Indira Gandhi National Open University
School of Continuing Education

MRDP-205
Research and Project Work

METHODS IN SOCIAL RESEARCH

UNIT 1
Designing a Research Study **5**

UNIT 2
Tools of Data Collection **27**

UNIT 3
Analysis of Data and their Presentation **43**

UNIT 4
Statistics on Rural Development **61**

Prof. V.C. Kulandai Swamy
Vice-Chancellor
IGNOU, New Delhi

Prof. H.P. Dikshit
Pro-Vice-Chancellor
IGNOU, New Delhi

EXPERT COMMITTEE

Prof. Kanta Ahuja
Department of Economics
University of Rajasthan, Jaipur

Dr. K.R. Ranganathan
Director, Institute of Public Admn.Member, Central Board for the Prevention & Control of Water Pollution, New Delhi

Prof. A.W. Khan
Director
Communication Division
IGNOU, New Delhi

Prof. Iqbal Narain
member Secretary
ICSSR, New Delhi

Dr. D.P. Singh
University of Lucknow
Lucknow

Prof. C.S. Murthy
School of Engineering & Technology
IGNOU, New Delhi

Shri R.N. Kaul
Consultant (Forestry)
New Delhi

Dr. N.K. Bansal
Centre of Energy Studies
IIT, New Delhi

Prof. A.P. Barnabas
Consultant
IGNOU, New Delhi

Smt. Sarala Gopalan
Joint Secretary (Training)
Department of Rural Development

Dr. Samir Chowdhry
Director, Child in Need Institute
West Bengal

Prof. Suresh Goel
Professor Life Sciences
School of Sciences, IGNOU

Shri R.P. Kapoor
Director Genral
NIRD, Hyderabad

Prof. C. Gopinath
Indian Institute of Management
Ahmedabad

Prof. B.N. Koul
Director, Division of Distance Education, IGNOU

Dr. P.B. Mathur
Assistant Director General, ICAR, New Delhi

Prof. A.B. Bose, Advisor,
Planning Commission, New Delhi

COURSE PREPARATION TEAM

Dr. Gurchain Singh
School of Continuing Education,
IGNOU, New Delhi

Dr. N.J. Kurien
Director (E & S), Dept. of Rural Development, Krishi Bhawan,
New Delhi

Mr. Saugatio Sen
School of Continuing Education,
IGNOU, New Delhi

COURSE EDITOR

Prof. A.B. Bose
Adviser, Planning Commission
Yojna Bhawan, New Delhi

COURSE COORDINATOR AND BLOCK EDITOR

Dr. Gurchain Singh
School of Continuing Education
IGNOU, New Delhi

Prof. Nageshwar Rao
Vice-Chancellor, IGNOU, New Delhi

COURSE PREPARATION TEAM (Revision)

Prof. R.P. Singh
Director, SOCE, IGNOU, New Delhi

Prof. Gurchain Singh and
Ms. Santosh Tanwar
SOCE, IGNOU, New Delhi

COURSE AND PROGRAMME COORDINATOR (Revision)

Dr. Balkar Singh, Assistant Professor
Rural Development, SOCE, IGNOU, New Delhi

PRINT PRODUCTION

Mr. Rajiv Girdhar
Assistant Registrar, MPDD, IGNOU, New Delhi

Mr. Hemant Parida
Section Officer, MPDD, IGNOU, New Delhi

August, 2023

© Indira Gandhi National Open University 2023

ISBN : 81-7263-0964-4

All rights reserved. No part of this work may be reproduced in any form, by mimeograph or any other means, without permission in writing from the Indira Gandhi National Open University, New Delhi

Further information on the Indira Gandhi National Open University Courses may be obtained from the University's Office at Maidan Garhi, New Delhi - 110068

Printed and Published on behalf of Indira Gandhi National Open University, New Delhi

BLOCK INTRODUCTION

The course MRDP-205 has only one block. The rest of this course will be project work. The reading materials of this block pertain to Methods in Social Research. We have made an effort to acquaint you with the fundamental issues related to research design, tools of data collection and analysis of data and their presentation. We have also furnished vital statistical information on rural development. The main objective of this block is to impart such knowledge to you as may facilitate your research work in the discipline of rural development.

Unit 1: Designing of Research Study – will introduce you to nature of scientific knowledge, characteristics of scientific method and the meaning of social research. You will learn about the basic elements of research design and main characteristics of different types of research designs employed in social research. The unit will also make you aware of the problems encountered in the preparation of a research design.

Unit 2: Tools of data collection – will acquaint you with the nature of main tools employed to conduct social research. You will learn that research strategies and the attendant need for data determine the selection of particular tools. We have discussed the salient features of such tools as observation, interview, schedule, mailed questionnaire and so on. You will acquire knowledge about the strengths and weakness of different tools. You will also become aware that a wide range of factors influence choice and selection of tools for the collection of data.

Unit 3: Analysis of Data and their Presentation – will give you useful insights about the basic procedures involved in the analysis of data. You will become aware of the important link that exists between empirical evidence (data) and the attendant theoretical formulations. You will learn about various factors which help in the processing, classification and interpretation of data. You will gain some knowledge of statistical analysis, tabulation and presentation of data. The unit will also give you a preliminary idea for the preparation of a research report.

Unit 4: Statistics on Rural Development – will furnish you with some statistical information on rural development. It will enable you to procure the requisite statistics from the relevant sources of information. You will learn about the nature and importance of statistical requirements for adequate planning and efficient implementation of research projects in the field of rural development.

