PREFACE

In the curricular structure introduced by this University for students of Post-Graduate degree programme, the opportunity to pursue Post-Graduate course in a subject is introduced by this University is equally available to all learners. Instead of being guided by any presumption about ability level, it would perhaps stand to reason if receptivity of a learner is judged in the course of the learning process. That would be entirely in keeping with the objectives of open education which does not believe in artificial differentiation. I am happy to note that university has been recently accredited by National Assessment and Accreditation Council of India (NAAC) with grade 'A'.

Keeping this in view, the study materials of the Post Graduate level in different subjects are being prepared on the basis of a well laid-out syllabus. The course structure combines the best elements in the approved syllabi of Central and State Universities in respective subjects. It has been so designed as to be upgradable with the addition of new information as well as results of fresh thinking and analysis.

The accepted methodology of distance education has been followed in the preparation of these study materials. Co-operation in every form of experienced scholars is indispensable for a work of this kind. We, therefore, owe an enormous debt of gratitude to everyone whose tireless efforts went into the writing, editing, and devising of a proper layout of the materials. Practically speaking, their role amounts to an involvement in 'layout of the materials. Practically speaking, their role amounts to an involvement in 'invisible teaching'. For, whoever makes use of these study materials would virtually derive the benefit of learning under their collective care without each being seen by the other.

The more a learner would seriously pursue these study materials, the easier it will be for him or her to reah out to larger horizons of a subject. Care has also been taken to make the language lucid and presentation attrative so that they may be rated as quality self-learning materials. If anything remains still obscure or difficult to follow, arrangements are there to come to terms with them through the counselling sessions regularly available at the network of study centres set up by the University.

Needless to add, a great deal of these efforts is still experimental—in fact, pioneering in certain areas. Naturally, there is evedry possibility of some lapse or deficiency here and there. However, these do admit of rectification and further improvement in due course. On the whole, therefore, these study materials are expected to evoke wider appreciation the more they receive serious attention of all concerned.

Professor (Dr.) Subha Sankar Sarkar
Vice-Chancellor

Netaji Subhas Open University

Under Graduate Degree Programme
Choice Based Credit System (CBCS)
Subject : Honours in Geography (HGR)
Paper : Skill Enhancement Course
Course Code : SE-GR-21
Course : Research Methods

Published — May, 2022

Printed in accordance with the regulations of the Distance Education
Bureau of the University Grants Commission.

Netaji Subhas Open University

Under Graduate Degree Programme

Choice Based Credit System (CBCS)

Subject : Honours in Geography (HGR)

Paper : Skill Enhancement Course

Course Code : SE-GR-21

Course : Research Methods

: Course Writer :

Dr. Pushpita Sengupta
Assistant Professor of Geography
Purash-Kanpur Haridas Nandi Mahavidyalaya

: Course Editor :

Smt. Tinki Kar Bhattacharya
Assistant Professor of Geography
Netaji Subhas Open University

: Format Editor :
Smt. Tinki Kar Bhattacharya
Assistant Professor of Geography
Netaji Subhas Open University

Netaji Subhas
Open University

UG : GEOGRAPHY
[HGR]

Course : Research Methods
Course Code : SE-GR-21

Research Methods

Unit 1: ❒ Geographic Enquire: Definition and Ethics, Literature Review, Framing Research Questions, Objective and Hypothesis, Preparing Sample Questionnaires and Inventories

STRUCTURE

1.1 ❒ Introduction

Research is a derivative of the French word "Researche" means quest, search, pursuit and search for truth. Therefore, research in common parlance refers to a search for knowledge. Once can also define research as a scientific and systematic search for pertinent information on a specific topic. In fact, research is an art of scientific investigation. Research is an academic activity and as such the term should be used in a technical sense. According to Clifford Woody research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organizing and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis. D. 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 of an art." Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the pursuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research. The systematic approach concerning generalization and the formulation of a theory is also research. As such the term 'research' refers to the systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analyzing the facts and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalizations for some theoretical formulation.

1.2 ❒ Objectives

The main objectives of this chapter are

1. Understand the meaning of research
2. Cognize and distinguish between different kinds of researches.
3. Discuss the process of framing literature review.
4. Learn to prepare the research questions.
5. Identify the objectives and the hypothesis of the research.
6. Learn the preparation of sample questionnaires and inventories to conduct the research.

1.3 ❒ Geographic Enquire: Definition and Ethics

Geographical enquiry deals with an approach of learning that accepts that knowledge has been constructed and prioritizes the need for students to make sense of things for themselves. Although geographical enquiry is often associated with sequences of skills, these skills need to be developed and refined in specific.

Roberts (2003) identified four important aspects of enquiry, these are as follows:

- **Creating a need to know** - In a traditional transmission approach to teaching, the teacher decides what the students are going to learn and then what the learning outcomes will be by presenting lesson objectives. This assumes that it is possible for all students to achieve the same objectives in spite of the different knowledge, experiences and skills they bring to the lesson. In contrast, an enquiry approach puts an emphasis on questions and encourages curiosity. At the start of a lesson, instead of identifying the end-points of learning, an enquiry approach identifies the starting point: a question to frame what is being studied. Riley (2000) characterizes a good enquiry question as one that captures interest and imagination, that places a historical concept or process at the forefront of students' minds, and that results in a tangible, lively, enjoyable activity. In an enquiry approach to learning, it is important that the big questions become the students' own. This involves provoking curiosity, possibly by setting up some puzzling situation or problem, or through a stimulus (Davidson, 2006).
- **Using data** - In an enquiry approach, students are expected to analyse, interpret and challenge data. Students are made aware that data has been selected, either because they deal with selection problems themselves in the field, library or at the computer, or through some activity, e.g. dealing with conflicting data on the same topic. Also, in this approach students learn to handle data presented in a relatively unprocessed form, so that it is the students rather than the teacher or textbook author who categorise data into advantages/disadvantages etc.
- **Making sense** - In an enquiry approach, opportunities are provided for students to make sense of data for themselves. First, they need time and space to relate new information and ideas to what they already know; they need time to think, rather than being hurried along by a need for 'pace'. Second, they need to be introduced to the big ideas of geography that will help them to make sense of new data – ideas such as sustainability, interdependence, globalisation. The role of classroom talk is essential in helping students make sense. This might include

whole-class discussion or role-play. It might involve students working in small groups, where activities might include categorising, ranking or evaluating pieces of information, or making links on spider diagrams and concept maps, or developing arguments for or against something. There is a need for new frameworks to help students with data, partly to extend the range of questions they consider to include the moral aspects of geography, 'e.g. What ought to happen?' and 'Who should be responsible?', and partly to help students develop different ways of thinking about sustainability, globalisation etc. Making sense might involve extended writing, preceded by class discussion and some draft writing. The key point about all these possible activities is that they require students to study the data and make links for themselves, supported in developing their understanding through conversations with the teacher and each other.

- **Reflecting on learning** - In an enquiry approach to teaching, the teacher returns to the questions that framed the enquiry and explores with a class the extent to which the questions have been answered. When debriefing, the emphasis for the teacher is on trying to understand what has gone on in the students' minds, what sense they have made of what they have studied, and whether they have had to rethink what they knew before. This approach is more likely to 'lead the student to unanticipated rather than predicted outcomes' and this, according to McKernan (quoted in 2009 in the GA's manifesto for Geography) would make the experience 'truly educational'.

1.4 ❒ Literature Review

A literature review is a survey of scholarly sources that provides an overview of a particular topic. Literature reviews are a collection of the most relevant and significant publications regarding that topic in order to provide a comprehensive look at what has been said on the topic and by whom. The basic components of a literature review include:

- a description of the publication;
- a summary of the publication's main points;
- a discussion of gaps in research;
- an evaluation of the publication's contribution to the topic.

The purpose of a literature review is to provide a review of writings on the given topic in order to establish the reviewer's own position in the existing field of scholarship on that

topic. A literature review provides a reader with a comprehensive look at previous discussions prior to the one the reviewer will be making in his/her own research paper, thesis, or dissertation.

A literature review has four main objectives:

- It surveys the literature in chosen area of study.
- It synthesises the information in that literature into a summary.
- It critically analyses the information gathered by identifying gaps in current knowledge; by showing limitations of theories and points of view; and by formulating areas for further research and reviewing areas of controversy.
- It presents the literature in an organised way.

Therefore, a literature review can be just a simple summary of the sources, but it usually has an organizational pattern and combines both summary and synthesis. A summary is a recap of the important information of the source, but a synthesis is a re-organization, or a reshuffling, of that information. It might give a new interpretation of old material or combine new with old interpretations or they might trace the intellectual progression of the field, including major debates. And depending on the situation, the literature review may evaluate the sources.

The different way of organizing the sources into a review are as follows:

- **Chronological:** If a review follows the chronological method, one could write about the materials on the basis of their publication date.
- **By trend:** A better way to organize the materials sources chronologically is to examine the sources under another trend, such as the history of the research item. Then the review report would have subsections according to eras within this period.
- **Thematic:** Thematic reviews of literature are organized around a topic or issue, rather than the progression of time. However, progression of time may still be an important factor in a thematic review. However, more authentic thematic reviews tend to break away from chronological order. The subsections might include how they are personified, how their proportions are exaggerated. A review organized in this manner would shift between time periods within each section according to the point made.

- **Methodological:** A methodological approach differs from the two above in that the focusing factor usually does not have to do with the content of the material. Instead, it focuses on the "methods" of the researcher or writer. A methodological scope will influence either the types of documents in the review or the way in which these documents are discussed.

Once you've decided on the organizational method for the body of the review, the sections you need to include in the paper should be easy to figure out. They should arise out of your organizational strategy. In other words, a chronological review would have subsections for each vital time period. A thematic review would have subtopics based upon factors that relate to the theme or issue.

Sometimes, though, you might need to add additional sections that are necessary for your study, but do not fit in the organizational strategy of the body. What other sections you include in the body is up to you. Put in only what is necessary. Here are a few other sections you might want to consider:

- **Current Situation:** Information necessary to understand the topic or focus on the literature review.
- **History:** The chronological progression of the field, the literature, or an idea that is necessary to understand the literature review, if the body of the literature review is not already a chronology.
- **Methods and/or Standards:** The criteria you used to select the sources in your literature review or the way in which you present your information. For instance, you might explain that your review includes only peer-reviewed articles and journals.

A literature review is a piece of **discursive prose**, not a list describing or summarizing one piece of literature after another. It's usually a bad sign to see every paragraph beginning with the name of a researcher. Instead, organize the literature review into sections that present themes or identify trends, including relevant theory. You are not trying to list all the material published, but to synthesize and evaluate it according to the guiding concept of your thesis or research question

1.5 ❒ Framing Research Questions

Framing a clear research question is a crucial part of developing your research proposal, and should be seen as emerging from a dialogue between a developing theoretical position

and decisions you need to take about research design and subsequent data analysis. Therefore, research questions play a central role in all of the research efforts that follow. Well-formulated research questions will not assure useful research results; however, poorly formulated research questions will almost certainly guarantee that the research results will at best be of limited value to the decision maker.

THEORETICAL AND EMPIRICAL ANTECEDENTS

↓

RESEARCH QUESTION

↓

RESEARCH DESIGN

(What, when, where, how, with/from whom)

One way of structuring research questions effectively utilizes a hierarchical approach. In this approach, researchers begin with very general questions. Then from the general questions researchers pose questions of increasing specificity, which are implied by a general question. When questions reach a level of specificity acceptable to the researcher, they become the basis for writing questionnaire items.

An excellent way to form general research questions begins with the information gaps the decision maker and researcher identified earlier when discussing the decision problem. Recall that information gaps are simply the difference between what a decision maker currently knows and would like to know about a problem in order to make a decision.

The information gaps represent an easy way to begin the process of writing the research questions that will ultimately produce the data that addresses the decision problem. And because the general research questions derive from the decision problem and information gaps, they are much more likely to be of use to the decision maker.

The preceding discussion covers only one way of formulating research questions. Researchers should use whatever methods suit them. The important point to all this is that good research questions are critical to the overall research effort. Therefore, formulating them should represent a substantial portion of the overall research project.

As noted earlier, it may be helpful to construct a hierarchy of research questions. That is, each general research question should imply a series of more specific research questions. For convenience, these more specific questions will be referred to as “implied research questions.” The structure of general questions producing implied questions suggests a

structure illustrated hypothetically in Chart 1 below. It shows only a hypothetical possibility and a small one at that. The exhibit shows only one general question when in reality several are likely. The exhibit shows only a handful of implied questions flowing from the general question, when a dozen or more are not at all uncommon. The main points to the exhibit are that general questions produce implied questions, and that implied questions may produce one or more additional implied questions – each level of increasing specificity.

