaching the 40% mark, he/she would still restrict entry of men into the sample and keep searching for women till the quota is fulfilled.

#### Snowball Sampling

This is another popular non-probability technique widely used, especially in academic research. In this technique, an initial group of respondents is selected, usually at random. After being interviewed, these respondents are asked to identify others who belong to the target population of interest. Subsequent respondents are selected based on the information provided by the selected group members. The group members may provide information based on their understanding about the qualification of the other prospective respondents. This method involves probability and non-probability methods. The initial respondents are chosen by a random method and the subsequent respondents are chosen by non-probability methods.

### **UNIT - 4**

#### **Data Collection:**

Methods of Data Collection: There are two types of data **Primary Data**— collected for the first time **Secondary Data**—those which have already been collected and analysed by someone else.

#### **Methods of Primary Data Collection**

**OBSERVATION METHOD**: Commonly used in behavioural sciences It is the gathering of primary data by investigator's own direct observation of relevant people, actions and situations without asking from the respondent. e.g. • A hotel chain sends observers posing as guests into its coffee shop to check on cleanliness and customer service. • A food service operator sends researchers into competing restaurants to learn menu items prices, check portion sizes and consistency and observe point-of purchase merchandising. • A restaurant evaluates possible new locations by checking out locations of competing restaurants, traffic patterns and neighborhood conditions. Observation can yield information which people are normally unwilling or unable to provide. e.g. Observing numerous plates containing uneaten portions the same menu items indicates that food is not satisfactory.

Types of Observation:

- 1. Structured for descriptive research
- 2. Unstructured—for exploratory research
- 3. Participant Observation
- 4. Non- participant observation
- 5. Disguised observation

Limitations: - feelings, beliefs and attitudes that motivate buying behaviour and infrequent behaviour cannot be observed. - expensive method Because of these limitations, researchers often supplement observation with survey research.

SURVEY METHOD Approach most suited for gathering descriptive information.

Structured Surveys: use formal lists of questions asked of all respondents in the same way.

**Unstructured Surveys**: let the interviewer probe respondents and guide the interview according to their answers. Survey research may be Direct or Indirect. Direct Approach: The researcher asks direct questions about behaviours and thoughts. e.g. Why don't you eat at MacDonalds? Indirect Approach: The researcher might ask: "What kind of people eat at MacDonald's?" From the

response, the researcher may be able to discover why the consumer avoids MacDonald's. It may suggest factors of which the consumer is not consciously aware.

ADVANTAGES: -can be used to collect many different kinds of information -Quick and low cost as compared to observation and experimental method.

LIMITATIONS: -Respondent's reluctance to answer questions asked by unknown interviewers about things they consider private. -Busy people may not want to take the time -may try to help by giving pleasant answers -unable to answer because they cannot remember or never gave a thought to what they do and why -may answer in order to look smart or well informed.

**CONTACT METHODS:** Information may be collected by Mail, Telephone, Personal interview <u>Mail Questionnaires:</u>

Advantages: -can be used to collect large amounts of information at a low cost per respondent. - respondents may give more honest answers to personal questions on a mail questionnaire -no interviewer is involved to bias the respondent's answers.

-convenient for respondent's who can answer when they have time - good way to reach people who often travel

Limitations: -not flexible

-take longer to complete than telephone or personal interview

-response rate is often very low –

researcher has no control over who answers.

Telephone Interviewing: - quick method -

more flexible as interviewer can explain questions not understood by the respondent –depending on respondent's answer they can skip some Qs and probe more on others – allows greater sample control –

response rate tends to be higher than mail

Drawbacks: -Cost per respondent higher -

Some people may not want to discuss personal Qs with interviewer

-Interviewer's manner of speaking may affect the respondent's answers – Different interviewers may interpret and record response in a variety of ways –under time pressure ,data may be entered without actually interviewing

<u>Personal Interviewing:</u> It is very flexible and can be used to collect large amounts of information. Trained interviewers are can hold the respondent's attention and are available to clarify difficult questions. They can guide interviews, explore issues, and probe as the situation requires. Personal interview can be used in any type of questionnaire and can be conducted fairly quickly.

Interviewers can also show actual products, advertisements, packages and observe and record their reactions and behaviour. This takes two forms-

Individual- Intercept interviewing Group -

Focus Group Interviewing Intercept interviewing: Widely used in tourism research. -allows researcher to reach known people in a short period of time. - only method of reaching people whose names and addresses are unknown -involves talking to people at homes, offices, on the street, or in shopping malls. -interviewer must gain the interviewee's cooperation -time involved may range from a few minutes to several hours( for longer surveys compensation may be offered) --involvesthe use of judgmental sampling i.e. interviewer has guidelines as to whom to "intercept", such as 25% under age 20 and 75% over age 60

Drawbacks: -Room for error and bias on the part of the interviewer who may not be able to correctly judge age, race etc. -Interviewer may be uncomfortable talking to certain ethnic or age groups.

Focus Group Interviewing: It is rapidly becoming one of the major research tool to understand people's thoughts and feelings. It is usually conducted by inviting six to ten people to gather for a few hours with a trained moderator to talk about a product, service or organization. The meeting isheld in a pleasant place, and refreshments are served to create a relaxed environment. The moderator needs objectivity, knowledge of the subject and industry, and some understanding of group and consumer behaviour. The moderator starts with a broad question before moving to more specific issues, encouraging open and easy discussion to bring out true feelings and thoughts. At the same time, the interviewer focuses the discussion, hence the name focus group interviewing. - often held to help determine the subject areas on which questions should be asked in a later, large-scale, structured-direct interview Comments are recorded through note taking or videotaped and studied later to understand consumer' buying process. This method is especially suited for managers of hotels and restaurants, who have easy access to their customers. e.g. Some hotel managers often invite a group of hotel guests from a particular market segment to have a free breakfast with them. Managers get the chance to meet the guests and discuss what they like about the hotel and what the hotel could do to make their stay more enjoyable and comfortable. The guests appreciate this recognition and the manager gets valuable information. Restaurant managers use the same approach by holding discussion meetings over lunch or dinner.

Drawbacks: -Cost: may cost more than telephone survey -Sampling: group interview studies keep small sample size to keep time and cost down, therefore it may be difficult to generalize from theresults. - Interviewer bias.