The reading materials of this block will equip you with suitable knowledge for pursuing your research interest in rural development.

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 1 DESIGNING A RESEARCH STUDY

Contents

1.0 Aims and Objectives
1.1 Introduction
1.2 Scientific Knowledge
1.3 Nature of Scientific Method
1.4 Social Research
1.5 Research-Design
1.6 Elements of Research-Design
1.7 Types of Research Design
1.8 Sampling
1.9 Let Us Sum Up
1.10 Key Words
1.11 Suggested Readings
1.12 Model Answers

1.0 AIMS AND OBJECTIVES

The main aim of this unit is to provide you with preliminary understanding which will be helpful to you in the designing of a research study. However, as a pre-requisite for developing such an understanding, the earlier part of this unit is devoted to acquaint you with such aspects as the nature of scientific knowledge, characteristics of scientific method and meaning of research in social sciences.

After having gone through this unit you should be able to:

- describe the nature of scientific knowledge, the important characteristics of scientific method and the meaning of social research;
- indicate some of the important elements of research design;
- state the characteristics of different types of research;
- identify the main issues and problems in preparing a research design.

1.1 INTRODUCTION

An important dimension of research in social sciences is the preparation of research design. Now, to design is to plan. In other words, designing is the process of making decisions before an actual situation arises in which such decisions have to be executed. Adequate fore-planning makes the researcher aware of the difficulties which may crop up during an envisaged research investigation. It also helps to avoid wasteful expenditure of money, time and energy, apart from protecting the researcher from possible pitfalls sufficiently in advance.

In this unit we will acquaint you with some of the fundamental issues which are crucial in the designing of a research study. At the very outset, the unit deals with the nature of scientific knowledge, main characteristics of scientific method and the meaning of research in social sciences. Later, it discusses the important elements of a research design. The last section gives a general view of the main types of research studies. The next two units will acquaint you with various aspects involved in the collection and analysis of data.

1.2 SCIENTIFIC KNOWLEDGE

The pursuit of knowledge is a fundamental condition of human existence. Knowledge is an important medium for social and cultural progress of human beings. A sound understanding of our natural as well as social environment has contributed to the

overall development of societies. Thus the acquisition of knowledge in various forms remains the most cherished goal of mankind. However, prior to the rise of modern science, the contents of knowledge mostly contained mystical elements. The explanations of natural or social events were usually unclear in the past. But over a period, the growth of scientific approach to knowledge has made it possible to develop a better and sound understanding of our environment.

Scientific approach to knowledge is grounded on a set of fundamental assumptions: the world exists; we can know the world with our senses; there is a definite order inherent in the universe; and the occurrence of events has a cause and effect relationship. Resting on such assumptions, scientific approach helps to discover **order** from the apparent confusion of events. Scientific approach relies on **observation** of concrete **facts** so that true causes of a phenomenon are clearly comprehended. Proceeding with the observation of real and particular events or situations it ultimately aims to generate general laws, or principles. Such general laws or principles must have sufficient potential for explaining similar events or phenomena in time and space.

Now the accumulation of a systematic body of knowledge through the application of scientific approach is called scientific knowledge. In other words, scientific knowledge is that knowledge which enhances our understanding of the entire **empirical** world, i.e., the world which is susceptible to human experience. The hallmark of scientific knowledge is that its validity rests on direct or indirect observation of reality. Thus the knowledge which has supernatural, spiritual or magical characteristics is not scientific knowledge.

Scientific knowledge can be grouped under two main divisions: Natural Sciences and Social Sciences. Natural Sciences deal with natural environment in which human beings live. For example disciplines such as Physics and Chemistry etc. which deal with laws of matter, motion, energy or chemical properties of matter belong to natural sciences. Included in this division are also such disciplines as Zoology, Botany and so on which deal with living beings. The second major division of scientific knowledge is constituted by Social Sciences. Social Sciences are broadly concerned with various aspects of group life of human beings. The scientific study of organized human groups is a relatively recent development. However, a vast amount of information has already been gathered about the social, historical, economic and political life of human beings. The accumulation of such scientific information has been used to build a system of knowledge which deals with the nature, growth and functioning of human societies. This system of knowledge constitutes the field of social sciences. For example, disciplines such as anthropology, sociology, economics and political science belong to the field of social sciences.

1.3 NATURE OF SCIENTIFIC METHOD

Now the subject matter of any scientific investigation can be anything ranging from natural, biological, social to psychological phenomena. However, basic to any such investigation is the scientific method so that knowledge thus generated attains the scientific character. In other words, scientific knowledge is united by the scientific method and cannot be characterized by any particular or general subject matter. It is erroneous to go by commonsense popular notions that scientific knowledge pertains only to the field of natural sciences. It must be clear to you at this stage that knowledge in social sciences when produced through the application of scientific method is no less scientific in character than the natural sciences. However, a word of caution is necessary. The subject matter of social sciences is constituted by living and thinking human beings. Human beings cannot be subjected to the same treatment (e.g. experimentation under laboratory conditions) as is the case of natural sciences. Despite such a limitation there is nothing supernatural, artificial or unreal about the social phenomenon. Social facts are as real as facts in natural sciences. Hence, there is no difficulty to subject them (social facts) to scientific treatment. Notwithstanding some limitations, scientific method is the best known approach to generate scientific knowledge in social sciences.