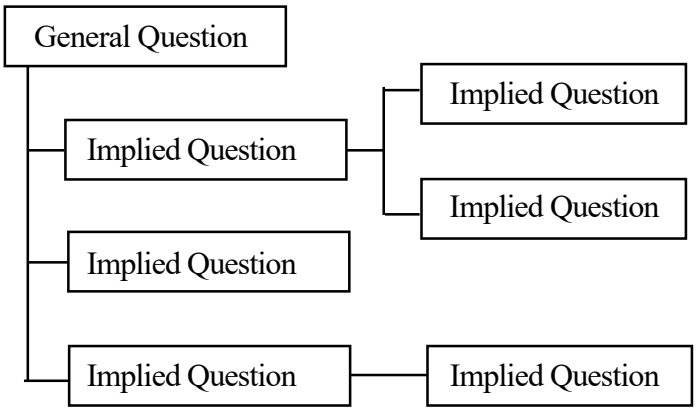


Chart 1. Hierarchy of Research Questions

Researchers should brainstorm in order to produce as comprehensive a list of questions as possible. Ultimately, the implied questions will become the basis for producing and screening questionnaire items. Therefore, produce a better list of questions, and you will produce a better questionnaire and in turn collect better data.

Many students wonder how specific they should be with their implied research questions. No simple rule exists to answer this question. Often, knowing when to stop picking apart a general question becomes a matter of experience and intuition. In general, it's better to be too detailed than not detailed enough.

Even experienced researchers find the process of developing useful general and implied research questions difficult. It requires creativity, patience, and a willingness to think deeply about a given subject. Difficult though it may be, the results are crucial for a successful research project. Therefore, it's certainly well worth the time and effort it takes.

Because research questions are so critical, it's essential that they be communicated in writing to everyone working on the research project. Research teams should not assume they know what questions the project addresses. Being explicit keeps all parties working together toward the same goals.

1.6 ❒ Objectives and Hypothesis

1.6.1 Objectives

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

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

Research objectives are the aims and goals of a session of study. It is what you hope to achieve from the study that you will do. When you clarify your objectives- preferably by writing them down - you sharpen your focus when you study.

A statement of research objectives can serve to guide the activities of research. Consider the following examples.

- **Objective:** To describe what factors farmers take into account in making such decisions as whether to adopt a new technology or what crops to grow.
- **Objective:** To develop a budget for reducing pollution by a particular enterprise.
- **Objective:** To describe the habitat of the giant panda in China.

In the above examples the intent of the research is largely descriptive.

- In the case of the first example, the research will end the study by being able to specify factors which emerged in household decisions.

- In the second, the result will be the specification of a pollution reduction budget.
- In the third, creating a picture of the habitat of the giant panda in China.

1.6.2 Hypothesis

A hypothesis is a tentative theory. The scientific process, for example, takes a hypothesis-which can range from a wild guess (unlikely) to a theory founded on solid evidence. Then by carefully designed experiments the process aims to uncover any flaws in the theory.

Research questions are formulated in conjunction with research hypotheses (also referred to as "conceptual hypotheses). Importantly, researchers should formulate research questions and hypotheses at the same time. Indeed, the process presented in this class essentially requires that these two steps occur simultaneously.

Research hypotheses serve a variety of purposes, however, most important at this stage of the research process is that research hypotheses help to "justify" whether a research question is sufficiently important to collect the data needed to answer it. If not, the question should be dropped. Another useful role of hypotheses is to help generate implied research questions. So although this topic is not covered just yet, bear in mind that research questions and research hypotheses work hand in hand and no discussion of research questions can be complete with incorporating research hypotheses into it.

Therefore, hypothesis is usually considered as the principal instrument in research. Its main function is to suggest new experiments and observations. In fact, many experiments are carried out with the deliberate object of testing hypotheses. Decision-makers often face situations wherein they are interested in testing hypotheses on the basis of available information and then take decisions on the basis of such testing. In social science, where direct knowledge of population parameter(s) is rare, hypothesis testing is the often used strategy for deciding whether a sample data offer such support for a hypothesis that generalisation can be made. Thus hypothesis testing enables us to make probability statements about population parameter(s). The hypothesis may not be proved absolutely, but in practice it is accepted if it has withstood a critical testing.

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.

1.6.3 Testing of Hypotheses

Basic concepts in the context of testing of hypotheses need to be explained.

(i) ***Null hypothesis and alternative hypothesis:*** In the context of statistical analysis, we often talk about null hypothesis and alternative hypothesis. If we are to compare method *A* with method *B* about its superiority and if we proceed on the assumption that both methods are equally good, then this assumption is termed as the null hypothesis. As against this, we may think that the method *A* is superior or the method *B* is inferior, we are then stating what is termed as alternative hypothesis. The null hypothesis is generally symbolized as H_0 and the alternative hypothesis as H_a. Suppose we want to test the hypothesis that the population mean is equal to the hypothesis mean (μ) is equal to the hypothesised mean (μH_0) = 100

Then we would say that the null hypothesis is that the population mean is equal to the hypothesized mean 100 and symbolically we can express as:

$H_0 : \mu = \mu_{H0} = 100$

If our sample results do not support this null hypothesis, we should conclude that something else is true. What we conclude rejecting the null hypothesis is known as alternative hypothesis. In other words, the set of alternatives to the null hypothesis is referred to as the alternative hypothesis. If we accept H_0, then we are rejecting H_a and if we reject H_0, then we are accepting H_a.

For H_0: $\mu = \mu_{H0} = 100$, we may consider three possible alternative hypotheses as follows:

Table 1

Alternative hypothesis	**To be read as follows**
$H_{a\,H} : \mu \neq \mu$	The alternative hypothesis is that the population mean is not equal to 100 i.e., it may be more or less than 100
$H_{a\,H} : \mu > \mu$	The alternative hypothesis is that the population mean is greater than 100
$H_{a\,H} : \mu < \mu$	The alternative hypothesis is that the population mean is less than 100

The null hypothesis and the alternative hypothesis are chosen before the sample is drawn (the researcher must avoid the error of deriving hypotheses from the data that he collects and then testing the hypotheses from the same data).

In the choice of null hypothesis, the following considerations are usually kept in view:

(a) Alternative hypothesis is usually the one which one wishes to prove and the null hypothesis is the one which one wishes to disprove. Thus, a null hypothesis represents the hypothesis we are trying to reject, and alternative hypothesis represents all other possibilities.

(b) If the rejection of a certain hypothesis when it is actually true, it is taken as null hypothesis because then the probability of rejecting it is true is α (the level of significance) which is chosen very small.

(c) Null hypothesis should always be specific hypothesis i.e., it should not state about or approximately a certain value. Generally, in hypothesis testing we proceed on the basis of null hypothesis, keeping the alternative hypothesis in view. Why so? The answer is that on the assumption that null hypothesis is true, one can assign the probabilities to different possible sample results, but this cannot be done if we proceed with the alternative hypothesis. Hence the use of null hypothesis (at times also known as statistical hypothesis) is quite frequent.

(ii) ***The level of significance:*** This is a very important concept in the context of hypothesis testing. It is always some percentage (usually 5%) which should be chosen with great care, thought and reason. In case we take the significance level at 5 per cent, then this implies that H_0 will be rejected when the sampling result (i.e., observed evidence) has a less than 0.05 probability of occurring if H_0 is true. In other words, the 5 per cent level of significance means that researcher is willing to take as much as a 5 per cent risk of rejecting the null hypothesis when it (H_0) happens to be true. Thus the significance level is the maximum value of the probability of rejecting H_0 when it is true and is usually determined in advance before testing the hypothesis.

(iii) ***Decision rule or test of hypothesis:*** Given a hypothesis H_0 and an alternative hypothesis H_a, we make a rule which is known as decision rule according to which we accept H_0 (i.e., reject H_a) or reject H_0 (i.e., accept H_a). For instance, if (H_0 is that a certain lot is good (there are very few defective items in it) against H_a) that the lot is not good (there are too many defective items in it), then we must decide the number of items to be tested and the criterion for accepting or rejecting the hypothesis. We might test 10 items in the lot and plan our decision saying that if there are none or only 1 defective item among the 10, we will accept H_0 otherwise we will reject H_0 (or accept H_a). This sort of basis is known as decision rule.

(iv) ***Type I and Type II errors:*** In the context of testing of hypotheses, there are basically two types of errors we can make. We may reject H_0 when H_0 is true and we may accept H_0 when in fact H_0 is not true. The former is known as Type I error and the latter as Type II error. In other words, Type I error means rejection of hypothesis which should have been accepted and Type II error means accepting the hypothesis which should have been rejected. Type I error is denoted by α (alpha) known as α error, also called the level of significance of test; and Type II error is denoted by β (beta) known as β error. In a tabular form the said two errors can be presented as follows:

Table 2

	Decision	
	Accept H_0	**Reject H_0**
H_0 (true)	Correct decision	Type I error (α error)
H_0 (false)	Type II error (β error)	Correct decision

The probability of Type I error is usually determined in advance and is understood as the level of significance of testing the hypothesis. If type I error is fixed at 5 per cent, it means that there are about 5 chances in 100 that we will reject H_0 when H_0 is true. We can control Type I error just by fixing it at a lower level. For instance, if we fix it at 1 per cent, we will say that the maximum probability of committing Type I error would only be 0.01.

But with a fixed sample size, *n*, when we try to reduce Type I error, the probability of committing Type II error increases. Both types of errors cannot be reduced simultaneously. There is a trade-off between two types of errors which means that the probability of making one type of error can only be reduced if we are willing to increase the probability of making the other type of error. To deal with this trade-off in business situations, decision-makers decide the appropriate level of Type I error by examining the costs or penalties attached to both types of errors. If Type I error involves the time and trouble of reworking a batch of chemicals that should have been accepted, whereas Type II error means taking a chance that an entire group of users of this chemical compound will be poisoned, then in such a situation one should prefer a Type I error to a Type II error. As a result one must set very high level for Type I error in one's testing technique of a given hypothesis. Hence, in the testing of hypothesis, one must make all possible effort to strike an adequate balance between Type I and Type II errors.

(v) ***Two-tailed and One-tailed tests:*** In the context of hypothesis testing, these two terms are quite important and must be clearly understood. A two-tailed test rejects the null hypothesis if, say, the sample mean is significantly higher or lower than the hypothesised value of the mean of the population. Such a test is appropriate when the null hypothesis is some specified value and the alternative hypothesis is a value not equal to the specified value of the null hypothesis. Symbolically, the two tailed test is appropriate when we have H_0: $\mu = \mu_{H0}$ and H_a:$\mu \neq \mu_{H0}$ which may mean $\mu > \mu H_0$ or $\mu < \mu H_0$. Thus, in a two-tailed test, there are two rejection regions, one on each tail of the curve.

Mathematically we can state:

Acceptance Region $A : Z \leq 1.96$

Rejection Region $R : Z \geq 1.96$

If the significance level is 5 per cent and the two-tailed test is to be applied, the probability of the rejection area will be 0.05 (equally splitted on both tails of the curve as 0.025) and that of the acceptance region will be 0.95. If we take m = 100 and if our sample mean deviates significantly from 100 in either direction, then we shall reject the null hypothesis; but if the sample mean does not deviate significantly from m , in that case we shall accept the null hypothesis.

But there are situations when only one-tailed test is considered appropriate. A *one-tailed test* would be used when we are to test, say, whether the population mean is either lower than or higher than some hypothesised value. For instance, if our $H_0 : \mu = \mu_{H0}$ and $Ha : \mu < \mu_{H0}$, then we are interested in what is known as left-tailed test (wherein there is one rejection region only on the left tail).

Mathematically we can state:

Acceptance Region A: $Z > -1.645$

Rejection Region R: $Z < -1.645$

If our $\mu = 100$ and if our sample mean deviates significantly from100 in the lower direction, we shall reject H_0, otherwise we shall accept H_0 at a certain level of significance. If the significance level in the given case is kept at 5%, then the rejection region will be equal to 0.05 of area in the left tail as has been shown in the above curve.

Mathematically we can state:

Acceptance Region A : $Z \leq 1.645$

Rejection Region A : $Z \geq 1.645$

If our $\mu = 100$ and if our sample mean deviates significantly from 100 in the upward direction, we shall reject H_0, otherwise we shall accept the same. If in the given case the significance level is kept at 5%, then the rejection region will be equal to 0.05 of area in the right-tail.

It should always be remembered that accepting H_0 on the basis of sample information does not constitute the proof that H_0 is true. We only mean that there is no statistical

evidence to reject it, but we are certainly not saying that H_0 is true (although we behave as if H_0 is true).

1.7 ❒ Preparing Sample Questionnaires and Inventories

1.7.1 Questionnaires

Well-designed questionnaire is essential for a successful survey. Unfortunately, questionnaire design has no theoretical base to guide the marketing researcher in developing a flawless questionnaire. All the researcher has to guide him/her is a lengthy list of do's and don'ts born out of the experience of other researchers past and present. Hence, questionnaire design is more of an art than a science.

The design of a questionnaire will depend on whether the researcher wishes to collect exploratory information (i.e. qualitative information for the purposes of better understanding or the generation of hypotheses on a subject) or quantitative information (to test specific hypotheses that have previously been generated).