**EXPERIMENTAL METHOD** Also called Empirical Research or Cause and Effect Method, it is a data-based research, coming up with conclusions which are capable of being verified with observation or experiment. Experimental research is appropriate when proof is sought that certain variables affect other variables in some way. e.g. -Tenderisers (independent variable) affect cooking time and texture of meat( dependent variable). - The effect of substituting one ingredient in whole or in part for another such as soya flour to flour for making high protein bread. -Developrecipes to use products. Such research is characterised by the experimenter's control over the variables under study and the deliberate manipulation of one of them to study its effects. In such aresearch, it is necessary to get at facts first hand, at their source, and actively go about doing certain things to stimulate the production of desired information. -Researcher

must provide self with a working hypothesis or guess as to the probable results. - Then work to get enough facts (data) to prove or disprove the hypothesis. -He then sets up experimental designs which he thinks will manipulate the persons or the materials concerned so as to bring forth the desired information. Evidence gathered through experimental or empirical studies today is considered to be the most powerful support possible for a given hypothesis.

#### **Testing of Hypotheses:**

A hypothesis test is a formal way to make a decision based on statistical analysis. <u>Hypotheses</u>. Hypothesis tests are tests about a population parameter .

Characteristics of hypothesis: Hypothesis must possess the following characteristics:

(i) Hypothesis should be clear and precise. If the hypothesis is not clear and precise, the inferences drawn on its basis cannot be taken as reliable.

(ii) Hypothesis should be capable of being tested. In a swamp of untestable hypotheses, many a time the research programmes have bogged down. Some prior study may be done by researcher in order to make hypothesis a testable one. A hypothesis "is testable if other deductions can be made from it which, in turn, can be confirmed or disproved by observation."

(iii) Hypothesis should state relationship between variables, if it happens to be a relational hypothesis.

(iv) Hypothesis should be limited in scope and must be specific. A researcher must remember that narrower hypotheses are generally more testable and he should develop such hypotheses.

(v) Hypothesis should be stated as far as possible in most simple terms so that the same is easily understandable by all concerned. But one must remember that simplicity of hypothesis has nothing to do with its significance.

(vi) Hypothesis should be consistent with most known facts i.e., it must be consistent with a substantial body of established facts. In other words, it should be one which judges accept as being the most likely.

(vii) Hypothesis should be amenable to testing within a reasonable time. One should not use even an excellent hypothesis, if the same cannot be tested in reasonable time for one cannot spend a life-time collecting data to test it.

(viii) Hypothesis must explain the facts that gave rise to the need for explanation. This means that by using the hypothesis plus other known and accepted generalizations, one should be able to deduce the original problem condition. Thus hypothesis must actually explain what it claims to explain; it should have empirical reference.

<u>The null hypothesis</u> (H0) is a statement involving equality  $(=, \leq, \geq)$  about a population parameter.

<u>The alternative hypothesis (Ha)</u> is a statement that contradicts the null hypothesis. The alternative hypothesis is what we conclude is true if the experimental results lead us to conclude that the null

hypothesis (our assumption) is false. The alternative hypothesis must not involve equality (6=, ). The exact statement of the null and alternative hypotheses depend on the claim that you are testing.

We decide to reject the null hypothesis if the sample outcome contradicts our assumption. The logic is as follows:

• To make this decision we calculate the p-value := the probability of our sample outcome or something more extreme occuring ASSUMING the null hypothesis is true.

If the p-value is very small, our sample outcome was very unlikely (a "rare event") based on ourassumption, so we reject our assumption.

Recall that our assumption is based on the null hypothesis, so rejecting the assumption is the same rejecting the null hypothesis.

The smaller the p-value, the larger (in absolute value) the z/t-score, so an "unusual" z/t-score givesus a hint about the p-value.

Example : Skittles. Our null hypothesis (assumption) is that a bag of Skittles contains 95% orangeSkittles (and the rest are green), because that is written on the label. We reach in a hand and pull out a big handful of Skittles. They are all green! • Describe the p-value in words.

We can't get the exact p-value, but what do you estimate it to be? • Should we decide to "reject" or "not to reject" our assumption (null hypothesis) that the bag contains 95% orange Skittles? Disclaimer: This scenario is a complete fabrication. Skittles are fabulous.

. Decision. How small must the p-value be to reject the null hypothesis? This can change and is set by the significance level,  $\alpha$ , which will be provided in each problem. Common significance levels are 0.01, or 1 100, and 0.05, or 5 100.

The decision process is as follows:

• If p-value  $< \alpha$ , reject the null hypothesis. (Our outcome contradicts the assumption.)

• If p-value  $\geq \alpha$ , do not reject the null hypothesis. (Our outcome seems reasonable based on the assumption.)

Skipper, p 6 This silly little rhyme is from the Triola textbook "Elementary Statistics" and has been helpful to many students: "If the p is low, the null must go. If the p is high, the null will fly."

Confidence intervals allowed us to find ranges of reasonable values for parameters we were interested in. Hypothesis testing will let us make decisions about specific values of parameters orrelationships between parameters.

- 1.  $\pi$  population proportion
- 2.  $\mu$  population mean

3.  $\mu 1 - \mu 2$  - difference in population means (example: compare average heights of men and women)

4.  $\mu$ d - population mean difference (for paired data) (example: compare average heights of fathersand sons)

5.  $\pi 1 - \pi 2$  - difference in population proportions

You will need to be able to distinguish these settings from each other. Notably, to distinguish proportion settings from mean settings, think about the question being asked. Are responses to the question of interest yes or no? If so, you are dealing with proportions. If the response is a numericalvalue, you are dealing with means.

In hypothesis testing, we use sample data to choose between two competing hypotheses. Think of it like a jury trial. There are two options: innocent and guilty. You assume innocence until shown guilty beyond a reasonable doubt. In hypothesis testing, there are 2 choices, the null hypothesis and the alternative hypothesis. You assume the null hypothesis is true until the alternative is shown beyond "chance". Hypothesis - a claim about a population characteristic (parameter)

Null Hypothesis - the status quo - initially assumed true

Alternative Hypothesis - the researcher's proposal - what you hope to show Main idea:

Reject the null hypothesis in favor of the alternative only with convincing/significant evidence. We do NOT say that we accept the alternative, only that we have significant evidence to reject thenull. This is because we could have made a mistake (see below). Forms of Hypotheses:

Null Hypothesis: H0 : population parameter = some hypothesized valueAlternative Hypothesis: HA : population parameter is **not equal to** that same hypothesized value (two-sided) ORHA : population parameter > that same hypothesized value (one-sided to the right) OR HA : population parameter < that same hypothesized value (one-sided to the left)

#### **ELEMENTS/TYPES OF ANALYSIS**

Analysis means the computation of certain indices or measures along with searching for patterns of relationship that exist among the data groups. Analysis, particularly in case of survey or experimental data, involves estimating the values of unknown parameters of the population and testing of hypotheses for drawing inferences. Analysis may, therefore, be categorised as <u>descriptive analysis and inferential analysis</u> (Inferential analysis is often known as statistical analysis).