Scientific method can be distinguished by adherence to several principles. These principles are: Objectivity, Empiricism, Relativism, Ethical Neutrality, Skepticism and Simplicity. In the following, we will briefly explain to you these salient principles of scientific method.

i) Objectivity: By objectivity is meant that the scientific investigation of social phenomenon must not be influenced by the subjective biases of the investigator. Social analysis must be over and above the personal feelings or attitudes of the social scientist. In other words, the process of acquiring knowledge in social sciences must remain independent of any considerations of race, colour, creed, nationality, religion, moral or political preferences. However, an ideal form of objectivity is difficult to attain in social sciences since there are always chances that personal or cultural preferences may influence a given social analysis. Thus true objectivity is hard to achieve in social sciences because the role of subjective factors cannot be totally eliminated. But efforts must be made to minimize subjective influences so that a near true state of objectivity is attained in social sciences.

ii) Empiricism: The second important characteristic of scientific method is empiricism. It means that a social investigation must be conducted empirically. In other words, our views about some or the other aspect of society must be based on clear and definite factual evidence. Such evidence must be produced by **observing** the relevant social reality with the help of human senses, viz., sight, hearing, taste, smell and touch. Nothing is left to speculation. Rather, the understanding of various social phenomena is developed through human experience. Now the observation of social reality can be made directly with the help of human senses or indirectly with the support of some instruments so as to aid and extend the ability to observe.

iii) Relativism: Relativism is that principle of scientific method whereby the results of a scientific investigation are never considered as permanent or absolute truths. Scientific method presupposes that propositions which are true in some situation ma become open to question (or come under criticism) on the face of new evidence. Further advances in knowledge may refute or reject the established notions or views of our social environment. By relativism in scientific method is meant that no notions are sacred, no propositions are privileged and no truths are absolute to the social scientist. Our views about social reality are tentative and have only a relative credibility *vis a vis* our past understanding of the social world. Our understanding of a social phenomena is subject to change on the face of fresh and more sound evidence. Scientifically speaking, a proposition or argument is valid or true for a period until it is refuted in the future.

iv) Ethical Neutrality: Scientific method demands that a social scientist maintains an ethically neutral attitude in his pursuit of knowledge. In his professional capacity he is not supposed to take sides on issues of moral or ethical nature. He must remain over and above all religious, political or moral issues. Even though as a citizen he may have some personal preferences, as a social scientist he must abandon them. His sole aim must be to reveal the true nature of various social phenomena so as to facilitate the systematic accumulation of scientific knowledge of societies. Whether or not such knowledge is used to make sensible reasonable decisions is another matter. Scientific method reserves silence on normative questions.

v) Skepticism: Scientific method presupposes that a social scientist must have the willingness or capacity to doubt the validity of existing social theories. He must have a critical bent of mind. He must not accept the truth of a given proposition or theory simply because of its popularity or general acceptance. He must be skeptical about the authenticity of each and every explanation, particularly when they are supported by insufficient evidence. All doubts must be cleared with the help of observable evidence. Doubt is an integral part of scientific method that helps in the growth of scientific knowledge.

vi) Simplicity: When an adequate explanation of a social phenomenon is available, to generate another similar one is superfluous. Duplication must be avoided. Out of a given range of social theories, the one which is scientifically superior must be accepted while others must be rejected. Duplicate or repetitive explanations are simply untenable and a waste of time.

1.4 SOCIAL RESEARCH

The field of social sciences is of crucial importance to human beings. Increasing our knowledge about the functioning of human societies is as important as learning about physics, chemistry, biology, or engineering. Developing a sound understanding of our social environment is necessary so that human beings can live happy, meaningful and satisfying lives. This presupposes that we must study different aspects of society. If we study only some aspect and neglect others we will build a distorted picture of society.

We have noted above that social sciences are divided into a number of disciplines. Each discipline is further subdivided into specialized areas so as to facilitate more intensive and deeper understanding. For example, social stratification is one such special area of study within the discipline of sociology. It is devoted to study the way societies are stratified into castes, classes, occupations, professions, and status groups. This kind of specialized knowledge in a given area of a discipline is generated by means of **research.** The knowledge thus produced is systematically related and interpreted in a meaningful way within the broader parameters of the concerned discipline. Furthermore, the knowledge acquired by various disciplines is used to build interdisciplinary perspectives in social sciences so as to develop a more balanced understanding of society.

Now, as a beginner you must be curious to know what research is all about. Broadly speaking, research can be viewed as an overall scheme of scientific activities in which scholars engage themselves in order to produce knowledge or to extend/modify/verify existing claims of knowledge. As in natural sciences, research in social sciences is basically a scientific exercise. Two most important characteristics of social research are that:

- it must be **social** and
- it must be **empirical** in nature.

By social is meant that the phenomenon studied must be somehow related to the group life of individuals. For example, research can be undertaken to study problems relating to the family, community, or organizational, and institutional aspects of society. By empirical is meant that the investigation must be based on direct or indirect observation of social reality. In other words, the findings of research must be supported by sufficient evidence.

Now social research operates at two distinct but interrelated levels — conceptual-theoretical and observational-empirical. At the conceptual-theoretical level research is concerned with abstract statements or generalizations of social reality. These are in the form of concepts, hypotheses, rules, principles, laws, and theories. On the other hand, at the observational-empirical level research is concerned with facts, evidence, and data. In research there is a constant interplay between these two levels. Each level has a bearing on the other. For example, the right type of theoretical constructs or hypotheses facilitate the process of procuring relevant data. Similarly, sound empirical data may help to generate adequate theoretical formulations of social reality.