Exploratory questionnaires: If the data to be collected is qualitative or is not to be statistically evaluated, it may be that no formal questionnaire is needed. For example, in interviewing the female head of the household to find out how decisions are made within the family when purchasing breakfast foodstuffs, a formal questionnaire may restrict the discussion and prevent a full exploration of the woman's views and processes. Instead one might prepare a brief guide, listing perhaps ten major open-ended questions, with appropriate probes/prompts listed under each.

Formal standardised questionnaires: If the researcher is looking to test and quantify hypotheses and the data is to be analysed statistically, a formal standardised questionnaire is designed. Such questionnaires are generally characterised by:

Prescribed wording and order of questions, to ensure that each respondent receives the same stimuli prescribed definitions or explanations for each question, to ensure interviewers handle questions consistently and can answer respondents' requests for clarification if they occur.

Prescribed response format, to enable rapid completion of the questionnaire during the interviewing process.

Given the same task and the same hypotheses, six different people will probably come up with six different questionnaires that differ widely in their choice of questions, line of

questioning, use of open-ended questions and length. There are no hard-and-fast rules about how to design a questionnaire, but there are a number of points that can be borne in mind:

1. A well-designed questionnaire should meet the research objectives. This may seem obvious, but many research surveys omit important aspects due to inadequate preparatory work, and do not adequately probe particular issues due to poor understanding. To a certain degree some of this is inevitable. Every survey is bound to leave some questions unanswered and provide a need for further research but the objective of good questionnaire design is to 'minimise' these problems.
2. It should obtain the most complete and accurate information possible. The questionnaire designer needs to ensure that respondents fully understand the questions and are not likely to refuse to answer, lie to the interviewer or try to conceal their attitudes. A good questionnaire is organised and worded to encourage respondents to provide accurate, unbiased and complete information.
3. A well-designed questionnaire should make it easy for respondents to give the necessary information and for the interviewer to record the answer, and it should be arranged so that sound analysis and interpretation are possible.
4. It would keep the interview brief and to the point and be so arranged that the respondent(s) remain interested throughout the interview.

It emphasises that writing of the questionnaire proper should not begin before an exploratory research phase has been completed. Even after the exploratory phase, two key steps remain to be completed before the task of designing the questionnaire should commence. The first of these is to articulate the questions that research is intended to address. The second step is to determine the hypotheses around which the questionnaire is to be designed.

It is possible for the piloting exercise to be used to make necessary adjustments to administrative aspects of the study. This would include, for example, an assessment of the length of time an interview actually takes, in comparison to the planned length of the interview; or, in the same way, the time needed to complete questionnaires. Moreover, checks can be made on the appropriateness of the timing of the study in relation to contemporary events such as avoiding farm visits during busy harvesting periods.

1.7.2 Preliminary decisions in questionnaire design

There are nine steps involved in the development of a questionnaire:

1. Decide the information required.
2. Define the target respondents.
3. Choose the method(s) of reaching your target respondents.
4. Decide on question.
5. Develop the question wording.
6. Put questions into a meaningful order and format.
7. Check the length of the questionnaire.
8. Pre-test the questionnaire.
9. Develop the final survey form.

1.7.3 Deciding on the information required

It should be noted that one does not start by writing questions. The first step is to decide 'what are the things one needs to know from the respondent in order to meet the survey's objectives?' One may already have an idea about the kind of information to be collected, but additional help can be obtained from secondary data, previous rapid rural appraisals and exploratory research. In respect of secondary data, the researcher should be aware of what work has been done on the same or similar problems in the past, what factors have not yet been examined, and how the present survey questionnaire can build on what has already been discovered. Further, a small number of preliminary informal interviews with target respondents will give a glimpse of reality that may help clarify ideas about what information is required.

At the outset, the researcher must define the population about which he/she wishes to generalise from the sample data to be collected. For example, in marketing research, researchers often have to decide whether they should cover only existing users of the generic product type or whether to also include non-users. Secondly, researchers have to draw up a sampling frame. Thirdly, in designing the questionnaire we must take into account factors such as the age, education, etc. of the target respondents.

It may seem strange to be suggesting that the method of reaching the intended respondents should constitute part of the questionnaire design process. However, a moment's

reflection is sufficient to conclude that the method of contact will influence not only the questions the researcher is able to ask but the phrasing of those questions. The main methods available in survey research are:

- personal interviews
- group or focus interviews
- mailed questionnaires
- telephone interviews.

Within this region the first two mentioned are used much more extensively than the second pair. However, each has its advantages and disadvantages. A general rule is that the more sensitive or personal the information, the more personal the form of data collection should be.

1.7.4 Putting questions into a meaningful order and format

Opening questions: Opening questions should be easy to answer and not in any way threatening to the respondents. The first question is crucial because it is the respondent's first exposure to the interview and sets the tone for the nature of the task to be performed. If they find the first question difficult to understand, or beyond their knowledge and experience, or embarrassing in some way, they are likely to break off immediately. If, on the other hand, they find the opening question easy and pleasant to answer, they are encouraged to continue.

Question flow: Questions should flow in some kind of psychological order, so that one leads easily and naturally to the next. Questions on one subject, or one particular aspect of a subject, should be grouped together. Respondents may feel it disconcerting to keep shifting from one topic to another, or to be asked to return to some subject they thought they gave their opinions about earlier.

Question variety: Respondents become bored quickly and restless when asked similar questions for half an hour or so. It usually improves response, therefore, to vary the respondent's task from time to time. An open-ended question here and there (even if it is not analysed) may provide much-needed relief from a long series of questions in which respondents have been forced to limit their replies to pre-coded categories. Questions involving showing cards/pictures to respondents can help vary the pace and increase interest.

Closing questions: It is natural for a respondent to become increasingly indifferent to the questionnaire as it nears the end. Because of impatience or fatigue, he may give careless answers to the later questions. Those questions, therefore, that are of special importance should, if possible, be included in the earlier part of the questionnaire. Potentially sensitive questions should be left to the end, to avoid respondents cutting off the interview before important information is collected.

In developing the questionnaire the researcher should pay particular attention to the presentation and layout of the interview form itself. The interviewer's task needs to be made as straight-forward as possible.

- Questions should be clearly worded and response options clearly identified.
- Prescribed definitions and explanations should be provided. This ensures that the questions are handled consistently by all interviewers and that during the interview process the interviewer can answer/clarify respondents' queries.

Ample writing space should be allowed to record open-ended answers, and to cater for differences in handwriting between interviewers.

In general it is best for a questionnaire to be as short as possible. A long questionnaire leads to a long interview and this is open to the dangers of boredom on the part of the respondent (and poorly considered, hurried answers), interruptions by third parties and greater costs in terms of interviewing time and resources. In a rural situation an interview should not last longer than 30-45 minutes.

1.7.5 Piloting the questionnaires

Even after the researcher has proceeded along the lines suggested, the draft questionnaire is a product evolved by one or two minds only. Until it has actually been used in interviews and with respondents, it is impossible to say whether it is going to achieve the desired results. For this reason it is necessary to pre-test the questionnaire before it is used in a full-scale survey, to identify any mistakes that need correcting.

The purpose of pretesting the questionnaire is to determine:

- Whether the questions as they are worded will achieve the desired results
- Whether the questions have been placed in the best order
- Whether the questions are understood by all classes of respondent

- Whether additional or specifying questions are needed or whether some questions should be eliminated
- Whether the instructions to interviewers are adequate.

Usually a small number of respondents are selected for the pre-test. The respondents selected for the pilot survey should be broadly representative of the type of respondent to be interviewed in the main survey.

If the questionnaire has been subjected to a thorough pilot test, the final form of the questions and questionnaire will have evolved into its final form. All that remains to be done is the mechanical process of laying out and setting up the questionnaire in its final form. This will involve grouping and sequencing questions into an appropriate order, numbering questions, and inserting interviewer instructions.

1.7.6 Inventory

An inventory of a study area involves an accounting of its land surface features. Such an objective is quantitative in nature and may be achieved by means of statistical sampling. Remote sensing based inventories are commonly conducted for natural resources such as forest timber or agriculture crops (Lillesand and Kiefer 2007; Ustin 2004). Student researchers may also wish to inventory built or other cultural features such as buildings or roads with high spatial resolution imagery. Even human resources (i.e., population) can be estimated through image analysis (Ridd and Hipple 2005). Since inventories involve enumerative estimates or listing of objects or resources, which can never be exact when conducted over large areas, a complete or "wall - to - wall" sampling may not be required. Some type of spatial sampling scheme can be implemented, such as sporadic image coverage of the study area and the use of transects or dot grids overlaid on each image. Examples of inventories that might be conducted by a student researcher with the aid of remotely sensed imagery are: (1) counts of automobiles in parking lots for transportation studies; (2) quantify cation of total stream length in a watershed; and (3) estimation of areal extent of land used for grazing in a state or country. In addition, the United State Environmental Protection Agency (EPA) has a variety of national, spatially contiguous environmental monitoring initiatives that use remotely sensed data both to inventory different properties across large regions and evaluate the status of large regions by sampling a subset of the total area.

1.8 ❒ Conclusion

To conduct a good, logical, systematic and scientific research the following steps should be followed:

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 are as 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.

1.9 ❒ Summary

(i) Research is a tool which is a building block and a sustaining pillar of every discipline. (ii) Research is an unbiased, structured and sequential method of enquiry, directed towards a clear implicit or explicit objective. This enquiry might lead to validating the existing postulates or arriving at new theories and models. (iii) Ideally research can be divided into various types like descriptive and analytical or applied and fundamental or qualitative and quantitative or conceptual and empirical. This differentiation is based on the nature and applicability of the research. (iv) Research, although not directly but is very important in the decision making in all the areas of life.

1.10 ❒ Glossary/Keywords

Research is simply the process of finding solution to a problem after thorough examination and analysis of factors.

Enquiry is an act of asking for information.

Hypothesis a supposition or proposed explanation made on the basis of limited evidence as a starting point for further investigation.

Questionnaire a set of printed or written questions with a choice of answers, devised for the purposes of a survey or statistical study.

1.11 ❒ References and Further Readings

Chandan, J.S. (2009) *Statistics for Business and Economics*, New Delhi: Vikas Publishing House Pvt. Ltd.

Chawla D. and Sondhi N. (2016) *Research Methodology: Concept and Cases,* Vikas Publishing House: New Delhi.

Davidson, G. (2006) 'Start at the beginning', *Teaching Geography*, 31 (3).

Gaur, A.S. and Gaur, S.S. (2009) *Statistical Methods for Practice and Research*, New Delhi: Response Books.

Geographical Association (2009) *A different view: a manifesto for the Geographical Association*. Sheffield: Geographical Association.

Kothari, C.R., (2004). '*Research Methodology: Methods and Techniques'*, New Age Publication

Krishnaswami, O.R. and Ranganatham, M. (2016) *Methodology of Research in Social Sciences*, 2/e, Mumbai: Himalaya Publishing House.

Punch, Keith, F., (1998) *Survey Research – the Basics*, New Delhi: Sage Publications.

Riley, (2000) 'Into the KS3 history garden: choosing and planting your enquiry questions', *Teaching History*, 99.

Roberts, M. (2010) Geographical enquiry http://www.open.edu/openlearn accessed on 22.02.2020

Sensisng, D.A., (2010) Remote Sensing '*Research Methods in Geography: A Critical Introduction'*. Edited by Basil Gomez and John Paul Jones III, *Wiley & Blackwell*, pp167

Walliman, N. (2016) *Social Research Methods,* New Delhi: Sage Publications.

1.12 ❒ Model Questions

1. Define research and examine its characteristics. Explain its significance in modern times.
2. Discuss in detail about the geographical enquiry.
3. Discuss the different approaches of enquiry.
4. Why literature review is a necessary part of a research?
5. What are the basic components of literature review?
6. What are the main objectives of the literature review?
7. What are the different ways of organizing the source of literature?
8. Discuss the structuring of the research questions.
9. What are the different research objectives? Discuss it with example.
10. Explain research hypothesis. What is the purpose of research hypothesis?
11. What are the main characteristics of research hypothesis?
12. Discuss the different concepts of testing hypothesis?
13. What are the relevant points while framing the questionnaire pertaining to the survey?
14. Discuss the preliminary steps to design questionnaire.
15. What are the different orders of putting question for a research?
16. Why is piloting the survey necessary for finalizing the questionnaires?
17. Discuss the different steps of a good, logical, systematic and scientific research.

Unit 2: ❒ Data Collection: Types and Sources of Data; Methods of Data Collection; Data Input and Editing

STRUCTURE

2.1 ❒ Introduction

'Data' is basically unorganized statistical facts and figures collected for some specific purposes, such as analysis. There can be different *sources* of data, such as statistical and non-statistical sources. Also, there are different *methods of data collection*, depending on the type of data. Data collection is a standout amongst the most essential stages in carrying on a research. Data collection is an extremely challenging work which needs exhaustive planning, diligent work, understanding, determination and more to have the capacity to complete the assignment effectively. Data collection begins with figuring out what sort of data is needed, followed by the collection of a sample from a certain section of the population. Certain tools are used to collectdata from the chosen sample.