"Descriptive analysis is largely the study of distributions of one variable. This study provides us with profiles of companies, work groups, persons and other subjects on any of a multiple of characteristics such as size. Composition, efficiency, preferences, etc.". this sort of analysis may be in respect of one variable (described as unidimensional analysis), or in respect of two variables (described as bivariate analysis) or in respect of more than two variables (described as multivariate analysis). In this context we work out various measures that show the size and shape of a distribution(s) along with the study of measuring relationships between two or more variables. We may as well talk of correlation analysis and causal analysis.

Correlation analysis studies the joint variation of two or more variables for determining the amount of correlation between two or more variables. Causal analysis is concerned with the study of how one or more variables affect changes in another variable. It is thus a study of functional relationships existing between two or more variables. This analysis can be termed as regression analysis

Causal analysis is considered relatively more important in experimental researches, whereas in most social and business researches our interest lies in understanding and controlling relationships between variables then with determining causes per se and as such we consider correlation analysis as relatively more important. In modern times, with the availability of computer facilities, there has been a rapid development of multivariate analysis which may be defined as "all statistical methods which simultaneously analyse more than two variables on a sample of observations"

Usually the following analyses are involved when we make a reference of multivariate analysis:

(a) Multiple regression analysis: This analysis is adopted when the researcher has one dependent variable which is presumed to be a function of two or more independent variables. The objective of this analysis is to make a prediction about the dependent variable based on its covariance with all the concerned independent variables.

(b) Multiple discriminant analysis: This analysis is appropriate when the researcher has a single dependent variable that cannot be measured, but can be classified into two or more groups on the basis of some attribute. The object of this analysis happens to be to predict an entity's possibility of belonging to a particular group based on several predictor variables.

(c) Multivariate analysis of variance (or multi-ANOVA): This analysis is an extension of two way ANOVA, wherein the ratio of among group variance to within group variance is worked out on a set of variables.

(d) Canonical analysis: This analysis can be used in case of both measurable and non-measurable variables for the purpose of simultaneously predicting a set of dependent variables from their joint covariance with a set of independent variables.

Inferential analysis is concerned with the various tests of significance for testing hypotheses in order to determine with what validity data can be said to indicate some conclusion or conclusions. It is also concerned with the estimation of population values. It is mainly on the basis of inferential analysis that the task of interpretation (i.e., the task of drawing inferences and conclusions) is performed.

#### The chi-squared test

This is a versatile non-parametric test.

Goodness of Fit: whether a sample fits an expected distribution (of arbitrary shape). Test for Independence: are paired observations on two categorical variables independent.

- Variable X: Gender [Male,Female]
- Variable Y: Voting Preference [Lab,Con,Lib]
- Does gender affect voting preference?

Goodness of Fit - Consider throwing a die many times to see whether or not it's fair. We get a frequency distribution. We expect this to be a uniform distribution: an equal number of 1s, 2s, 3s,etc.

#### **DATA COLLECTION**

The next step in the research process after identifying the type of research the researcher intends to do is the deciding on the selection of the data collection techniques. The data collection technique is different for different types of research design. There are predominantly two types of data: (i) the primary data and (ii) the secondary data.

Primary data is one a researcher collects for a specific purpose of investigating the research problem at hand. Secondary data are ones that have not been collected for the immediate study at hand but for purposes other than the problem at hand. Both types of data offer specific advantages and disadvantages.

a) Secondary data offer cost and time economies to the researcher as they already exist in variousforms in the company or in the market.

b) It is feasible for a firm to collect.

c) Since they are collected for some other purposes, it may sometimes not fit perfectly into the problem defined.

d) The objectives, nature and methods used to collect the secondary data may not be appropriate to the present situation.

Mostly secondary data helps to:

- a) Identify the problem.
- b) Better define the problem.
- c) Develop an approach to the problem.
- d) Formulate an appropriate research design by identifying the key variables.
- e) Answer certain research questions and formulate hypotheses.
- f) Interpret the primary data more in-depth.

#### SECONDARY DATA

Secondary data are the data that are in actual existence in accessible records, having been alreadycollected and treated statistically by the persons maintaining the records. In other words, secondary data are the data that have been already collected, presented tabulated, treated with necessary statistical techniques and conclusions have been drawn. Therefore, collecting secondary data doesn't mean doing some original enumeration but it merely means obtaining data that have already been collected by some agencies, reliable persons, government departments, research workers, dependable organisations etc. Secondary data are easily obtainable from reliable records, books, government publications and journals.

When once primary data have been originally collected, moulded by statisticians or statistical

machinery, then it becomes secondary in the hands of all other persons who may be desirous of handling it for their own purpose or studies. It follows, therefore, that primary and secondary data are demarcated separately and that the distinction between them is of degree only. It a person 'X' collects some data originally, then the data is primary data to 'X' whereas the same data when used by another person 'Y' becomes secondary data to 'Y'.

#### SOURCES OF SECONDARY DATA

The following are some of the sources of secondary data:

- 1. Central and State government publications.
- 2. Publications brought out by international organisation like the UNO, UNESCO, etc.
- 3. Foreign government publications.
- 4. Official publications as well as reports of municipalities, district parishads, etc.

5. Reports and publications of commissions - like U.G.C. education commission, tariff commission, chambers of commerce, co-operative societies, trade associations, banks, stock exchanges, business houses etc.

6. Well-know newspapers and journals like the *Economic Times, The Financial Express, Indian Journal of Economics, Commerce, Capital, Economical Eastern Economist,* etc. Further Year Books such as *Times of India Year Book, Statesman's Year Book* also provide valuable data.

7. Publications brought out by research institutions, universities as well as those published by research workers give considerable secondary data.