Social research begins with the selection and formulation of the research problem and ends by drawing some conclusions about the research finding. In between there are a number of other stages involved in research: preparation of research design, data-collection, data processing and data analysis. All stages are important in conducting social research. However, the present unit is mainly devoted to familiarize you with different aspects of designing of a research study.

Check Your Progress I

Notes: a) Use the space given below for your answer

b) Check your answer with the model answer given at the end of this unit.

1) The following are some statements based on text you have already read. State whether these statements are true or false by putting ✓ in the relevant box.

a) Scientific approach to knowledge relies on observation and helps to enhance our understanding of the entire empirical world.

True ☐ False ☐

b) Scientific knowledge has supernatural and spiritual characteristics.
True ☐ False ☐

c) Scientific method is used to study natural sciences as well as social sciences.
True ☐ False ☐

d) Social facts are unreal and artificial and cannot be studied by scientific method.
True ☐ False ☐

e) Scientific method relies on objectivity
True ☐ False ☐

f) By objectivity in social science is meant that social analysis must not be influenced by any subjective factors.
True ☐ False ☐

g) Empiricism is that characteristic of scientific method whereby a social investigation is based on the observation of social facts.
True ☐ False ☐

h) Element of doubt is an integral part of scientific method.
True ☐ False ☐

i) Scientific research in social sciences is both social and empirical in nature.
True ☐ False ☐

j) Social research helps to generate extend, correct and verify knowledge in social sciences.
True ☐ False ☐

2) List the main principles of Scientific Method.

.....

.....

.....

.....

.....

3) What is the aim of research?

.....

.....

.....

.....

.....

1.5 RESEARCH DESIGN

All research begins with a question leading to selection of a research problem. Problems initially chosen generally require more precise formulation so as to make them worthy of research. Starting from a general idea one must decide specifically what one wants to know; for what purpose one wants to know; how to carry out a particular social investigation and such other questions. Now, from the very beginning a researcher is supposed to make several crucial decisions such as, for instance:

- What entities (i.e. individuals , groups, organization or institutions) are to be studied;
- What should be an adequate theoretical framework of the study;
- What kinds of hypotheses are relevant and useful;
- What types of data are relevant;
- How and from where to collect the required data;
- What kinds of relationships among different sets of data are to be anticipated.

All such matters relating to research must be decided in advance.

The meaning of 'design' in research will be clear to you with the help of an analogy. Imagine, for a moment, how an architect designs a building, say a house. To begin with, the architect conceives the layout of the building. He decides how large it will be, how many rooms it will have, what materials will be used and so on. He considers each decision that will have to be made during the construction of the house. He does all this **before** the actual construction begins. Thus he prepares a design so as to have a clear picture of the house. The preparation of such a design helps to eliminate the difficulties which may be encountered at the time of construction. It also avoids wastage of time, money and material. Further, such an exercise enables the construction plan to be executed successfully.

The above analogy is equally relevant to the field of research. Before conducting a research investigation, it is necessary for a researcher to make a number of decisions concerning aspects such as:

- Selection of the topic and formulation of the research problem (including review of literature and setting out the objectives of the study);
- Data to be collected;
- Tools of data collection;
- Tabulation and Analysis;
- Time frame of study;
- Cost estimates;
- Writing the Report.

These decisions are articulated in such a manner that the research purpose is achieved with minimum expenditure of money, time and energy. Thus, broadly speaking, the research design is a symbolic representation of such decisions as help in the creation of conditions for the successful completion of a research investigation combining relevance and purpose of research with economy in procedure. In the following we will acquaint you with such aspects as facilitate the process of decision-making at various levels of research and are the concerns of research design.

1.6 ELEMENTS OF RESEARCH DESIGN

We have already stated above that research begins with a specific problem or topic. So let us take a brief look at how problems are selected for study.

1.6.1 Selection of Topic for Research

We have discussed earlier that social research is both social and empirical. Thus, anything that is social and empirical is a relevant topic for scientific social research However, these are too broad criteria which suggest a wide range of potential top for social research. These do not indicate how specific topics are likely to emerge. Hence, in the following, we will acquaint you with those factors which affect decisions on topics selected in social sciences.

i) The Structure and State of a Discipline

The existing state of theory and research in a particular discipline generally influences the process of selection of topic. The broad parameters of a discipline also determine the framework of a topic for research. For example, sociologists usually study various aspects of such institutions as religion, politics, education and family, to mention some. It is so because the discipline has emerged, developed and got organized in such a way as to facilitate the study of **social relations** within organizations, institutions, communities and so on. Now, as knowledge in a particular area of a discipline develops, consistencies and gaps in information are revealed or become visible. A number of unresolved problems or issues also arise. The researcher selects such topics as help to fill the gaps in information or attempts to resolve certain issues or problems through research.