2.2 ❒ Objectives

1. Understand the meaning of data.
2. Identify the various types of data.
3. Know the difference between primary and secondary data and their sources.
4. Conversant with the various data collection methods.
5. Know the advantages and disadvantages of each method.
6. Demonstrate skills in interviewing others while collecting data.
7. Design questionnaires to tap different variables.
8. Learn the process of data input and editing

2.3 ❒ Data Collection: Types and Sources of Data

2.3.1 Types of Data Collection

On the basis of collection, there are two types of data viz., primary data and secondary data.

- The primary data are those which are collected afresh and for the first time, and thus happen to be original in character.
- The secondary data,on the other hand, are those which have already been collected by someone else and which have already been passed through the statistical process.

The researcher would have to decide which sort of data he would be using (thus collecting) for his study and accordingly he will have to select one or the other method of data collection. The methods of collecting primary and secondary data differ since primary data are to be originally collected, while in case of secondary data the nature of data collection work is merely that of compilation.

On the basis of **nature**, there are two types of data – quantitative and qualitative and both are equally important.

- **Quantitative data** is information that can be measured i.e one can count. Because it is countable.

Examples:How much did it cost? Or what is the average attendance at each programme session? In this type of data, structured surveys are done involving multiple-choice, rating scales question.

- **Qualitative data** is information about qualities of attributes which is not countable. That is, it provides information about how people feel about something.

Examples:Sharing what people like about a programme? Or how they think it could be improved? Or whether they would recommend the programme to others? The literature reviews, observation or participant's experience and impact fall under this category.

2.3.2 Sources of Data

There are two sources of data in Statistics.

- Statistical sources refer to data that are collected for some official purposes and include censuses and officially conducted surveys.
- Non-statistical sources refer to the data that are collected for other administrative purposes or for the private sector, statistical survey.

A statistical survey is normally conducted using a sample. It is also called **Sample Survey**. It is the method of collecting sample data and analyzing it using statistical methods. This is done to make estimations about population characteristics. The advantage is that it gives the full control over the data. One can ask questions suited for his study that are carrying out. But, the disadvantage is that there is a chance of sample error creeping up. This is because a sample is chosen and the entire population is not studied. Leaving out some units of the population while choosing the sample causes this error to arise.

Opposite to a sample survey, a census is based on all items of the population and then data are analyzed. Data collection happens for a specific reference period. For example, the Census of India is conducted every 10 years. Other censuses are conducted roughly every 5-10 years. Data is collected using questionnaires that may be mailed to the respondents.

Responses can also be collected over other modes of communication like the telephone. An advantage is that even the most remote of the units of the population get included in the census method. The major disadvantage lies in the high cost of data collection and that it is a time-consuming process.

Registers are basically storehouses of statistical information from which data can be collected and analysis can be made. Registers tend to be detailed and extensive. It is

beneficial to use data from here as it is reliable. Two or more registers can be linked together based on common information for even more relevant data collection.

From agriculture to business, all industries maintain registers for record-keeping. Some administrative registers also serve the purpose of acting as a repository of data for other statistical bodies in a country (Kaishnaswamy).

2.4 ❒ Methods of Data Collection

2.4.1 Primary Data Collection

Primary data is collected from the first-hand experience and is not used in the past. The data gathered by primary data collection methods are specific to the motive of the research, and highly authentic and accurate.

Primary data collection methods can be divided into two categories: quantitative methods and qualitative methods. There are several methods of collecting primary data, particularly in surveys and descriptive researches. Important ones are:

- **Observation Method**

The observation method is the most commonly used method especially in studies relating to behavioural sciences. Observation becomes a scientific tool and the method of data collection for the researcher, when it serves a formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability. In case the observation is characterised by a careful definition of the units to be observed, the style of recording the observed information, standardised conditions of observation and the selection of pertinent data of observation, then the observation is called as **structured observation**. But when observation takes place without these characteristics to be thought of in advance, the same is termed as **unstructured observation**. Structured observation is considered appropriate in descriptive studies. If the observer observes by making himself, more or less, a member of the group he is observing so that he can experience what the members of the group experience, the observation is called as the **participant observation**. But when the observer observes as a detached emissary without any attempt on his part to experience through participation what others feel, the observation of this type is often termed as **non-participant observation**. If the observation takes place in the natural setting, it may be termed as uncontrolled observation, but when observation takes place according to definite pre-arranged plans, involving experimental procedure, the same is then termed **controlled observation.**

The main advantage of this method is that subjective bias is eliminated, if observation is done accurately. Secondly, the information obtained under this method relates to what is currently happening; it is not complicated by either the past behaviour or future intentions or attitudes. Thirdly, this method is independent of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents as happens to be the case in the interview or the questionnaire method. This method is particularly suitable in studies which deal with subjects (i.e., respondents) who are not capable of giving verbal reports of their feelings for one reason or the other.

However, observation method has various limitations. Firstly, it is an expensive method. Secondly, the information provided by this method is very limited. Thirdly, sometimes unforeseen factors may interfere with the observational task. At times, the fact that some people are rarely accessible to direct observation creates obstacle for this method to collect data effectively.

- **Interview Method**

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

a) **Personal interview method** requires a person known as the interviewer asking questions generally in a face-to-face contact to the other person or persons. This sort of interview may be in the form of direct personal investigation or it may be indirect oral investigation. It is of various types which is discussed below: **structured interviews** involve the use of a set of predetermined questions and of highly standardised techniques of recording. As against it, the **unstructured interviews** are characterised by a flexibility of approach to questioning. Unstructured interviews do not follow a system of pre-determined questions and standardised techniques of recording information. **Focused interview**is meant to focus attention on the given experience of the respondent and its effects. Under it the interviewer has the freedom to decide the manner and sequence in which the questions would be asked and has also the freedom to explore reasons and motives. Such interviews are used generally in the development of hypotheses and constitute a major type of unstructured interviews. The **clinical interview** is concerned with broad underlying feelings or motivations or with the course of individual's life experience. In case of **non-directive interview**, the

interviewer's function is simply to encourage the respondent to talk about the given topic with a bare minimum of direct questioning.

The chief merits of the interview method are as follows:

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

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

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

(iv) Observation method can as well be applied to recording verbal answers to various questions.

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

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

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

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

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

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

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

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

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

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

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

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

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

(vii) Interviewing at times may also introduce systematic errors.

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

b) Telephone interviews method of collecting information consists ofcontacting respondents on telephone.

The chief merits of such a system are:

1. It is more flexible in comparison to mailing method.
2. It is faster than other methods i.e., a quick way of obtaining information.
3. It is cheaper than personal interviewing method; here the cost per response is relatively low.
4. Recall is easy; callbacks are simple and economical.
5. There is a higher rate of response than what we have in mailing method; the non-responseis generally very low.
6. Replies can be recorded without causing embarrassment to respondents.
7. Interviewer can explain requirements more easily.
8. At times, access can be gained to respondents who otherwise cannot be contacted for onereason or the other.
9. No field staff is required.
10. Representative and wider distribution of sample is possible.

But this system of collecting information is not free from demerits. Some of these are highlighted below:

1. Little time is given to respondents for considered answers; interview period is not likely to exceed five minutes in most cases.
2. Surveys are restricted to respondents who have telephone facilities.
3. Extensive geographical coverage may get restricted by cost considerations.
4. It is not suitable for intensive surveys where comprehensive answers are required to various questions.
5. Possibility of the bias of the interviewer is relatively more.
6. Questions have to be short and to the point; probes are difficult to handle.

- **Through questionnaires**

In this method a questionnaire is sent (usually by post) to the persons concerned with a request to answer the questions and return the questionnaire. A questionnaire consists of a number of questions printed or typed in a definite order on a form or set of forms. This method of data collection is quite popular, particularly in case of big enquiries. It is being adopted by private individuals, research workers, private and public organisations and even by governments. Before using this method, it is always advisable to conduct 'pilot study' (Pilot Survey) for testing the questionnaires. In a big enquiry the significance of pilot survey is felt very much. Pilot survey is in fact the replica and rehearsal of the main survey.

Quite often questionnaire is considered as the heart of a survey operation. Hence it should be very carefully constructed. If it is not properly set up, then the survey is bound to fail. This fact requires us to study the main aspects of a questionnaire viz., the general form, question sequence and question formulation and wording.

So far as the general form of a questionnaire is concerned, it can either be structured or unstructured questionnaire. Structured questionnaires are those questionnaires in which there are definite, concrete and pre-determined questions. The form of the question may be either closed (i.e., of the type 'yes' or 'no') or open (i.e., inviting free response) but should be stated in advance and not constructed during questioning. Thus a highly structured questionnaire is one in which all questions and answers are specified and comments in the respondent's own words are held to the minimum. More specifically, we can say that in an unstructured questionnaire, the interviewer is provided with a general guide on the type of information to be obtained, but the exact question formulation is largely his own

responsibility and the replies are to be taken down in the respondent's own words to the extent possible; in some situations tape recorders may be used to achieve this goal.

In order to make the questionnaire effective and to ensure quality to the replies received, a researcher should pay attention to the question-sequence in preparing the questionnaire. The first few questions are particularly important because they are likely to influence the attitude of the respondent and in seeking his desired cooperation. The opening questions should be such as to arouse human interest. Ideally, the question sequence should conform to the respondent'sway of thinking.

Question should also be impartial in order not to give a biased picture of the true state of affairs. Questions should be constructed with a view to their forming a logical part of a well thought out tabulation plan. In general, all questions should meet the following standards—(a) should be easily understood; (b) should be simple i.e., should convey only one thought at a time; (c) should be concrete and should conform as much as possible to the respondent's way of thinking.

Concerning the form of questions, we can talk about two principal forms, viz., multiple choice question and the open-end question. In the former the respondent selects one of the alternative possible answers put to him, whereas in the latter he has to supply the answer in his own words. The question with only two possible answers (usually 'Yes' or 'No') can be taken as a special case of the multiple choice question, or can be named as a 'closed question.'

Researcher must pay proper attention to the wordings of questions since reliable and meaningful returns depend on it to a large extent. Since words are likely to affect responses, they should be properly chosen. Simple words, which are familiar to all respondents should be employed. Words with ambiguous meanings must be avoided. To be successful, questionnaire should be comparatively short and simple i.e., the size of the questionnaire should be kept to the minimum. Questions should proceed in logical sequence moving from easy to more difficult questions

The merits claimed on behalf of this method are as follows:

1. There is low cost even when the universe is large and is widely spread geographically.
2. It is free from the bias of the interviewer; answers are in respondents' own words.
3. Respondents have adequate time to give well thought out answers.

4. Respondents, who are not easily approachable, can also be reached conveniently.
5. Large samples can be made use of and thus the results can be made more dependable and reliable.

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

1. Low rate of return of the duly filled in questionnaires; bias due to no-response is often indeterminate.
2. It can be used only when respondents are educated and cooperating.
3. The control over questionnaire may be lost once it is sent.
4. There is in built inflexibility because of the difficulty of amending the approach once questionnaires have been despatched.
5. There is also the possibility of ambiguous replies or omission of replies altogether to certain questions; interpretation of omissions is difficult.
6. It is difficult to know whether willing respondents are truly representative.
7. This method is likely to be the slowest of all.

- **Through Schedules**

This method of data collection is very much like the collection of data through questionnaire, with little difference which lies in the fact that schedules (proforma containing a set of questions) are being filled in by the enumerators who are specially appointed for the purpose. These enumerators along with schedules, go to respondents, put to them the questions from the proforma in the order the questions are listed and record the replies in the space meant for the same in the proforma. This method of data collection is very useful in extensive enquiries and can lead to fairly reliable results. It is, however, very expensive and is usually adopted in investigations conducted by governmental agencies or by some big organisations. Population census all over the world is conducted through this method.

Both questionnaire and schedule are popularly used methods of collecting data in research surveys. The important points of difference are as under:

1. The questionnaire is generally sent through mail to informants to be answered as specified in a covering letter. The schedule is generally filled out by the research worker or the enumerator, who can interpret questions when necessary.
2. To collect data through questionnaire is relatively cheap and economical. Here no field staff required. To collect data through schedules is relatively more

expensive since considerable amount of money has to be spent in appointing enumerators and in importing training to them.

3. Non-response is usually high in case of questionnaire as many people do not respond and many return the questionnaire without answering all questions. Bias due to non-response often remains indeterminate. As against this, non-response is generally very low in case of schedules because these are filled by enumerators who are able to get answers to all questions. But there remains the danger of interviewer bias and cheating.
4. In case of questionnaire, it is not always clear as to who replies, but in case of schedule the identity of respondent is known.
5. The questionnaire method is likely to be very slow since many respondents do not return the questionnaire in time despite several reminders, but in case of schedules the information is collected well in time as they are filled in by enumerators.
6. Personal contact is generally not possible in case of the questionnaire method as questionnaires are sent to respondents by post who also in turn return the same by post. But in case of schedules direct personal contact is established with respondents.
7. Questionnaire method can be used only when respondents are literate and cooperative, but in case of schedules the information can be gathered even when the respondents happen to be illiterate.
8. Wider and more representative distribution of sample is possible under the questionnaire method, but in respect of schedules there usually remains the difficulty in sending enumerators over a relatively wider area.
9. Risk of collecting incomplete and wrong information is relatively more under the questionnaire method. But in case of schedules, the information collected is generally complete and accurate as enumerators can remove the difficulties. As a result, the information collected through schedules is relatively more accurate than that obtained through questionnaires.
10. The success of questionnaire method lies more on the quality of the questionnaire itself, but in the case of schedules much depends upon the honesty and competence of enumerators.