8. Through the Internet/website sources.

Though the given list of secondary data cannot be said to be thorough or complete, yet it can be pointed out that it fairly indicates the chief sources of secondary data. Also, besides the above mentioned data there are a number of other important sources, such as records of governments in various departments, unpublished manuscripts of eminent scholars, research workers, statisticians, economists, private organisations, labour bureaus and records of business firms.

#### **TYPES OF SECONDARY DATA**

Secondary data are of two types. Data that are originated from within the company are called as internal data. If they are collected for some other purpose, they are internal secondary data. This poses significant advantage as they are readily available in the company at low cost. The most convenient example internal secondary data is the figures relating sales of the company. Important internal source of secondary data is database marketing, Database marketinginvolves the use of computers to capture and track customer profiles and purchase details. The information about customer profile would serve as the foundation for marketing programmes or as an internal source of information related to preference of customer"s preference of a particularproduct.

Published external secondary data refers to the data available without the company. There is such a pool of published data available in the market that it is sometimes easy to underestimate what available and thereby bypass relevant information. Several sources of external data are available. They are:

#### General Business Data

Guides or small booklets containing information about a particular trade or business.

Directories are helpful for identifying individuals or organisations that collect specific data.

Indexes used to locate information on a particular topic in several different publications by using an index.

Non-governmental statistical data refers to published statistical data of great interest to researchers. Graphic and statistical analyses can be performed on these data to draw meaning inference.

#### **Government Sources**

Census data is a report published by the Government containing information about the population of the country.

Other Government publications may be pertaining to availability of train tickets just before it leaves.

#### Computerised Databases

Online databases are databases consisting of data pertaining to a particular sector (e.g., banks) that is accessed with a computer through a telecommunication network

 $\Box$  Internet databases are available in internet portals that can be accessed, searched, and analysed on the internet.

□ Offline databases are databases available in the form of diskettes and CD-ROM disks.

Bibliographic databases comprises of citations in articles published in journals, magazines, newspapers etc.

Numeric databases contain numerical and statistical information. For example, time series data about stock markets.

Directory databases provide information on individuals, organisations and service. E.g. Getit Yellow pages.

Special-purpose databases are databases developed online for a special purpose.

*External Data-syndicated* In response to the growing need for data pertaining to markets, consumer etc., companies have started collecting and selling standardised data designed to serve the information needs of the shared by a number of organisations. Syndicated data

sources can be further classified as (a) consumer data (b) retail data (c) wholesale data (d) industrial data (e) advertising evaluation data and (f) media and audience data.

Consumer data relates to data about consumers purchases and the circumstances surrounding the purchase.

Retail data rely on retailing establishments for their data. The data collected focus on the products or services sold through the outlets and / or the characteristics of the outlets themselves.

□ Wholesale data refers to data on warehouse shipment data to estimate sales at retail.

 $\Box$  Industrial data refers to substantially more syndicated data services available to consumer goods manufacturers than to industrial goods suppliers.

 $\Box$  Advertising evaluation data refers to money spent each year on media such as magazines and television with the expectation that these expenditures will result in sales.

#### VERIFICATION OF SECONDARY DATA

Before accepting secondary data it is always necessary to scrutinize it properly in regard to its accuracy and reliability. It may perhaps happen that the authorities collecting a particular type of data may unknowingly carry out investigations using procedures wrongly. Hence it is always necessary to carry out the verification of the secondary data in the following manner:

(i) Whether the organization that has collected the data is reliable.

(ii) Whether the appropriate statistical methods were used by the primary data enumerators and investigators.

(iii) Whether the data was collected at the proper time.

#### **COLLECTION OF PRIMARY DATA**

By primary data we mean the data that have been collected originally for the first time. In other words, primary data may be the outcome of an original statistical enquiry, measurement of facts or a count that is undertaken for the first time. For instance data of population census is primary. Primary data being fresh from the fields of investigation is very often referred to as raw data. In the collection of primary data, a good deal of time, money and energy are required.

#### QUESTIONNAIRE

A questionnaire is defined as a formalised schedule for collecting data from respondents. It may be called as a schedule, interview form or measuring instrument.

Measurement error is a serious problem in questionnaire construction. The broad objective of a questionnaire include one without measurement errors. Specifically, the objectives of a questionnaire are as follows:

a) It must translate the information needed into a set of specific questions that the respondents can and will answer.

b) The questions should measure what they are supposed to measure.

c) It must stimulate the respondents to participate in the data collection process. The respondents should adequately motivated by the virtual construct of the questionnaire.

d) It should not carry an ambiguous statements that confuses the respondents.

#### **Questionnaire Components**

A questionnaire consists typically of five sections. They are:

- a) Identification data
- b) Request for cooperation
- c) Instruction
- d) Information sought
- e) Classification of data

a) Identification data occupation is the first section of a questionnaire where the researcher intends to collect data pertaining to the respondent's name, address and phone number.

b) Request for cooperation refers to gaining respondent's cooperation regarding the datacollection process.

c) Instruction refers to the comments to the respondent regarding how to use the questionnaire.

d) The information sought form the major portion of the questionnaire. This refers to theitems relating to the actual area of the study.

e) Classification data are concerned with the characteristics of the respondent.

#### **OBSERVATION METHODS**

This is another type of method used when the researcher feels that survey type of methods may not be so relevant in data collection. In subjective issues, respondents need to be observed rather than asked lest biases and prejudices happen in their response. Observation method may be either

structured or unstructured. Structured observation method involves having a set of items to be observed and how the measurements are to be recorded. In unstructured observation, the observer monitors all aspects of the phenomena that seem relevant to the problem at hand. In this context, the observer may have an open mind to study the persons or object.

#### DATA PREPARATION AND ANALYSIS

#### DATA ANALYSIS AND INTERPRETATION

A researcher's important function is the appropriate interpretation of different types of statistical data with the help of his tools. The preliminary statistical work consists of collection, classification, tabulation, presentation and analysis of data. The most important part of the statistical work consists in the proper use of the statistical tools in the interpretation of data.

The most commonly used tools are 'Mean, Median, Mode; Geometric Mean, Measures of Dispersion such as Range; Mean Deviation, Standard Deviation and also other measures such as Coefficient of Correlation, Index Numbers etc. It is necessary to note that technical interpretation of data has to be combined with a high degree of sound judgement, statistical experience, skill and accuracy. After all figures do not lie, they are innocent. But figures obtained haphazardly, compiled unscientifically and analyzed incompetently would lead to general distrust in statistics by the citizens. It should be understood that "statistical methods are the most dangerous tools in the hands of an expert".