The knowledge of the existing state of research in a discipline helps to refine the problem and to make it precise and worthy of research. Too broad a problem or too

ambitious a research project may consume a lot of money, time and energy and yield limited results. On the other hand, a narrowed down, precisely formulated and sharply defined problem may be successfully researched within a time bound schedule. In this context, a thorough review of previous research works on the topic of interest is considered absolutely essential. Besides helping to narrow and refine the research problem, review of literature may reveal the broader significance of the problem. Such an exercise will also help to suggest relevant methods and procedures by indicating how other researchers have addressed the problem. But only experienced researchers are able to tap this source of ideas since it requires a thorough knowledge of literature in the concerned discipline.

ii) Social Problems

Many people are attracted to the social sciences because of their perceived relevance to social problems, and this is the most potential source of research ideas or topics for the beginning researcher. The focus and development of social sciences are intimately related to interest in basic problems of "human condition". Throughout the growth of social science disciplines this has remained a major source for topics of research. For example, in the Indian context considerable interest has been displayed in research problems relating to the situation of the disadvantaged groups (the scheduled castes, scheduled tribes, landless labour, women etc.), migration, social inequalities and exploitation, gender inequalities, rural power structures, violence, social tensions etc. While selecting a topic, the researcher is generally motivated by the idea of arriving at a better understanding of the nature and dimensions of the problems.

iii) Other Determinants

A number of other factors also influence the selection of a research topic. The availability of grants for particular themes, the prestige and popularity of a research area of the subject, public interest and the motivations of the researcher (to name some) place a premium on different research topics at different times. Many a times these factors reinforce one another. For example, since independence, particular emphasis has been laid to study problems relating to landless and rural poor in India. Recently, women's studies, bonded labour and child labour have emerged as a priority area of research because of contemporary concern about these problems. Some research studies are sponsored by agencies interested in a particular area. In recent years, government departments have been funding a number of research studies through grants to institutions for information about a problem area or for evaluating programmes under implementation.

iv) Practical Considerations

Research requires time, money and personnel. Constraints in these resources also influence decisions concerning the research. For example, a research project requiring long duration cannot be pursued with small funds. Again, to survey rural areas so as to determine the state of malnutrition may require a large number of trained personnel. Furthermore, skill and knowledge of the researcher and availability of data also affect decisions regarding the formulation and selection of a research problem. Sometimes, a topic is chosen for its theoretical or practical relevance. However, the usual practice is to justify social research on both practical and theoretical grounds.

Now, once a general topic has been chosen, it needs to be stated in researchable terms. This means that the topic must be translated into one or more clearly defined specific areas of inquiry. However, formulating a researchable problem involves several decisions. We look at some of these now.

1.6.2 Selection of Units of Analysis

In social research, the objects or events (entities) under study are referred to as units of analysis. Let us say, a researcher wants to determine, the nature and extent of poverty in rural areas. In such a case a village can be treated as the relevant unit. On the other hand, if a researcher wants to know the nature of social relationships among various castes or communities living in the village, then 'caste' or community will constitute the relevant unit of analysis. Further, if one wants to know what sort of people are involved in active politics in rural areas, then one would obviously need to gather information on the characteristics of individual persons. The unit of analysis in such a case will be an individual. Generally, the purpose of the study dictates what is the appropriate unit of analysis.

1.6.3 Choice of Variable

A social scientist observes units of analysis in the process of research. But he is primarily interested to study the relationships among some characteristics or properties of the observed units. Now, characteristics of units that **vary,** taking on different values or attributes for **different observations,** are called **variables.** Variables may vary over cases, over time or over both cases and time. For example, i you wish to observe individual people then any set of characteristics that may differ for different people, such as age (range of years), gender (male and female), and marital status (single, married, divorced, widowed, separated) is a variable.

Many a time there is confusion between variables and the attributes or categories of the variable. For example, gender is a variable consisting of two categories (male/female) because they are simply categories into which a person's gender will fall. Again, 'divorced' or 'married' are not variables but attributes of the variable marital status. To keep this distinction clear, note that any term you would use to describe a person is an attribute or category of a variable.

Types of variables

It is necessary for a researcher to decide which variables should be the focus of research. Variables which are under focus are called *explanatory* variables. All other variables which are not the direct focus of research are called *extraneous* variables.

Explanatory variables are further classified as: *dependent* and *independent.* The dependent variable is the one the researcher is interested in explaining and predicting. Variation in the dependent variable is thought to depend on or to be influenced by certain other variables. If you think in terms of cause and effect, the independent variable is the presumed cause and the dependent variable the presumed effect. For example, if you wish to study the relationship between educational qualifications and income, then educational attainment is an independent variable and income is the dependent variable. Keeping all other things as equal, information about a person's qualifications can help to roughly predict his level of income. Now a variable is not intrinsically independent or dependent. An independent variable in one study may well be a dependent variable in another. Further, there is also a convention to symbolize the independent variable with the letter 'X' and dependent variable with letter 'Y'.

Extraneous variables are classified as *controlled* and *uncontrolled.* Controlled or more commonly control variables are held constant or prevented from varying during the course of observation or analysis. This helps to limit the focus of research. For example, this can be done by selecting only individuals of the same age and gender, observing groups of the same size. Controlling variables or holding them constant simplifies complex social situations. It is a means of ruling out variables which are not of immediate interest even though such variables might otherwise explain part of the phenomenon that the investigator wishes to understand. Indeed, one aim of an efficient research design is to identify potentially relevant extraneous variables in order to control as many as is feasible.

A distinction is also made between *quantitative* and *qualitative* variables. A variable is quantitative if its values or categories consist of numbers, i.e. if differences between its categories can be expressed numerically. For example, the variable 'income' measured in rupees represents a quantitative variable. Different people have different incomes which are quantified and measured in terms of rupees. Qualitative variables represent certain qualities, attributes or discrete categories. They are usually designated by words or concepts and signify non-numerical differenes between categories as, for example, male and female. One can make categorical (but not numerical) distinctions between people of different gender. In other words, a person can be classified either as male or female.