11. In order to attract the attention of respondents, the physical appearance of questionnaire must be quite attractive, but this may not be so in case of schedules as they are to be filled in by enumerators and not by respondents.

12. Along with schedules, observation method can also be used but such a thing is not possible while collecting data through questionnaires.

- **Some Other Methods of Data Collection**

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

i) Warranty cards are usually postal sized cards which are used by dealers of consumer durables to collect information regarding their products. It is placed inside the package along with the product with a request to the consumer to fill in the card and post it back to the dealer.

ii) Distributor or store audits are performed by distributors as well as manufactures through their salesmen at regular intervals. The data are obtained in such audits not by questioning but by observation.

iii) Pantry audit technique is used to estimate consumption of the basket of goods at the consumer level. In this type of audit, the investigator collects an inventory of types, quantities and prices of commodities consumed. Thus in pantry audit data are recorded from the examination of consumer's pantry. The usual objective in a pantry audit is to find out what types of consumers buy certain products and certain brands, the assumption being that the contents of the pantry accurately portray consumer's preferences.

iv) An extension of the pantry audit approach on a regular basis is known as 'consumer panel', where a set of consumers are arranged to come to an understanding to maintain detailed daily records of their consumption and the same is made available to investigator on demands.

v) The use of mechanical devices has been widely made to collect information by way of indirect means. Eye camera, Pupilometric camera, Psychogalvanometer, Motion picture camera and Audiometer are the principal devices so far developed and commonly used by modern big business houses, mostly in the developed world for the purpose of collecting the required information.

vi) Projective techniques (or what are sometimes called as indirect interviewing techniques) for the collection of data have been developed by psychologists to

use projections of respondents for inferring about underlying motives, urges, or intentions which are such that the respondent either resists to reveal them or is unable to figure out himself.

vii) Depth interviews are those interviews that are designed to discover underlying motives and desires and are often used in motivational research. Such interviews are held to explore needs, desires and feelings of respondents. As such, depth interviews require great skill on the part of the interviewer and at the same time involve considerable time.

viii) Content-analysis consists of analysing the contents of documentary materials such as books, magazines, newspapers and the contents of all other verbal materials which can be either spoken or printed. Content-analysis prior to 1940's was mostly quantitative analysis of documentary materials concerning certain characteristics that can be identified and counted. But since 1950's content-analysis is mostly qualitative analysis concerning the general import or message of the existing documents (Krishnaswamy).

2.4.2 Secondary Data Collection

Secondary data means data that are already available i.e., they refer to the data which have already been collected and analysed by someone else. Secondary data may either be published data or unpublished data. Usually published data are available in: (a) various publications of the central, state are local governments; (b) various publications of foreign governments or of international bodies and their subsidiary organisations; (c) technical and trade journals; (d) books, magazines and newspapers; (e) reports and publications of various associations connected with business and industry, banks, stock exchanges, etc.; (f) reports prepared by research scholars, universities, economists, etc. in different fields; and (g) public records and statistics, historical documents, and other sources of published information. The sources of unpublished data are many; they may be found in diaries, letters, unpublished biographies and autobiographies and also may be available with scholars and research workers, trade associations, labour bureaus and other public/ private individuals and organisations. Researcher must be very careful in using secondary data. Dr. A.L. Bowleyvery aptly observes that it is never safe to take published statistics at their face value without knowing their meaning and limitations and it is always necessary to criticise arguments that can be based on them.

Before using secondary data, the following characteristics must be checked –

1. **Reliability of data can be checked from following aspect**
   - Who collected the data?
   - From what source?
   - Which methods?
   - Time?
   - Possibility of bias?
   - Accuracy?

2. Suitability of data – The object, scope and nature of the original enquiry must be studied and then carefully scrutinize the data for suitability.

3. Adequacy – The data is considered inadequate if the level of accuracy achieved in data is found inadequate or if they are related to an area which may be either narrower or wider than the area of the present enquiry.

Table 1: Differences between Primary and Secondary Data

	Primary Data		Secondary Data
1.	Primary data are those data which are collected from the primary sources.	1.	Secondary data are those data which are collected from the secondary sources.
2.	Primary data are known as basic data.	2	Secondary data are known as subsidiary data.
3.	The collection of primary data is more expensive.	3.	The collection of secondary data is comparatively less expensive.
4.	It takes more time to collect the data.	4.	It takes less time to collect the data.
5.	Primary data are more accurate.	5.	Secondary data are less accurate than the primary data.
6.	Primary data are known as first hand data.	6.	Secondary data are known as second hand data.
7.	Primary data are not readily available.	7.	Subsidiary data are readily available.
8.	It is required to take much care at the time of collecting data	8.	It is not required to take much care a the time of collecting data.

2.5 ❒ Data Input and Editing

2.5.1 Data Input

In case of Social Science, data can be either Analogue (non-digital) spatial data or Digital spatial data. Analogue data are normally in paper form and include paper maps, tables of statistics and hardcopy of aerial photographs. All these forms of data need to be converted to digital form before use in a GIS (Geographical Information System).Digiltal data like remote sensing data are already in computer-readable formats and are supplied on diskette, magnetic tape or CD- ROM or across a computer network.There are four methods of data input which are widely used: –

i) **Keyboard entry** - This method is also called as key-coding. It is the entry of data into file for GIS at a computer terminal. This technique is used for attribute data that are available only on paper. This technique can be mixed with digitizing process for the creation of GIS database. The attribute data, 'once in digital format, are linked to the relevant map features in the spatial database using identification codes. There are unique codes that are allocated to each point, line and area feature in the dataset. The coordinates of spatial entities like point, line and area features can be encoded by keyboard entry. However, it will not be convenient when the number of features are more. This method leads to obtain very high level of precision data by entering the actual surveying measurements. Used for entering land information during digitization can be used for drawing maps using survey measurements.

ii) **Manual digitizing** - It is the most common method of encoding spatial features from paper maps. It is a process of converting the spatial features on a map into a digital format. Point, line, and area features that form a map, are converted into (x, y) coordinates. A point is represented by a single coordinate, a line by a string of coordinates, and, when one or more lines are combined with a label point inside an outline, then an area (polygon) is identified. Hence, digitizing is called as the process of capturing a series of points and lines.

In this process, the map is affixed to a digitizing table. Three or more control points are to be identified and digitized for each map sheet, example intersection of roads, historic monuments etc. The points are called reference points or tics or control points. The

coordinates of these points are known to the person who is doing digitizing. The coordinates of these points are used by the system to perform necessary mathematical operations and also to calculate coordinates of remaining features present in the map. Some of the common problems in digitizing paper maps are: Paper maps are unstable; each time the map is removed from the digitizing table, the reference points must be re-entered when the map is affixed to the table again. If the map has stretched or shrunk in the interim, the newly digitized points will be slightly off in their location.

iii) **Automatic digitization and** iv) **Scanning** are the most commonly used method. This method is normally used when raster type of data is to be produced from analogue maps. And the scanned image can be used as the base for digitizing the features in vector format.

Other than the above mentioned methods of files are directly transferred into GIS which is called direct file translation (Krishnaswamy).

2.5.2 Data Editing

The data which has been collected from various primary/secondary sources is raw in nature, this means that there are likely chances of errors and inconsistencies in it. Since data collected is of pivotal importance to policy and decision makers everywhere be it governmental departments, business organizations, health or educational institutions etc. it would be better to have a team of experts at hand who know how to scrutinize, review and edit this data before it is finally fed into the data bases and the required statistics are generated.

The process through which the data is reviewed to check for consistency, adequacy, detect errors and outliers (values that are either too big or too small from the rest of the data) and the correction of errors within the data in order to maximize its usefulness for the purpose for which it was collected is called data editing.

The basic purpose served by data editing is that it improves the quality, accuracy and adequacy of the collected data thereby making it more suitable for the purpose for which the data was collected. The following can therefore be identified as the main objectives of the data editing process:

- Detection of errors in the data that otherwise affect the validity of outputs.

- Validation of data for the purposes it was collected.
- Provision of information that would help access the overall level of accuracy of the data.
- Detection and identification of any inconsistencies in the data and outliers and to make adjustments for them.

The different types of data editing are discussed below:

- **Validity and completeness of data** refers to correctness and completeness of obtained responses. This helps ensure that there are no missing values or empty fields in the data bases.
- **Range** verifies that data within a field fall between the boundaries specified for the particular field.
- **Duplicate data entry** helps ensure that there is no repetition or duplication of data and each unit on the data base or register was filled only once.
- **Logical consistency:** Through this type of editing connections between data fields or variables are taken into account.
- **Outliers:** This type of editing helps to detect values that are too extreme or unusual so that they can be verified and checked.

The data editing process can be divided intofollowing three parts or stages:

- The first stage is the stage where rules are set for editing. This stage is further subdivided into following two steps.

In step one, you provide instructions to desk editors who then check the data for coherence and consistency.

In step two, you set the rules by establishing logical relations between the variables according to various criteria. This set of rules is called automated validation rules and this type of editing seeks to detect errors during data entry and to screen them.

- The **manual desk editing** stage is a traditional method that is put into effect by a specialized editing team. The data,(if) on paper is checked after the data has been collected and before it is fed into the data bases. If however, electronic means have been used to collect the data, the forms entered into the database are revised individually.

- The **automated data editing** method makes use of computer programs and systems for checking the data all at once after it has been entered electronically. These programmes and systems contain **Audit rules** which validate the data, detect errors and determine unacceptable responses.

There are certain things that must be kept in mind when editing the data, they are as follows :

- Who should make or set the editing rules? The answer to this question would be that such rules should be made by professionals who are **experts** in data collection, questionnaire design and analysis.
- The editing rules need to be consistent and free from any contradictions.
- When setting the editing rules, it must be established whether the variable is qualitative or quantitative because the rules for editing either one are different from the other.
- Give enough time to each of the various stages of the process, that is, data collection entry and analysis and at the end of each make a quick check to see that all the necessary edits have been made and that there are no empty places within the questionnaire form.
- The questionnaire must be edited in full during the early stages of editing. If however it is found that some errors remain, a sample of forms should be subjected to re editing. The size of the sample is determined according to the expected number of the remaining errors.
- The desk editing stage isto ensure that the data is almost free from all errors.
- The questionnaire must be subjected to desk editing and also to automated rules within the built in data editing/computer software programs.

There are certain factors that can influence or pose limitations to the data editing process, these can be summed up as follows:

- Data editing can be influenced by the amount of time available, the budget, the presence or absence of other resources and also by the group of people involved in the editing process.

- The available computer software programs.
- Follow up with the respondents is of critical importance in the data editing process because they are often the best source of information in many cases. However, the respondents might feel this to be stressful and burdensome thereby causing limitations to the data editing process.
- Some types of data do not require extensive editing, therefore it would be better to keep in mind the intended uses of data and make sure that the more important part of data is kept free from all errors. In this way, the intended use of data does play an important role in influencing the data editing process.
- Methods and procedures must be established and must be followed while correcting or handling the data errors, in the survey plan, right at the start of the project otherwise the process would be of no or little use.
- It must be remembered that if one plan is to edit the data manually, one must develop and document the methods that are to be followed. The team must be trained, a method must be established to check their work progress and the impact of the edits on the original data must also be assessed.
- In case of automated editing, the rules for editing must be developed and well documented. Develop software or customize an existing computer programme as per data editing demand or requirement.

2.6 ❒ Conclusion

Data collection must be done enormously to predict correct result of analysis in a research. One have decided how to get information whether by direct observation, interviews, surveys, experiment and testing or other methods. Recording and arranging of information or detain a distinctive structure, contingent upon the nature of data that anyone collected. Research is an endless process because as time changes strategy of reports containing details also vary due to respondent are not same in nature. A research about a topic gives overview, detailed and explanation according to research types. At last collection of data are most important for research because it acts as proof or evidence of the valuable reports.

2.7 ❒ Summary

(i) To understand the multitude of choices available to a researcher for collecting the information, one needs to be fully cognizant of the resources available. (ii) The information units are known as the data. (iii) On the basis of sources, the data is divided into two types i.e. primary and secondary. (iv) The data sources could be either contextual and primary or historical and secondary in nature. The secondary data can be collected from various published or written or available sources. (v) Although it is not very safe to use this type of data, therefore some amount of check is required for the usage of secondary data. (vi) After collecting the data, data input and editing are involved. Then the data is useful for the research.