#### DATA EDITING AND CODING

Authenticity and relevance of a research investigation is based on the assurance of error-free qualitative reliability of the collected data. Data processing has to be carried out in an appropriate manner. Processing comprises the task of editing, coding classification and tabulation.

In spite of a careful collection by a researcher, there may be a possibility for errors of omission and commission arising and it is for this purpose that the process of editing becomes necessary.

The editor, while examining certain responses of the respondents, may find some mistakes in the form of incomplete, vague or irrelevant answers. Such inconsistent answers have to be eliminated or suitably and reasonably modified. Further, there should be no room for fictitious data to creep in. Hence the editor has to take care to see that the authenticity of the data is in a perfect shape.

For the purpose of classification of the data into meaningful and useful classes, the procedure of coding has to be used. This procedure would be advantageous in dealing with the data having a number of characteristics. Also, a large volume of data can be processed accurately.

Manual processing and analysis can be carried out by using measures of central tendency, dispersion, correlation regression and other statistical methods if the volume of data is not very large.

In case a researcher is confronted with a very large volume of data then it is imperative to use 'computer processing'. For this purpose necessary statistical packages such as SPSS etc. may be

used. Computer technology can prove to be a boon because a huge volume of complex data can be processed speedily with greater accuracy.

# INTERPRETATION OF DATA IN GENERAL MANAGEMENT AND SOCIAL SCIENCES

Data pertaining to economic, psychological sociological or managerial phenomena necessarily requires appropriate interpretation through the use of analytical procedures based on inductive or deductive logical reasoning. Further, proper statistical methods will have to applied for scientific analysis. Depending upon the nature of the data which may be nominal, ordinal, interval or ratio level, a researcher has to carefully evaluate the appropriateness and precision in the use of 'Parametric' or 'Non-parametric' tests of hypothesis. It may be noted that generally the nominal level data is weak whereas the ratio level data is comparatively strong.

Statistical analysis can be classified as (i) descriptive and (ii) inferential.

Descriptive data provides quantitative information for analytical interpretation for instance: with respect to the wage distribution of 500 workers in a factory, we can calculate various measures of central tendency, dispersion, skewness etc. Inferential data relates to statistical inference where conclusions are drawn on the basis of samples taken randomly from a population, which is assumed to be normal. Population parameters are estimated on the basis on the basis of sample statistics.

Depending upon the nature of researcher's problem, relevant sampling methods are used for obtaining data. However, for the purpose of hypothesis testing, parametric or non-parametric tests may be used depending upon the fact whether the assumptions in regard to population are based on 'distribution' or 'distribution-free characteristics'.

#### **INTERPRETATION OF FINANCIAL RATIOS**

Financial ratio analysis is a study of ratios between various items or groups of items in financial statements. Financial ratios can be broadly classified into the following categories:

- 1. Liquidity ratios
- 2. Leverage ratios
- 3. Turnover ratios
- 4. Profitability ratios
- 5. Valuation ratios

#### **Liquidity Ratios**

Liquidity refers to the ability of a firm to meet its obligations in the short run, usually one year. Liquidity ratios are generally based on the relationship between current assets and current liabilities.

The important liquidity ratios are:

(a) *Current Ratio:* Current assets include cash, current investments, debtors, inventories (stocks), loans and advances, and prepaid expenses. Current liabilities represent liabilities that are expected to mature in the next twelve months. These comprise (i) loans, secured or unsecured, that are due in the next twelve months and (ii) current liabilities and provisions. The current ratio thus measures the ability of the firm to meet its current liabilities.

(b) *Acid-Test Ratio (also called the quick ratio):* Quick assets are defined as current assets excluding inventories.

It is a fairly stringent measure of liquidity. It is based on those current assets, which are highly liquid. Inventories are excluded because they are deemed to be the least liquid component of the current assets.

(c) *Cash Ratio:* Because cash and bank balance and short term marketable securities are the mostliquid assets of a firm.

#### Leverage Ratios

Financial leverage refers to the use of debt finance. While debt capital is a cheaper source of finance, it is also a riskier source of finance. Leverage ratios help in accessing the risk arising from the use of debt capital. Two types of ratios are commonly used to analyze financial leverage:

- (i) Structural ratios
- (ii) Coverage ratios

Structural ratios are based on the proportions of debt and equity in the financial structure of the firm. Coverage ratios show the relationship between debt serving commitments and sources for meeting these burdens.

The important structural ratios are:

(a) *Debt-Equity Ratio:* It shows the relative contributions of creditors and owners.

The numerator of this ratio consists of all debt, short-term as well as long-term, and the denominator consists of net worth plus preferential capital.

(b) *Debt-Assets Ratio:* It measures the extent to which borrowed funds support the firm's assets.

The numerator of this ratio includes all debts, short-term as well long-term, and the denominator of this ratio is total of all assets.

(c) *Interest Coverage Ratio (also called "times interest earned"):* A high interest coverage ratio means that the firm can easily meet the interest burden even if profit before interest and taxes suffer a considerable decline. A low interest coverage ratio may result in financial embarrassment when profit before interest and taxes decline.

Though widely used, this ratio is not a very appropriate measure because the source of interest payment is cash flow before interest and taxes.

(d) *Fixed Charges Coverage Ratio:* This ratio shows how many times the cash flow before interest and taxes covers all fixed financing charges. In the denominator of this ratio, only the repayment of loan is adjusted upwards for the tax factor because the loan repayment amount, unlike interest, is not tax deductible.

(e) Debt Service Coverage Ratio

#### **Turnover Ratios**

Turnover ratios also referred to as activity ratios or assets management ratios, measure how efficiently the assets are employed by a firm. The important turnover ratios are:

(a) *Inventory Turnover:* It measures how fast the inventory is moving through the firm and generating sales. It reflects the efficiency of inventory management.

(b) **Debtors' Turnover:** It shows how many times accounts receivable (debtors) turnover during the year.

(c) *Average Collection Period:* It represents the number of days' worth of credit sales that is locked in debtors.

(d) *Fixed Assets Turnover:* This ratio measures sales per rupee of investment in fixed assets.

This ratio is supposed to measure the efficiency with which fixed assets are employed.

(e) Total Assets Turnover: This ratio measures how efficiently assets are employed overall.