1.6.4 Relationships

The ultimate objective of a social scientist is to make sense out of the social world by identifying enduring relationships among phenomena. In fact, much research in socia sciences is directly aimed at developing and testing relationships. There may also be other more immediate purposes for which research is conducted, viz., gaining

preliminary insights into or familiarity wih a phenomenon; description of groups or communities; or exploration of an event or situation. But on the whole, research findings will depend to a large extent on what particular relationships are anticipated. The anticipated relationships among variables can only be revealed by systematically observing a phenomenon. Research investigations, as you must have realised, do not make random observations. Thus, keeping in view the goals of research, a researcher must decide what to observe or ignore, how to go about making observations, and how to interpret them. Such decisions are based on the researcher's expectations about how variables are related to one another; they constitute an important feature of a good research design. Since all research carries expectations about the nature of what is being investigated, it is important that anticipated relationships and the guiding theoretical premises are identified as far as possible.

Now the relationships with which social scientists are concerned are the relationships among variables. Two or more variables are related, or form a relationship, to the extent that changes in one variable are accompanied by predictable changes in the others. When we say variable X and variable Y are related we mean that there is something common to both variables. For example, if we say that education and income are related, we mean that the two "go together", i.e., that they co-vary. If individuals with higher education have higher incomes then covariation is what education and income have in common. Establishing a relationship in empirical research consists of determining which values of one variable co-vary with values of one or more other variables.

Check Your Progress II

Notes : Use the space provided below for your answer.

1) Describe very briefly the factors which affect the selection of a topic for research. (Hint: See the Text)

.....

.....

.....

.....

.....

.....

2) Explain very briefly, the meaning of the following terms : (a) Independent variable (b) quantitative and qualitative variable. (Hint: See the Text.)

.....

.....

.....

.....

.....

.....

1.6.5 The Nature of Causal Relationship

For purposes of explanation and prediction, scientific research finds it helpful to think in terms of cause and effect relationship. But how does one identify such a relationship? What is meant by the term causality? In lay terms, a cause is something that makes some other thing to happen or change. For example, a stone thrown at the window pane will cause the pane of glass to shatter. Here throwing of stone is the cause while the shattered glass is the effect. We all find such an implicit understanding of causality in everyday life. But the meaning of causality in scientific research rests on the assumption that one can observe a constant or stable association between the occurrence of events and from such an association we infer a causative connection.

Causal relationships are generally regarded as the heart of scientific understanding. Thus, it is important to understand the bases of causal inferences. Social scientists generally require at least three kinds of evidence to establish causality. These requisites are association, direction of influence and nonspuriousness.

Association: For one variable to be a cause of another, the variables must be statistically associated. If the pattern of changes in one variable is not related to changes in another, then the former is not the cause of the latter. Associations, of course, are almost never perfect. Perfect associations may be expected only under pure theoretical inference but not in the real world of observations. This is so because in real social situations it is not possible to eliminate completely the influence of extraneous factors. In actual practice, thus, causal relationships are determined in terms of strong or weak associations. For instance statistical measures only indicate the association of two variables. Cause and effect influence/deducation is based on the research design's formulation and theoretical premises. In social sciences causal relationships are sometimes implied from comparatively weak associations. One reason for this is that many measurements in social sciences are imprecise. The primary reason though is that in explaining human action, it is assumed that there can be multiple causes operating independently or jointly to produce the same or similar effects.

Direction of Influence: The second criterion needed to establish a causal relationship between events is that cause must precede its effect. At least the direction of influence should be from cause to effect. In other words, changes in the causal factor or independent variable, must influence changes in the effect, or dependent variable, but not *vice versa*. For example, the formation of clouds is always expected to precede the rains.

The requirement of determining direction of influence has two major implications for designing a research study — (a) hypothetical relationships should always specify the direction of influence among variables; and (b) whenever the direction cannot be established theoretically it should be tested empirically. However, this task is relatively easier in experiments but often difficult in other types of social research.

Nonspuriousness: When a correlation or association is produced by an extraneous third factor and neither of the variables involved in the correlation has influenced the other then the relationship is called a spurious relationships. The idea of nonspuriousness in research means that to infer a causal relationship from an observed correlation there should be good reason to believe that there are no hidden factors which could have created an accidental or spurious relationship. To infer nonspuriousness, the researcher ideally must show that the relationship between the variables are held constant. However, circumstances seldom allow researchers to control all variables. The greater the number of variables that are controlled without altering a relationship, the greater the likelihood that the relationship is not spurious.

1.6.6 Operationalization of Concepts and Definitions

Any discipline in social sciences involves the use of a special language. Such a language as it relates to research is basically constituted by concepts. Now a concept is an abstraction which represents an object, quality or property of something or a certain phenomenon. For example, social status, role, power, bureaucracy are concepts in political science and sociology.

Concepts serve a number of important functions:

- They enable us to relate to some aspect of reality. In other words they represent a way of looking at empirical phenomena.
- Concepts are the foundation of communication — the developments within a discipline can be better understood without elaborations and further explanation.
- Concepts are means for classification and generalisation. Researchers categorize, structure, order and generalize their experiences and observations in terms of concepts.
- Concepts are the building blocks of theories. Consequently they help in the process of explanation and prediction of social phenomena.