2.8 ❒ Glossary/Keywords

Primary Data are the first hand information gathered through interaction or interview is known as primary data.

Secondary data are the information collected from the sources where it is already available.

2.9 ❒ References and Further Readings

Chandan, J.S. (2009) *Statistics for Business and Economics*, New Delhi: Vikas Publishing House Pvt. Ltd.

Chawla D. and NeenaS. (2016)*Research Methodology: Concept and Cases,* Vikas Publishing House: New Delhi.

Churchill, Gilbert A (1983) Marketing Research: Methodological Foundations, The Dryden Press, New York.

Gaur, A.S. and Gaur, S.S., (2009) *Statistical Methods for Practice and Research*, New Delhi: Response Books.

Kothari C.R. (2004) *Research Methodology: Methods and Techniques*, New Delhi: New Age International Publishers.

Krishnaswami, O.R. and Ranganatham, M.(2016) *Methodology of Research in Social Sciences*, 2/e, Mumbai: Himalaya Publishing House.

Mahalotra N.K. (2002) Marketing Research: An Applied Orientation. Pearson Education Asia.

Mustafi, C.K. (1981) Statistical Methods in Managerial Decisions, Macmillan: New Delhi. Punch, Keith, F., *Survey Research – the Basics*, New Delhi: Sage Publications.

Raj, D. (1968), "Sampling Theory," McGraw-Hill Book Company, New York.

Singh, D. and F.S. Chaudhary, (1986) Theory and Analysis of Sample Survey Designs, Wiley Eastern: New Delhi.

Walliman, N. (2016)*Social Research Methods,* New Delhi: Sage Publications.

Yates, E (1960), "Sampling Methods for Censuses and Surveys," Charles Griffin & Company, Ltd., London.

2.10 ❒ Model Questions

1. What is meant by data?
2. What are different methods of data classification? Discuss it elaborately.
3. What is sample survey? Write the difference between sample survey and census.
4. How is secondary data economical?
5. Enumerate the different methods of collecting primary data.
6. Write the difference between structured and unstructured questionnaire.
7. Write short notes:
   (a) Observation method of primary data collection
   (b) Interview method of primary data collection
   (c) Primary data collection through questionnaires
   (d) Primary data collection through schedules
8. Write the difference between primary and secondary data.

9. Write the difference between questionnaires and schedules.
10. Discuss interview as a technique of data collection.
11. What are the guiding considerations in the construction of questionnaire? Explain.
12. Discuss the methods of data input in details.
13. What is data editing? Discuss the various types of data editing.
14. Elaborate the process of data editing.
15. Write the limitations of data editing.

Unit 3: ❒ Data Analysis: Qualitative and Quantitative Analysis; Techniques of Data Representation

STRUCTURE

3.1 ❒ Introduction

The data, after collection, has to be processed and analysed in accordance with the outline laid down for the purpose at the time of developing the research plan. Technically speaking, processing implies editing, coding, classification and tabulation of collected data so that they are amenable to analysis. The term analysis refers to the computation of certain measures along with searching for patterns of relationship that exist among data-groups. Different techniques are used to depict the data analysis so that anyone can understand the analysis.

3.2 ❒ Objectives

The objectives of this chapter are as follows:

1. Elaborating different types of data analysis.
2. Describing the different methods of data analysismeasures like central tendency, frequency distribution etc.

3. Describing the different methods of diagrammatic representation of data like histogram, bar chart, pie chart etc.

3.3 ❐ Data Analysis: Qualitative and Quantitative Analysis

Data analysis refers to a variety of specific procedures and methods. Data analysis involves goals,relationships, decision making, and ideas in addition to working with the actual data itself. Simply speaking data analysis includes ways of working with information (data) to support the work, goals and plans for anyresearch work. From this perspective, a data analysis process includes the following key components: Purpose, Questions, Data Collection, Data Analysis Procedures and Methods, Interpretation/Identification of Findings, Writing, Reporting, and Dissemination and Evaluation.

3.3.1 Quantitative Procedures for Data Analysis –

There are a variety of procedures that may be used to analyze quantitative data. The two most basic types are:

1. Summary Measures:

In case of summary measures, the following questionsneed to be considered:

- How do the data converge (come together)?
- What is a “typical” (average) value?
- Where is the middle (center) of a group?

Types of Summary Measures: Mean, Median and Mode

Mean: It is also known as arithmetic average, is the most common measure of central tendency and may be defined as the value which we get by dividing the total of the values of various given items in a series by the total number of items.

$$\text{Mean (or } \bar{X}) = \frac{\Sigma X_i}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

where $\bar{X}$ = The symbol we use for mean (pronounced as X bar)

Σ = Symbol for summation

X_i = Volue of the i-th item X, i = 1, 2, ..., n

n = total number of items

Median: Median is the value of the middle item of series when it is arranged in ascending or descending order of magnitude. It divides the series into two halves, in one half all items are less than median value, whereas in the other half all items have values higher than median. It is a positional average.

$$\text{Median } (M) = \text{Value of } \left(\frac{n+1}{2}\right) \text{ th item}$$

Mode: Mode is the most commonly or frequently occurring value in a series. The mode in a distribution is that item around which there is maximum concentration. In general, mode is the size of the item which has the maximum frequency, but at items such an item may not be mode on account of the effect of the frequencies of the neighboring items. Like median, mode is a positional average and is not affected by the values of extreme items.

2. Variance Measures:

Questions need to be considered when working with variance measures include:

- How do scores differ?
- What are the differences between individuals in a group?
- What is the range of outcomes?

Example of variance measures: 50 participants rate a training using the following template:

The answer of the questions can be displayed in different ways like:

I Acquire new knowledge and/or skills	**Response in 1st training**	**%**	**Response in 2nd training**	**%**	**Response in 3rd training**	**%**
Strongly Disagree	45	90	0	0	15	30
Disagree	3	6	5	10	13	26
Agree	1	2	5	10	14	28
Strongly Agree	1	2	40	80	8	16
Total	50	100	50	100	50	100

The importance of measuring variance is that it demonstrates differences. Once the difference is identified, then you can seek an explanation and determine if it is significant. In the above examples, one might determine that the first training was not valuable to staff

and one may reconsider the topic or the person delivering the training; the second training was valuable and one could consider a follow-up and continue to provide the training; for the third training you could consider the target audience and re-evaluate the material being covered and appropriateness for some staff such as teachers versus family service workers. In these cases further information may be needed as to why there was low or high variability in the responses (Kothari).

(1) Measures of Dispersion

Averages can represent a series only as best as a single figure can, but it certainly cannot reveal the entire story of any phenomenon under study. Especially it fails to give any idea about the scatter of the values of items of a variable in the series around the true value of average. In order to measure this scatter, statistical devices called measures of dispersion are calculated. Important measures of dispersion are (a) range, (b) mean deviation, and (c) standard deviation.

(a) Range is the simplest possible measure of dispersion and is defined as the difference between the values of the extreme items of a series.

$$Range = \begin{pmatrix} Highest\ value\ of\ an \\ item\ in\ a\ series \end{pmatrix} - \begin{pmatrix} Lowest\ value\ of\ an \\ item\ in\ a\ series \end{pmatrix}$$

(b) Mean Deviation is the average of difference of the values of items from some average of the series. Such a difference is technically described as deviation. In calculating mean deviation we ignore the minus sign of deviations while taking their total for obtaining the mean deviation. When mean deviation is divided by the average used in finding out the mean deviation itself, the resulting quantity is described as the coefficient of mean deviation. Coefficient of mean deviation is a relative measure of dispersion and is comparable to similar measure of other series.

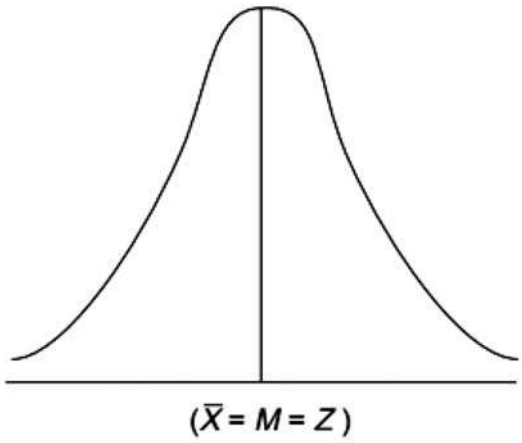
Mean deviation from mean $(\delta_x) = \dfrac{\Sigma \left| X_i - \bar{X} \right|}{n}$, if deviations, $\left| X_i - \bar{X} \right|$, are obtained from arithmetic average.

(c) Standard Deviation is most widely used measure of dispersion of a series and is commonly denoted by the symbol 'σ' (pronounced as sigma). Standard deviation is defined as the square-root of the average of squares of deviations,

when such deviations for the values of individual items in a series are obtained from the arithmetic average. When we divide the standard deviation by the arithmetic average of the series, the resulting quantity is known as coefficient of standard deviation which happens to be a relative measure and is often used for comparing with similar measure of other series. When this coefficient of standard deviation is multiplied by 100, the resulting figure is known as coefficient of variation. Sometimes, we work out the square of standard deviation, known as variance, which is frequently used in the context of analysis of variation.

$$\text{Standard deviation } (\sigma) = \sqrt{\frac{\Sigma\left(X_i - \bar{X}\right)^2}{n}}$$

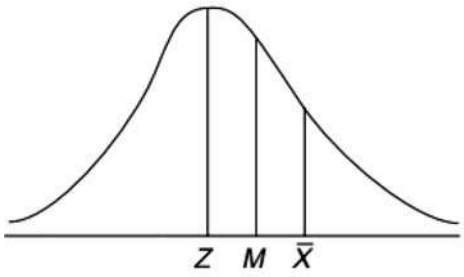
(2) Measures of asymmetry (skewness): When the distribution of item in a series happens to be perfectly symmetrical and bell shaped, then it is called normal curve.



Curve showing no skewness in which case we have $\bar{X} = M = Z$

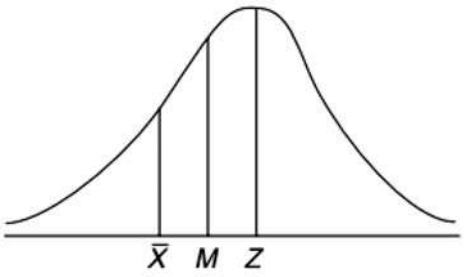
But if the curve is distorted (whether on the right side or on the left side), we have asymmetrical distribution which indicates that there is skewness. If the curve is distorted on the right side, we have positive skewness but when the curve is distorted towards left, we have negative skewness. In case of positive skewness, we have Mode (Z) < Median (M) < Mean (X) and in case of negative skewness we have Mean < Median < Mode.

Kurtosis is the measure of flat-toppedness of a curve. A bell shaped curve or the normal curve is Mesokurtic because it is kurtic in the centre; but if the curve is relatively



Curve showing positive shewness In case of positive skewness we have:

$$Z < M < \bar{X}$$



Curve showing positive shewness In case of negative skewness we have:

$$\bar{X} < M < Z$$

more peaked than the normal curve, it is called Leptokurtic whereas a curve is more flat than the normal curve, it is called Platykurtic.

(3) Measures of Relationship: We have dealt with those statistical measures that we use in context of univariate population i.e., the population consisting of measurement of only one variable. But if we have the data on two variables, we are said to have a bivariate population and if the data happen to be on more than two variables, the population is known as multivariate population. If for every measurement of a variable, X, we have corresponding value of a second variable, Y, the resulting pairs of values are called a bivariate population (Kothari).

In simple regression, we have only two variables, one variable (defined as independent) is the cause of the behaviour of another one (defined as dependent variable). Regression can only interpret what exists physically i.e., there must be a physical way in which independent variable X can affect dependent variable Y. The basic relationship between X and Y is given by

$$\hat{Y} = a + bX$$

where the symbol $\hat{Y}$ denotes the estimated value of Y for a given value of X. This equation is known as the regression equation of Y on X (also represents the regression line of Y on X when drawn on a graph) which means that each unit change in X produces a change of b in Y, which is positive for direct and negative for inverse relationships. Then generally used method to find the 'best' fit that a straight line of this kind can give is the least-square method. To use it efficiently, we first determine

$$\Sigma x_i^2 = \Sigma x_i^2 - n\bar{X}^2$$

$$\sum y_i^2 = \sum Y_i^2 - n\bar{Y}^2$$

$$\sum x_i y_i = \sum X_i Y_i - n\bar{X} \cdot \bar{Y}$$

$$b = \frac{\sum x_i y_i}{\sum x_i^2}, a = \bar{Y} - b\bar{X}$$

These measures define a and b which will give the best possible fit through the original X and Y points and the value of r can then be worked out as under:

$$r = \frac{b\sqrt{\sum x_i^2}}{\sqrt{\sum y_i^2}}$$

In addition, we may also have a corresponding value of the third variable, Z, or the forth variable, W, and so on; the resulting pairs of values are called a multivariate population.