#### **Profitability Ratios**

They reflect the final result of business operations. There are two types of profitability ratios:

(i) Profit margin ratios

(ii) Rate of return ratios

The important profit margin ratios are:

(a) *Gross Profit Margin Ratio:* The ratio shows the margin left after meeting manufacturingcosts. It measures the efficiency of the production as well as pricing.

(b) *Net Profit Margin Ratio:* This ratio shows the earnings left for shareholders as a percentageof net sales.

(c) *Return on Total Assets:* It is measure of how efficiently the capital is employed. To ensureinternal consistency, the following variant of return on total assets may be employed:

(a) *Earning Power:* It is a measure of operating profitability.

(b) **Return on Equity:** it is a measure of great interest to equity shareholder. The numerator of this ratio is equal to profit after tax less preference dividends. The denominator

includes all contributions made by equity shareholders. It is also called the return on net worth.

#### Valuation Ratios

Valuation ratios indicate how the equity stock of the company is assessed in the capital market:

(a) *Price-earnings Ratio:* The market price per share may be the price prevailing on a certain day or the average price over a period of time. The earnings per share are simply: profit after tax less preference divided by the number of outstanding equity shares.

(b) *Yield:* It is a measure of the rate of return earned by shareholders.

#### (c) Market Value to Book Value Ratio

(d) *"q" Ratio:* Proposed by James Tobin, this ratio resembles the market value to book value ratio. However, there are two key differences:

(i) The numerator of the 'q' ratio represents the market value of equity as well as debt, not justequity.

(ii) The denominator of the' q' ratio represents all assets. Further, these assets are reckoned attheir replacement cost, not book value.

#### CLASSIFICATION AND TABULATION

Classification is the process of sorting 'similar' things from among a group of objects with different characteristics. In other words, heterogeneous data is divided into separate homogeneous classes according to characteristics that exist amongst different individuals or quantities constituting the data. Thus, fundamentally classification is dependent upon similarities among the items in the data.

The main object of classification is to present vividly, in a simplified and quickly intelligible form, a mass of complex data. Without condensing details in a classified form it is difficult to compare quickly, interpret thoroughly and analyse properly different sets of quantitative and qualitative phenomena. The basic requirements of good classification are stability, nonambiguity, flexibility and comparability.

#### **Descriptive and Quantitative Classification**

Depending on the characteristics of the data, they can be broadly categorized into two separate and distinct groups - descriptive and numerical. Descriptive characteristics are those that can be described in words and are expressible in qualitative terms. Numerical characteristics are quantitative in nature. For instance, literacy, sex, caste and religion are descriptive characteristics. Height, weight, age, income and expenditure are numerically expressible characteristics. Descriptive or qualitative classification is termed classification according to attributes. Numerical or quantitative classification of data in certain class intervals is termed as classification in terms of classes with certain intervals, or classification according to class intervals.

#### Simple and Manifold Classification

Classification based on attributes may be either simple or manifold. In the case of simple classification, only one attribute is studied. That is, the data is classified into two separate classes under a single attribute. For instance, data collected on literacy in the country can be classified into two distinct classes: literate and illiterate. Since this process is quite simple, it is known as simple classification.

On the other hand, analysing and classifying collected data under several attributes in different classes is called manifold classification. For example, if each of the two classes, literate and illiterate, is divided into males and females, then there would be four classes. If classified further on a regional basis, there would be a number of other classes. Such a process of classification of data into a number of classes and classes within classes is known as manifold classification.

#### **Classification According to Class Intervals**

Phenomena like income, heights and weights are all quantitatively measurable and data on them can be classified into separate class intervals of uniform length. For instance, the marks obtained by a group of 50 candidates in a subject at an examination can be classified into the following classes: 0-10, 10-20, 20-30, 30-40, 40-50, 50-60, 60-70 etc. Each class has a lower and an upper limit and the number of candidates getting marks between these two limits of the same class interval is called the frequency of the respective class. To give an example, if 12 candidates get between 40 and 50 marks, 12 is the frequency of the class 40-50.

#### Number of Classes

The number of classes into which particular data should be classified depends upon the mass of data. The larger the mass, the more should be the number of classes. Usually data is classified into not less than six classes and not more than 20 classes, depending upon the mass and size of the data and the length of the class intervals. The fundamental object of classifying data is to get the maximum possible amount of information most precisely. According to Sturges' Rule, the number of class intervals (n) =  $1 + 3.322 \log N$ , where N = total number of observations.

#### Length of Class Intervals

The uniform length of class intervals depends upon the difference between the extreme items in the data-the largest item and the smallest item-and the number of classes required. For example, if in the data on marks secured by 250 candidates in a subject at an examination, 0 and 93 are the lowest and highest marks respectively and 10 classes are required, each class would then have a class interval length of 10. Ordinarily class intervals are fixed in such a way as to enable easy calculation and precision.

#### Class Limits

The choice of class limits is determined by the mid-value of a class interval, which should as far as possible be identical with the arithmetic average of the items occurring in that class interval.

#### Tabulation

Tabulation is the process of arranging given quantitative data based on similarities and common

characteristics in certain rows and columns so as to present the data vividly for quick intelligibility, easy comparability and visual appeal.

#### Components of a Statistical Table

A statistical table comprises a title, a head-note, a stub head and stub details, captions and columns under the captions, field of the table under different column heads, footnotes and, source notes.

*Here's a sample:* 

*Title:* Students studying in different classes in X, Y, Z Commerce College.

Head-Note: Data relates to the academic year for ex.1998-99.

#### Purpose of Statistical Tables

Statistical tables are of two types: general purpose table and special purpose table.

1. *General Purpose Table:* This is primarily meant to present the entire original data on a subject. Such presentation of numerical data in a tabular form is especially useful as a source of information and reference for constructing different special

purpose tables.

2. *Special Purpose Table:* As its name implies, this is a statistical table that specially presents and emphasizes certain phases or significant aspects of the information given in a general purpose table. Presenting data in a special table not only makes it easy to understand specific data, it also facilitates easy comparison and clear-cut interpretation.

#### 4.5.5 Types of Tabulation

1. *One-way Table (single tabulation):* A one-way table gives answers to questions about one characteristic of the data.

2. *Two-way Table (double tabulation):* A two-way table gives information about two interrelated characteristics of a particular type of data.

3. *Three-way Table (Triple Tabulation):* A three-way table answers questions relating to three interrelated characteristics of a given data.