Concepts are the most critical elements in any theory because they define the shape and content of theories. For example, concepts of caste and class are crucial in the

shaping of theories of social stratification. Concepts when linked in a systematic way lead to the formation of theories. Hence, concept formation and theory formation are closely related.

Now clarity and precision in the usage of concepts are achieved by definitions. Definitions that describe concepts are conceptual definitions. A conceptual definition must denote the distinctive characteristics or qualities of the concerned phenomena. It should be stated in clear and unequivocal terms and must enhance communication and understanding among scholars belonging to a particular discipline.

It is often the case that empirical properties or events that are represented by concepts cannot be directly observed. The empirical existence of a concept is established with the help of operational definitions. Operational definitions specify the meaning of concepts and provide criteria for their empirical application. They tell us what to do and what to observe in order to bring the given phenomena within the range of a researcher's experience and observation. An operational definition is a set of procedures that describe the activities one should perform in order to establish empirically the existence or degree of existence of a concept. In short, the operational definitions bridge the conceptual theoretical level with the observational empirical level.

To make this clear, let us suppose that the concepts of poverty and malnutrition are involved in a research study where the relationship between the two is to be determined. It is necessary for the researcher to operationalize the concepts of both poverty and malnutrition. This can be made possible, say, by linking poverty with income and malnutrition with the intake of nutrients or calories. The range of family incomes below a certain level of income may be construed as representative of the degree of poverty prevalent among various families. Again, the range of calorie intake below the level of a certain specified number may speak of the nature and extent of malnutrition. This way it is possible to observe the nature of poverty and malnutrition and their attendant interrelationships.

1.6.7 Formulation of Hypotheses

We will acquaint you now with another significant aspect of research design which is crucial in the designing of a research study. This aspect relates to the formulation of hypotheses. Scientific investigation always starts with a problem or question that can be solved or answered empirically. This means that the statement of the research problem should be prepared in such a way so that it becomes possible to make observations of related facts. Also, it should suggest about the nature and type of observations necessary to solve the problem. In other words, the research questions must be stated in a precise manner so that they give clear indication of what is to be observed and what is not to be observed. This is so because collection of all types of facts will not yield results. To achieve this aim, the research questions need to be stated in the form of hypotheses. In fact, hypotheses serve as a guide to the kind of data that must be collected in order to answer the research questions. Also, they help in the efficient organisation of data.

At this stage, an important fact needs to be mentioned. All research do not necessarily begin or are carried out with the help of hypotheses. Sometimes, the major outcome of a research study may be merely to generate hypotheses.

Hypotheses are formulated in the form of tentative answers to research questions. Formally defined, a *hypothesis* is an expected but unconfirmed relationship between two or more variables. In fact, for most research this means that a problem should ask a question about such a relationship among variables. Thus, hypothesis should speculate about the nature of such relationship; and the variables involved should be observable or at least potentially observable. In simple terms, an adequate hypothesis involving two variables should indicate which variable is a cause for the other or how changes in one variable are related to the changes in the other. We may hypothesize that an increase in education will result in a decrease in prejudice. This statement implies two features about the relationship:

- which variable causes, explains or predicts the other (e.g. education predicts prejudice);

- how changes in one variable are related to changes in the other (e.g. as education increases, prejudice decreases).

Now hypotheses which specify the form of relationship among variables are said to be testable. In such a case, it is possible to determine whether the hypothesis is true or at least whether it is probably true or probably false.

A good usable hypothesis is one which satisfies the following criteria:

- A hypothesis should be related to a body of theory. It must provide the theoretical rationale.
- The concepts should be clearly defined; the statement of the hypothesis must be precise and specific. Hypotheses stated in general or vague terms do not lead anywhere.
- A hypothesis should be formulated in such a way so that their conceptual content can be easily translated to understand the observable reality. In other words, the concepts involved should be such that it is possible to generate operational definitions.
- A hypothesis should be empirically testable. It should be formulated in such a way so that it is possible to logically deduce certain inferences which in turn can be tested by observations in the field.

You would like to know something about the potential sources of hypotheses. Earlier studies or the research already done in a given discipline helps in the formulation of hypotheses. Hints for generating hypotheses can also come from the direct observation of social reality. Hypotheses can, however, even be structured on quesses or simply by intuition. Sometimes, the principal goal of a research study, as stated earlier, may be merely to generate hypotheses.

Check Your Progress III

Notes : a) Use the space provided below for your answer.
b) Compare your answer with the model answer given at the end of the Unit.

1) After carefully reading the preceding text complete the following statements:

a) Making advance decisions regarding the various concerns of research is the object of .....

b) A good research design helps to achieve the desired research purpose with minimum .....

c) The selection of a topic for research in social sciences is influenced by considerations such as (i)..... (ii)..... (iii).....

d) The usual practice is to justify social research on both ..... and ..... grounds.

e) Individuals, groups, communities, organisations, social events or social situations which the social scientist wants to study are called .....

f) Variables which are the direct focus of research are called ..... while all others are .....

g) Much research in social sciences is directly concerned with revealing ..... between variables.

h) It is generally regarded that the heart of scientific understanding is .....

i) Research questions formulated in the form of tentative answers are called .....

j) Control variable are held ..... during the course of investigation.

k) A variable is ..... if its values or categories consist of numbers.

2) Explain briefly the functions of concepts.
(Hint: See the text)

.....

.....

.....

3) What are the criteria which a hypothesis should satisfy?
(Hint: See the text)

.....

.....

.....