In case of bivariate population: Correlation can be studied through

(a) Cross Tabulation: Cross tabulation approach is especially useful when the data are in nominal form. Under it we classify each variable into two or more categories and then cross classify the variables in these subcategories. A symmetrical relationship is one in which the two variables vary together, but we assume that neither variable is due to the other.

(b) Charles Spearman's Coefficient of Correlation: It is the technique of determining the degree of correlation between two variables in case of ordinal data where ranks are given to the different values of the variables. The main objective of this coefficient is to determine the extent to which the two sets of ranking are similar or dissimilar.

$$\text{Spearman's coefficient of correlation (or } r_s) = 1 - \left[\frac{6\sum d_i^2}{n\left(n^2 - 1\right)}\right]$$

where d_i = difference between ranks of ith pair of the two variables;

n = number of pairs of observations,

(d) Karl Pearson's Coefficient of Correlation : It is the most widely used method of measuring the degree of relationship between two variables. This coefficient assumes the following: (i) that there is linear relationship between the two variables; (ii) that the two variables are casually related which means that one of the variables is independent and the other one is dependent; and (iii) a large number of independent causes are operating in both variables so as to produce a normal distribution.

Karl Pearson's coefficient of correlation (or r)* = $\frac{\Sigma\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{n.\sigma_X.\sigma_Y}$

* Alternatively, the formula can be written as :

$$r = \frac{\Sigma\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\Sigma\left(X_i - \bar{X}\right)^2 \cdot \Sigma\left(Y_i - \bar{Y}\right)^2}}$$

Or

$$r = \frac{\sum X_i Y_i - n \cdot \bar{X} \cdot \bar{Y}}{\sqrt{\sum X_i^2 - n\bar{X}^2}\ \sqrt{\sum Y_i^2 - n\bar{Y}^2}}$$

where X_i = i th value of X variable

$\bar{X}$ = mean of X

Y_i = i th value of Y variable

$\bar{Y}$ = Mean of Y

n = number of pairs of observations of X and Y

Karl Pearson's coefficient of correlation is also known as the product moment correlation co-efficient. The value of 'r' lies between ± 1. Positive values of r indicate positive correlation between the two variables (i.e., changes in both variables take place in the statement direction), whereas negative values of 'r' indicate negative correlation i.e., changes in the two variables taking place in the opposite directions. A zero value of 'r' indicates that there is no association between the two variables. When r = (+) 1, it indicates perfect positive correlation and when it is (–)1, it indicates perfect negative correlation, meaning

thereby that variations in independent variable (X) explain 100% of the variations in the dependent variable (Y).

In case of multivariate population: Correlation can be studied through (a) coefficient of multiple correlation; (b) coefficient of partial correlation; whereas cause and effect relationship can be studied through multiple regression equations.

(5) Other measures.

a) Index numbers – When series are expressed in same units, we can use averages for the purpose of comparison, but when the units in which two or more series are expressed happen to be different, statistical averages cannot be used to compare them. In such situations we have to rely upon some relative measurement which consists in reducing the figures to a common base. Once such method is to convert the series into a series of index numbers. This is done when we express the given figures as percentages of some specific figure on a certain data. We can, thus, define an index number as a number which is used to measure the level of a given phenomenon as compared to the level of the same phenomenon at some standard date. The use of index number weights more as a special type of average, meant to study the changes in the effect of such factors which are incapable of being measured directly. But one must always remember that index numbers measure only the relative changes. Changes in various economic and social phenomena can be measured and compared through index numbers.

However index numbers have their own limitations with which researcher must always keep him aware. For instance, index numbers are only approximate indicators and as such give only a fair idea of changes but cannot give an accurate idea.

b) Analysis of Time Series – Series of successive observations of the given phenomenon over a period of time are referred to as time series. There are various methods of isolating trend from the given series viz., the free hand method, semi-average method, method of moving averages, method of least squares and similarly there are methods of measuring cyclical and seasonal variations and whatever variations are left over are considered as random or irregular fluctuations. The analysis of time series is done to understand the dynamic conditions for achieving the short term and long-term goals of business firm(s). The past trends can be used to evaluate the success or failure of management policy or policies practiced hitherto. On the basis of past trends, the future patterns can be predicted and policy or policies may accordingly be formulated.

Apart from these the co-efficient of contingency is another measure that may as well be used by a researcher, depending upon the nature of the problem under study.

Other Used Terminology for Quantitative Data Analysis:

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 categorized as descriptive analysis and inferential analysis (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 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 analyses 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 ways 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.

Descriptive statistics concern the development of certain indices from the raw data, whereas inferential statistics concern with the process of generalisation. Inferential statistics are also known as sampling statistics and are mainly concerned with two major type of problems: (i) the estimation of population parameters, and (ii) the testing of statistical hypotheses (Kothari).

3.3.2 Qualitative Procedures for Data Analysis

Qualitative data analysis is a process that seeks to reduce and make sense of vast amounts of information, often from different sources, so that impressions that shed light on a research question can emerge. There are also a number of ways that qualitative data can be analyzed. We will concentrate on a narrative analysis method using questionnaires and interviews (discuss in previous chapter).

3.4 ❐ Techniques of Data Representation

Methods of presentation must be determined according to the data format, the method of analysis to be used and the information to be emphasized. If one wishes to compare or introduce two values at a certain time point, it is appropriate to use text or the written language. However, a table is the most appropriate when all information requires equal attention, and it allows readers to selectively look at information of their own interest. Graphs allow readers to understand the overall trend in data, and intuitively understand the

comparison results between two groups. One thing to always bear in mind regardless of what method is used, however, is the simplicity of presentation.

Therefore, data can be presented in one of the three ways:

i) Text Presentation

Text is the main method of conveying information as it is used to explain results and trends, and provide contextual information. Data are fundamentally presented in paragraphs or sentences. Text can be used to provide interpretation or emphasize certain data. If quantitative information to be conveyed consists of one or two numbers, it is more appropriate to use written language than tables or graphs. If more data are to be presented, or other information such as that regarding data trends are to be conveyed, a table or a graph would be more appropriate. By nature, data take longer to read when presented as texts and when the main text includes a long list of information, readers and reviewers may have difficulties in understanding the information.

ii) Table Presentation

Tables, which convey information that has been converted into words or numbers in rows and columns, have been used for nearly 2,000 years. Anyone with a sufficient level of literacy can easily understand the information presented in a table. Tables are the most appropriate for presenting individual information, and can present both quantitative and qualitative information. The strength of tables is that they can accurately present information that cannot be presented with a graph. A number such as "132.145852" can be accurately expressed in a table. Strength is that information with different units can be presented together. Finally, tables are useful for summarizing and comparing quantitative information of different variables. However, the interpretation of information takes longer in tables than in graphs, and tables are not appropriate for studying data trends. Furthermore, since all data are of equal importance in a table, it is not easy to identify and selectively choose the information required.

iii) Graph Presentation

Graphs simplify complex information by using images and emphasizing data patterns or trends, and are useful for summarizing, explaining, or exploring quantitative data. While graphs are effective for presenting large amounts of data, they can be used in place of tables to present small sets of data. A graph format that best presents information must be

chosen so that readers and reviewers can easily understand the information. In the following, we describe frequently used graph formats:

Scatter plots present data on the x- and y-axes and are used to investigate an association between two variables. A point represents each individual or object, and an association between two variables can be studied by analyzing patterns across multiple points. A regression line is added to a graph to determine whether the association between two variables can be explained or not. If multiple points exist at an identical location, the correlation level may not be clear. In this case, a correlation coefficient or regression line can be added to further elucidate the correlation.

A bar graph is used to indicate and compare values in a discrete category or group, and the frequency or other measurement parameters (i.e. mean). Depending on the number of categories, and the size or complexity of each category, bars may be created vertically or horizontally. The height (or length) of a bar represents the amount of information in a category. Bar graphs are flexible, and can be used in a grouped or subdivided bar format in cases of two or more data sets in each category. By comparing the endpoints of bars, one can identify the largest and the smallest categories, and understand gradual differences between each category. One form of vertical bar graph is the stacked vertical bar graph. A stack vertical bar graph is used to compare the sum of each category, and analyze parts of a category. While stacked vertical bar graphs are excellent from the aspect of visualization, they do not have a reference line, making comparison of parts of various categories challenging.

A pie chart is used to represent nominal data (in other words, data classified in different categories), visually represents a distribution of categories. It is generally the most appropriate format for representing information grouped into a small number of categories. It is also used for data that have no other way of being represented aside from a table (i.e. frequency table).

A line plot is useful for representing time-series data such as monthly precipitation and yearly unemployment rates; in other words, it is used to study variables that are observed over time. Line graphs are especially useful for studying patterns and trends across data that include climatic influence, large changes or turning points, and are also appropriate for representing not only time-series data, but also data measured over the progression of a continuous variable such as distance. In a line graph, the x-axis represents the continuous variable, while the y-axis represents the scale and measurement values. It is also useful to

represent multiple data sets on a single line graph to compare and analyze patterns across different data sets.

A box and whisker chart does not make any assumptions about the underlying statistical distribution, and represents variations in samples of a population; therefore, it is appropriate for representing non parametric data. A box and whisker chart consists of boxes that represent interquartile range (one to three), the median and the mean of the data, and whiskers presented as lines outside of the boxes. Whiskers can be used to present the largest and smallest values in a set of data or only a part of the data (i.e. 95% of all the data). Data that are excluded from the data set are presented as individual points and are called outliers. The spacing at both ends of the box indicates dispersion in the data (Kothari).

3.5 ❒ Conclusion

In order to represent the data, it is necessary to identify the nature of data. Once the data analysis is done the effective data presentation tools are required. Text, tables, and graphs are effective communication media that present and convey data and information. They aid readers in understanding the content of research, sustain their interest, and effectively present large quantities of complex information.

3.6 ❒ Summary

This chapter provides the basics of data analysis. (i) The various measures of qualitative and quantitative data analysis are defined. (ii) It also described the different methods of data presentation for the target audience to enhance uptake of the findings.

3.7 ❒ Glossary/Keywords

Analysis: An investigation of the sample of a population and their relations in making up the population

Correlation: A statistical relation between two or more variables such that systematic changes in the value of one variable are accompanied by systematic changes in the other

Data: A collection of facts from which conclusions may be drawn (Example: “Statistical data”)

Mean: An average of n numbers computed by adding some function of the numbers and dividing by some function of n

Median: The middle most value of the data set when the data is arranged in ascending or descending order

Mode: The most frequent value of a random variable

Qualitative: Involving distinctions based on qualities (Example: "Qualitative change") or relating to or involving comparisons based on qualities.

Quantitative: Expressible as a quantity of relating to or susceptible of measurement (Example: "Export wheat without quantitative limitations") or relating to the measurement of quantity (Example: "Quantitative studies")

Standard Deviation: The square root of the variance.

Variance: The second moment around the mean; the expected value of the square of the deviations of a random variable from its mean

3.8 ❒ References and Further Readings

Chandan, J.S. (2009) *Statistics for Business and Economics*, New Delhi: Vikas Publishing House Pvt. Ltd.

Chawla D. and NeenaS. (2016)*Research Methodology: Concept and Cases,* Vikas Publishing House: New Delhi.

Gaur, A.S. and Gaur, S.S., (2009) *Statistical Methods for Practice and Research*, New Delhi: Response Books.

Kothari C.R. (2004) *Research Methodology: Methods and Techniques*, New Delhi: New Age International Publishers.

Krishnaswami, O.R. and Ranganatham, M.(2016) *Methodology of Research in Social Sciences*, 2/e, Mumbai: Himalaya Publishing House.

Punch, Keith, F., (1998) *Survey Research – the Basics*, New Delhi: Sage Publications.

Walliman, N. (2016)*Social Research Methods,* New Delhi: Sage Publications.

3.9 ❒ Model Questions

1. Write a brief note on different types of analysis of data pointing out the significance of each.
2. Discuss the quantitative data analysis methods.
3. Write short notes on:
   a) Multiple regression analysis
   b) Multiple discriminant analysis
   c) Multivariate analysis of variance
   d) Canonical analysis
   e) Skewness
   f) Cross tabulation
   g) Index number
   h) Time series analysis
   i) Value of r in Karl Pearson's coefficient of correlation
   j) Descriptive analysis and inferential analysis
   k) Correlation analysis and Causal analysis
   l) Multivariate analysis.
   m) Univariate, bivariate and multivariate data
4. What do you mean by multivariate analysis? Explain how it differs from bivariate analysis.
5. What are the demerits of text representation of data?
6. What are the merits of table representation of data?
7. What are the different types of graph to represent data? Explain.
8. What does a measure of central tendency indicate? Describe the important measures of central tendency pointing out the situation when one measure is considered relatively appropriate in comparison to other measures.

9. "The analysis of time series is done to understand the dynamic conditions for achieving the short-term and long-term goals of business firms." Discuss.

10. "Changes in various economic and social phenomena can be measured and compared through index numbers". Explain this statement pointing out the utility of index numbers.