4. *Higher Order Table (Manifold Tabulation):* This table gives information under several mainheads and subheads on questions relating to a number of interrelated characteristics.

#### **Rules and Precautions on Tabulation**

1. Every statistical table should be given an appropriate title to indicate the nature of the data.

The title should be simple, intelligible and unambiguous and should not be too lengthy or too short.

2. If necessary, the title may be further explained with a suitable head-note.

3. Different types of data require different types of tabulation. It has to be decided at the outset whether one or more tables would be necessary to fit in the data precisely and suitably. A single simple table is appealing to the eye provided it is prepared properly. Several tables or a large table make comparisons difficult.

4. The stub heads and the main heads should be consistent with the nature of the data and be veryclear.

5. The main headings under the caption should be as few as possible in keeping with the requirements of space and type of data. If the main headings are few, comparison between different sets of data becomes easy.

6. The entire arrangement of data should be appropriate, compact and self-explanatory so that it is not necessary to rearrange the data in any manner.

7. Comparisons between different figures such as totals and averages-are easier if they are arranged vertically and not horizontally.

8. In order to show important parts of the data (under main heads) distinctly, it is necessary todraw thick double or multiple ruled lines.

9. Depending upon the nature of the data, items in the stub column may be arranged accordingto:

(i) Alphabetical order.

- (ii) Geographical importance.
- (iii) Customary classification.
- (iv) Chronological order.
- (v) Size or magnitude.

10. Figures in the data that are estimates, approximate or revised should be indicated by an alphabet, asterisk, number or any other symbol. An explanation should be given in the footnote.

11. The different units used in the data should be indicated in the column heads. For example: 'figures in rupees', 'figures in metres', and so on.

12. The source of the data should be indicated under the footnote. It is necessary to mention the source for further references and other details and also for indicating the reliability of the data.

#### STATISTICAL DATA ANALYSIS

The data generated using the questionnaire is analysed and inference made out of the data could be used by the decision maker. The fundamental question that arises in the minds of the researcher is: "What technique should be used to analyse the collected data?"

The collected data may be coded as per the description given in the scaling lesson. The researcher should ensure that he/she does not deviate from the scaling principles enumerated in the scaling lesson. The researcher can create a master file containing the coded information of allthe items included in the questionnaire.

The choice of technique to analyse the collected data could be pictorially represented as given in figure 11.1. Data analysis technique depends on the level of measurement and the type of sample the researcher uses. An overview of the choice of techniques used is given in figure 11.1. Descriptive statistics such as mode and relative and absolute frequencies are used on nominal data. Further chi-square test and Mcnemer test is used as inferential statistics. Ordinal data may be subjected to median and interquartile range. Under inferential statistics, non parametric techniques such as Kolmogorov Smirnov test, Mann Whitney test, Kruskal Wallis, and Friedman two- way ANOVA are used. Interval and ratio scale may be subjected to mean and standard deviation. Under inferential statistics, z test, t –test, one-way ANOVA, correlation and regression.

#### **HYPOTHESIS TESTING**

Basic analysis of the data involves testing of hypothesis. Lot of confusion prevails in developing a hypothesis. In simple terms, hypothesis refers to assumption of a relationship between two variables or difference between two or more groups. Hypothesis also contains the direction of relationship between the variables concerned.

Examples for hypothesis is given below:

(a) The purchasing power of the consumers is positively related to the availability of surplusincome.

(b) Customers belonging to the Northern states in India have a different taste preference thanthose from Northern States.

Hypotheses are of two types: (a) Null hypothesis and (b) Alternative hypothesis. A simple rule may be followed to develop a hypothesis:

1. What we hope or expect to be able to conclude as a result of the test usually should be placedin alternative hypothesis.

2. The null hypothesis should contain a statement of equality (=) and an alternative hypothesis contains a > or < than sign.

- 3. The null is the hypothesis that is tested.
- 4. The null and alternate hypothesis are complementary.

#### HOW TO SELECT A PARTICULAR TEST

An appropriate statistical test for analysing a given set of data is selected on the basis of: Scaling of the data: Is the measurement scale nominal, ordinal, interval or ratio;

Dependence, Independence of the measurements; Types of samples: Independent or dependent samples; Number of samples (groups) studied and;

Specific requirements such as sample size, shape of population distribution, are also used for considering the choice of a statistical test.

There are two types of samples: Independent and dependent samples. Two samples are independent sample if the sample selected from one of the populations has no effect or bearing on the sample selected from the other population. E.g., responses collected from Tamilians, Keralites, Kannadigas etc. They are exclusive groups of respondents where a Tamilian is exclusive in nature in that he does not take part in the other groups. Similarly, a Kannadiga is exclusive in nature in his membership in his group in that he does not take part in any other groups.

Dependent samples, also called related or correlated or matched samples, are ones in which the response of the nth subject in one sample is partly a function of the response of the nth subject in an earlier sample. Examples of dependent samples include before-during-after samples of the same people or matched response of similar people.

The nature of the samples is also considered while deciding on the appropriateness of the statistical test. The following are the conditions to be followed while choosing the tests:

Does the test involve one sample, two samples or k samples

If 2 samples or k samples are involved, are the individual cases independent or related.

The selection of an appropriate statistical test rests with two criteria:

(a) Type of scale used (Nominal, ordinal, interval or ratio)

(b) Type and the size of the samples. Type relates to whether the samples are independent ordependent.

The hypothesis of type two mentioned in the example above could be tested using two types of statistical tests. They are:

- (a) Parametric tests
- (b) Non-parametric tests

A simple understanding of the characteristics of the tests reveal that the term "parametric" is derived from the term parameter which is a descriptive measure computed from or used to describe a population of data. Parametric tests are used to test hypothesis with interval and ratio measurements and non parametric tests are used to test hypothesis involving nominal and

ordinal data. Parametric tests are more powerful than non-parametric tests. Explanation of parametric and non parametric tests in detail is beyond the scope of this study material.

There are few simple, easy to understand assumptions made while applying a parametric test. They are:

The observations must be independent – that is, the selection of any one case should notaffect the chances for any other case to be included in the sample.

The observations should be drawn from normally distributed populations. These populations should have equal variances.

The measurement scales should be at least interval so that arithmetic operations can be used with them.

Non-parametric tests do not have any assumptions of such kind. This is the advantage of non-parametric tests over parametric tests.

Hypothesis of the type 1 may be tested using Correlation and regression. Correlation is a test of association only between two variables. It uses only interval and ratio scale. Such correlations are called as Karl Pearson bi–variate correlation. Correlation of a special type employed on ordinal data is called Rank Correlation. This is otherwise called as Spearman Rank correlation. However, correlation will never tell the researcher about the independent – dependent relationship. Correlation analysis will give a result r called the correlation coefficient. R value ranges from -1 to +1 through a O. As r value approaches 1, the strength of the association increases and as it approaches 0, it decreases. R value will be associated with a positive or negative sign. Positive sign refers to positive correlation where the change in one variable causes change in the other variable in the same direction whereas a negative sign indicates inverse relationship.

Regression is a powerful technique dealing with two or more than two number of variables. Regression analysis will tell the researcher about the independent and dependent relationship. It deals with one dependent variable and any number of independent variables. Regression analysis involving only one independent variable, is called simple regression and that involves more than one independent variables is called multiple regression. Regressionresults in  $r^2$  value which explains the amount of variance accounted for, by the independent variables on the dependent variable. Standardized  $\beta$  coefficient determines the strength and the direction of relationship between the independent and dependent variables.

#### **REPORT DESIGN AND WRITING IN BUSINESS RESEARCH**

#### **INTRODUCTION**

Much has been dealt in detail in the previous lesson about the processes involved in research. The researcher may be glued into the technicalities in doing a research, however, the research effort goes in vain, if it is reported in a systematic manner to concerned decision makers. The report should be presented in a way what the decision maker needs and wishes to know. The decision maker is interested only in the results rather than complicated tables and he/she should be convinced of the usefulness of the findings. He / she must have sufficient appreciation of the method to realize its strengths and weaknesses. Research report is the only one which communicates with the decision maker.

Research reports are the only tangible products of a research project and only

documentary evidence on which the decision maker can make decisions. Management decisions on the problem concerned are guided by the report and presentation. Moreover, the report should be meticulously presented as this would form part of a secondary data at a later stage. Any reference to this report should convey information in an unambiguous manner with clarity.

#### **CRITERIA FOR REPORT WRITING**

The research report should be made as per the requirement of the decision maker meaning that it should purely and simply tailor made for the decision maker with due regard for their technical sophistication, interest in the subject area, circumstances under which they will read the report, and use they will make of it. The report should be made keeping in mind the technical sophistication of the decision maker. A decision maker with little technical sophistication may sometimes distort the inference that could be made from the result. Sometimes use of sophisticated technical jargons may result in the decision maker looking at the researcher with suspicion that he / she has used his high flair knowledge to prove his supremacy in the area of research.

The researcher may be confronted with a situation where the report he or she makes is meant for several others in the organization. In such a case, preparing a report that would satisfy everyone in the organization would be a tough task. In this regard, the researcher should have an understanding of the technical capacity and level of interest in the report by everyone concerned.

It may be appropriate if the researcher discusses the major findings, conclusions and recommendations with the decision makers before sitting down to prepare. Discussions before submission may prevent major discord among the targets to whom the research report is concerned. This would also result in the researcher knowing the needs of the concerned

decision makers and ensures that the report meets the client"s needs and finally the report is ultimately accepted. The discussion on the results should confirm specific dates for the delivery of the written report and other data.

#### **REPORT FORMAT**

Research formats may vary from researcher to researcher as well depending on the need of the decision maker. However, any researcher could not violate the fundamental contents a report should have. They should include the following:

i) Title page includes the title of the report, name, address and telephone number of the researcher or organization conducting the research, the name of the client for whom the report was prepared and the date of release.

ii) Letter of transmittal refers to a summary of the researcher's overall experience with the research, without mentioning the findings.

iii) Letter of authorization contains the authorization given by the decision maker to the researcher to do the project.

iv) Table of contents include the list of topics covered and appropriate page number.

v) Executive summary is important in a research report as this presents the report in a shortened form. Sometimes, the decision maker would read only this portion of the report when constrained by time. This should describe the problem, approach, and research design that was adopted. A small portion of the summary section should be devoted to the major results, conclusions and recommendations.

vi) Problem definition shows the background to the problem, highlights the discussion with the decision makers and industry experts and discusses the secondary data analysis, the qualitative research that was conducted, and the factors that were considered.

vii) Approach to the problem discusses the broad approach that was adopted in addressing the problem. This should contain a description of the theoretical foundations that guided the research, any analytical models formulated, research questions, hypothesis and the factors that influenced the research design.

viii) Research design shows the details of the nature of the research design adopted, information needed, data collection from secondary and primary sources, scaling techniques, questionnaire development and pretesting, sampling techniques, and field work.

ix) Data analysis describes the plan of the data analysis conducted on the data. It justifies the choice of the technique for a particular objective and hypothesis.

x) Results comprise of the results presented not only at the aggregate level but also at the subgroup level. The results, as mentioned earlier, should be presented in the most simpler way, enabling the decision maker to understand in the right sense.

xi) Limitations and Caveats contain the limitations caused by the research design, cost, time and other organizational constraints. However, a research should not contain many limitations. The researcher should have controlled many of the limitations during the research process.

xii) Conclusions and recommendations involve interpretation of the results in light of the

problem being addressed to arrive at major conclusions. The decision maker makes decision based on the conclusion and recommendations of the researcher.

#### **GUIDELINES FOR TABLES**

Data analysed should be presented in the research report in a tabular form. The guidelines for tables are as follows:

i) Title and number should be given for every table such as 1a. The title should be very brief just explaining the description of the information provided in the table.

ii) Arrangement of data items indicate that the data should be arranged in some order either pertaining to time or data etc.

iii) Leaders, ruling and spaces should be made in such a way that they lead the eye horizontally, impart uniformity, and improve readability.

iv) Explanations and comments: explanations and comments clarifying the table may be provided in the form of captions, stubs and footnotes. Designations placed on the vertical columns are headings; those placed in the left-hand are called stubs. Information that cannotbe incorporated in the table should be explained by footnotes.

v) Sources of the data refer to citing the source of secondary data used in the research.

#### **GUIDELINES FOR GRAPHS**

The researcher may have used graphical interpretation of the results. Use of graphs complements the textand the table adding clarity of communication and impact. The researcher may use any type of graphssuch as pie or round charts, line charts, pictographs, histograms and bar charts. While presenting the graphs, the researcher should ensure that each section or line or bar of the charts should be represented in different colours or shades.