11. What are the different methods of graphical data representation? Discuss in details.

Unit 4: ❒ Structure of a Research Report: Preliminaries, Text, Citation, Notes, References, Bibliography, Abstract and Key Words

STRUCTURE

4.1 ❒ Introduction

Once the task of data collection and data analysis is accomplished, the researcher has to initiate the task of drawing inferences followed by report writing. The analytical information and consequential inference(s) may well be communicated, preferably through research

report, to the consumers of research results who may be either an individual or a group of individuals or some public/private organization. Therefore, research report is considered a major component of the research study for the research task remains incomplete till the report has been presented and/or written. As a matter of fact even the most brilliant hypothesis, highly well designed and conducted research study, and the most striking generalizations and findings are of little value unless they are effectively communicated to others.

The purpose of research is not well served unless the findings are made known to others. Research results must invariably enter the general store of knowledge. All these explain the significance of writing research report. Some researchers do not consider writing a report as an integral part of the research process. But the general opinion is in favour of treating the presentation of research results or the writing a report as part and parcel of the research project. Writing a report is the last step in a research study and requires a set of skills and competencies somewhat different from those called for in respect of the earlier stages of research. This task should be accomplished by the researcher with utmost care and one may seek the assistance and guidance of experts for the purpose.

4.2 ❒ Objectives

The main objective of this chapter is to offer a set of guidelines intended to help researchers plan and write research reports that are well-organized, readable and presented in formats consistent with generally accepted practice.

4.3 ❒ Structure of a Research Report

Research reports are the product of slow, painstaking, accurate inductive work. The usual steps involved in writing report are:

(a) Logical analysis of the subject-matter;

(b) Preparation of the final outline;

(c) Preparation of the rough draft;

(d) Rewriting and polishing;

(e) Preparation of the final bibliography and references; and

(f) Writing the final draft.

All readers of the research report, must necessarily be conveyed enough about the study so that one can place it in its general scientific context, judge the adequacy of its methods and thus form an opinion of how seriously the findings are to be taken. For this purpose there is the need of proper layout of the report. The layout of the report means as to what the research report should contain. A comprehensive layout of the research report should comprise of the following three aspects:

(A) Preliminary pages;

(B) The main text; and

(C) The end matter.

The above three aspects are discussed below:

4.3.1 Preliminaries

In its preliminary pages the report should carry a title and date, followed by acknowledgements in the form of 'Preface' or 'Foreword'. Then there should be a table of contents followed by list of tables and illustrations so that the decision-maker or any researcher interested in reading the report can easily locate the required information in the report.

4.3.2 Main Text

The main text provides the complete outline of the research report along with all details. Title of the research study is repeated at the top of the first page of the main text and then follows the other details on pages numbered consecutively, beginning with the second page. Each main section of the report should begin on a new page. The main text of the report should have the following sections:

(i) Introduction;

(ii) Statement of findings and recommendations;

(iii) The results;

(iv) The implications drawn from the results; and

(v) The summary.

(i) Introduction : The purpose of introduction is to introduce the research project to the readers. It should contain a clear statement of the objectives of research i.e., enough

background should be given to make clear to the reader why the problem was considered worth investigating. A brief summary of other relevant research may also be stated so that the present study can be seen in that context. The hypotheses of study, if any, and the definitions of the major concepts employed in the study should be explicitly stated in the introduction of the report.

The methodology adopted in conducting the study must be fully explained. The scientific reader would like to know in details about the following :

- How was the study carried out?
- What was its basic design?
- If the study was an experimental one, then what were the experimental manipulations?
- If the data were collected by means of questionnaires or interviews, then exactly what questions were asked (The questionnaire or interview schedule is usually given in an appendix)?
- If measurements were based on observation, then what instructions were given to the observers?
- Regarding the sample data used in the study the reader should be told:
  - (i) How many sample data were there?
  - (ii) How wasthe sample data selected?

All these questions are crucial for estimating the probable limitations of the data. The statistical analysis adopted must also be clearly stated. In addition, the scope of the study should be stated and the boundary lines be demarcated. The various limitations, under which the research project was completed, must also be narrated.

(ii) Statement of findings and recommendations: After introduction, the research report must contain a statement of findings and recommendations in non-technical language so that it can be easily understood by all concerned. If the findings happen to be extensive, at this point they should be put in the summarised form.

(iii) Results:A detailed presentation of the findings of the study, with supporting data in the form of tables and charts together with a validation of results, is the next step in writing the main text of the report. This generally comprises the main body of the report, extending over several chapters. The result section of the report should contain statistical summaries and reductions of the data rather than the raw data. All the results should be

presented in logical sequence and split into readily identifiable sections. All relevant results must find a place in the report. But how one is to decide about what is relevant is the basic question. Quite often guidance comes primarily from the research problem and from the hypotheses, if any, with which the study was concerned. But ultimately the researcher must rely on his own judgement in deciding the outline of his report. Nevertheless, it is still necessary that he states clearly the problem with which he was concerned, the procedure by which he worked on the problem, the conclusions at which he arrived, and the bases for his conclusions.

(iv) Implications of the results: At the end of the main text, the researcher should again put down the results of his research clearly and precisely. He should, state the implications that flow from the results of the study, for the general reader is interested in the implications for understanding the human behaviour. Such implications may have three aspects as stated below:

(a) A statement of the inferences drawn from the present study which may be expected to apply in similar circumstances.

(b) The conditions of the present study which may limit the extent of legitimate generalizations of the inferences drawn from the study.

(c) The relevant questions that still remain unanswered or new questions raised by the study along with suggestions for the kind of research that would provide answers for them.

It is considered a good practice to finish the report with a short conclusion which summarizes and recapitulates the main points of the study. The conclusion drawn from the study should be clearly related to the hypotheses that were stated in the introductory section. At the same time, a forecast of the probable future of the subject and an indication of the kind of research which needs to be done in that particular field is useful and desirable.

(v) Summary: It has become customary to conclude the research report with a very brief summary, resting in brief the research problem, the methodology, the major findings and the major conclusions drawn from the research results.

4.3.3 End Matter

At the end of the report, appendices should be enlisted in respect of all technical data such as questionnaires, sample information, mathematical derivations and the like ones.

Bibliography of sources consulted should also be given. Index (an alphabetical listing of names, places and topics along with the numbers of the pages in a book or report on which they are mentioned or discussed) should invariably be given at the end of the report. The value of index lies in the fact that it works as a guide to the reader for the contents in the report.

4.3.4 Other Important criterion of the Report

The following criterions are also important part of the research report:

4.3.4.1. Citation

A "citation" is the way of providing information to the readers that certain material in the research work came from another source. It also gives the readers necessary information to find that source again, the citation includes:

- information about the author
- the title of the work
- the name and location of the publisher that published the copy of the source
- the date of publication of the research work.
- the page numbers of the subject matter.

Giving credit to the original author by citing sources is the only way to use other people's work without plagiarizing. But there are a number of other reasons to cite sources:

- Citations are extremely helpful to anyone who wants to find out more about researcher's idea and the source of such idea.
- Not all sources are good or right –researcher's own idea may often be more accurate or interesting than those of theother sources. Proper citation will help the researcher to conceive the best idea from some others' citation.
- Citing sources shows the amount of research anyone has done.
- Citing sources strengthens the research work by lending outside support to the researcher's idea.

Proper citation is essential under the following circumstances:

- Whenever quotes are used.

- Whenever paraphrase are used.
- Whenever an idea that someone else has already expressed.
- Whenever specific reference to the work of another are used.
- Whenever someone else's work has been critical in developing any ideas.

4.3.4.2. Notes

Research notes are small, often preliminary studies, descriptions of unexpected and perhaps unexplained observations or lab protocols that can be described in a short report with a few illustrations (figures/tables), or even a single figure.

Research Notes can be as short as a single-figure paper. In such cases, all that is required is a short introduction describing the question or hypothesis that led to the presented figure, followed by a description of the methods used.

4.3.4.3. References

References can be listed in any standard referencing style as long as it is consistent between references within a given article. However, basic requirements include:

- Journal abbreviations should follow the Index Medicus/MEDLINE abbreviation approach. Only articles, books and book chapters, datasets and abstracts that have been published or are in press, or are available through public e-print/ preprint servers/data repositories, may be cited. Unpublished abstracts, papers that have been submitted but not yet accepted, and personal communications should instead be included in the text; they should be referred to as 'personal communications' or 'unpublished reports' and the researchers involved should be named. Authors are responsible for getting permission to quote any personal communications from the cited individuals.
- Web links, URLs, and links to the authors' own websites should be included as hyperlinks within the main body of the article, and not as references.
- References to trials on a clinical trial database should be as follows: [Authors/ name of group], [title of the trial], In: ClinicalTrials.gov [cited year month date], Available from [URL of the link from ClinicalTrials.gov]. Example: Kovacs Foundation, The Effect of Ozone Therapy for Lumbar Herniated Disc. In: ClinicalTrials.gov [cited 2012 Aug 30], Available from: http://clinicaltrials.gov/ct2/show/NCT00566007

- Datasets published or deposited elsewhere (for example, in figshare, Dryad, etc.) should be listed in the 'References' section and the citation to the dataset should follow one of these examples.

4.3.4.4. Bibliography

Bibliography is a list of documents, books, periodicals, manuscripts, etc - which have some useful information on the given subject. A bibliography should include only those works which the participant has consulted for the study. This should be written in alphabetical order. Bibliography is to be written sequentially as Books, Journals/ Magazines, Articles and Internet.

4.3.4.5. Abstract

This is the 'shop window' for the research work. It is the first (and sometimes the only) section to be read and should be the last to be written. It should enable the reader to make an informed decision about whether they want to read the whole report. The length will depend on the extent of the work reported but it is usually a paragraph or two and always less than a page.

A good way to write an abstract is to think of it as a series of brief answers to questions. These would probably include:

- What is the purpose of the work?
- What methods did researcher use for his research work?
- What were the main findings and conclusions of the research?
- Did the research work lead to make any recommendations for future actions?
- What is the purpose of the work?
- What methods did you use for your research?
- What were the main findings and conclusions of the research work ?

Therefore, abstracts provide a basic summary of the article. Although the abstract should explain why the article might be interesting, the importance of the work should not be over-emphasized. Abstracts formatted with bullet point lists and separate headings are allowed, but the text will be included in the overall word limit. Citations should not be used in the abstract. Abbreviations, if needed, should be spelled out.

4.3.4.6. Key Words

Authors should supply relevant keywords that describe the subject of their article. These will improve the visibility of the article.

4.4 ❒ Conclusion

The conclusions should be a short section with no new arguments or evidence. The researcher needs to sum up the research work in the conclusion part.

This section may also include:

- Recommendations for action
- Suggestions for further research

In spite of all that has been stated above, one should always keep in mind the fact that,report-writing is an art which is learnt by practice and experience, rather than by mere doctrination.

4.5 ❒ Summary

(i) Writing a research report is an artistic job, wherein the researcher has to present what he/she has done, so that the research is understood well by the stakeholders and becomes useable. (ii) There are many rules which are generally follows to write a concise and analytical research report. (iii) Initially the preliminaries contain the introductory part then main texts discuss the different analysis and then come the conclusion. (iv) Besides these other important aspects are citation, notes, references and bibliography.

4.6 ❒ Glossary/Keywords

Abstract: Summarizes the report including the hypotheses, procedures, and major findings.

Acknowledgment: Include only if special help was received from an individual or group.

Appendix: Any tables, figures, forms, or other materials that are not totally central to the analysis but that need to be included are placed in the Appendix.

End Notes: These are like footnotes but are located at the back rather than the bottom of each page. These would include all of the references for all works cited in the Review of Related Literature or any other sections of the report as well as the references for quotations, direct or indirect, taken from other sources, or any footnote comments that might have been included. These are listed in numeric order as presented in the text.

Glossary: The glossary offers explanations of technical terms used. The terms are arranged in an alphabetical order. This saves the reader from the trouble of looking into the dictionary for the meaning of those words.

4.7 ❒ References and Further Readings

Cooper, D.R. and Schindler, P.S. (2013) Business Research Methods, Tata McGraw-Hill, New Delhi

Moser, C.A. and Kalton, G. (1971) Survey, Methods in Social Investigation, Heinemann Educational Books Ltd.

Kothari, C.R., (2004). 'Research Methodology: Methods and Techniques', New Age Publication

Walliman, N. (2016)*Social Research Methods,* New Delhi: Sage Publications.

4.8 ❒ Model Questions

1. What is research report?
2. Explain the significance of a research report and narrate the various steps involved in writing such a report.
3. What do you understand by the preliminary of a report?
4. Describe, in brief, the layout of a research report, covering all relevant points.
5. What are the different sections of the main text? Discuss them elaborately.
6. What points should be discussed in the introduction part of a research report?
7. What points should be kept in mind while preparing a research report?
8. Mention the cases where citation is required.

9. Write short notes on the following:

(a) The techniques of report writing;

(b) Characteristics of a good research report;

(c) Bibliography and its importance in context of research report;

(d) Rewriting and polishing of report.

(e) Citation

(f) Research Notes

(g) Reference

(h) Abstract