.....

Activity I

Now that you have read the section on hypothesis, try and formulate three hypotheses on different areas of rural life. Show these to your Academic Counsellor at the Study Centre.

1.7 TYPES OF RESEARCH DESIGN

Social scientists try to make sense about the social world and express it in terms of relationships among variables. Until now we have made efforts to make clear certain aspects which are essential for the preparation of a research design so that anticipated relationships between variables can be studied. It is hoped that knowledge of such aspects as the selection of topic, units of analysis, variables, operationalization of concepts, and formulation of hypotheses will enhance your capacity to conduct an empirical research study. In some empirical studies the main aim is to test hypotheses. But all research is not confined to the immediate purpose of testing hypothesis. In fact, research is conducted for various purposes and the corresponding research designs differ depending on the nature and purpose of a research study.

Broadly speaking, research studies can be categorised as follows based on the main purposes for which the research is being conducted:

a) Exploratory studies whose main aim is to explore a phenomenon, to develop an understanding of the phenomenon so that it helps to formulate a precise research problem for some future study.

b) To describe a particular group, community or situation through descriptive studies as precisely and completely as possible.

c) Diagnostic studies which determine the frequency with which some phenomenon occurs and the manner in which phenomenon is associated with some other factors which influence it.

d) Experimental studies which test a hypothesis so that a causal relationship between variables is revealed.

In the following we will acquaint you with the nature of each of these.

1.7.1 Exploratory Studies

Exploratory studies are undertaken when relatively little is known about the phenomenon to be researched. In the main, the purpose of an exploratory study is to collect facts which will later help to formulate a precise research problem or generate sound hypotheses for further research. The other functions of such a study are that the researcher gets acquainted with the phenomena which he wishes to study later on or simply to know about the conditions under which such future research is to be conducted. Thus an exploratory study may be aimed at clarifying concepts, establishing future research priorities and gathering information about the actual conditions affecting an intended research study.

When attempting to explore a topic or phenomenon about which one knows very little, one necessarily begins with a general description of the phenomenon. This may sound easy but in fact this is probably a rather difficult kind of study for the beginner to undertake. There are hardly any previous studies to guide and no clearly delineated independent and dependent variables and therefore, no categories within which to classify what one observes. The researcher has few guidelines to determine what is important, whom to interview or how to proceed further in a consistent fashion. For these reasons, the research design of an exploratory study is more open

and flexible than other kinds of research. Main decisions involved before taking up such a study generally pertain to the kinds of mechanical aids needed (e.g. photographic equipment, tape recorders etc.) the key knowledgeable persons with whom one can speak at first, and the selection of insight stimulating cases. But the future paths that these initial steps may lead to are often not possible to foresee.

Under exploratory research may fall observational studies of unique situations, first hand knowledge of radical political or religious movements, clinical case histories of persons, groups and events, and anthropological accounts of least explored tribal societies. It should be kept in mind that an exploratory study is just an initial step in a continuous research process and not a detached self-sufficient entity. The most difficult decisions pertain to its initiation. If an incorrect or irrelevant start is made then the efforts made during the later stages will be of limited use.

1.7.2 Descriptive Studies

The objective of a descriptive study is to describe some phenomenon, situation, event, people, group or community. Basically a fact finding exercise, a descriptive study focuses on relatively few dimensions of a well defined entity. The main aim is to measure systematically and precisely some dimensions of a phenomenon, usually with detailed numerical descriptions. For example, while studying the people of a community, a researcher may be interested to know the distribution by age, sex, caste or education. A descriptive study may also be undertaken to estimate the proportion of people in a given population who hold certain specific views or attitudes. The procedure followed in descriptive studies is broadly the same as for diagnostic studies.

1.7.3 Diagnostic Studies

Diagnostic studies are concerned with discovering and testing whether certain variables are associated as, for example, do people belonging to a particular caste usually vote for a particular political party? The aim of a diagnostic study may also be to determine the frequency with which something occurs or the ways in which a phenomenon is associated with some other factors.

So far as the issue of research design is concerned both descriptive and diagnostic studies share some common features. Both differ considerably from exploratory research and their design is much more structured. The research questions characteristic of these studies presuppose much prior knowledge about the nature of research problem to be studied. The unit of analysis is specified and the researcher is expected to be clear about what he wants to study and how he plans to conduct the research investigation. Decisions regarding the variables, nature of observations to be made in the field, the type of evidence to be collected, and tools of data collection are usually made beforehand. The research design of such studies demands strict adherence to objectivity so as to eliminate any chances of personal bias. The design must be prepared with utmost care keeping in view the crucial aspect of research economy. Faulty research design may lead to waste of time, labour and money.

As usual, the first step involved in the designing of such studies is that the research problem should be formulated accurately. If research objectives are not precisely stated and the principal areas of investigation not linked to them, systematic collection of required data cannot be ensured. There are chances that a lot of extraneous data is procured and the whole exercise later on turns out to be a wasteful activity. Concepts must be clearly defined and their operationalisation must be ensured so that they are able to speak about the corresponding observed reality in measurable terms.

Once the research problem is precisely formulated, (which implies preparatory work like review of work already done and specifying the objectives of the study) the next step involves certain decisions regarding the methods of data collection. While selecting the method of data collection, the researcher should keep in mind that a particular method — observation/interview/schedule/mail questionnaire has its own advantages and disadvantages. Decisions concerning tools of data collection should take into account the nature of research problem, the scope of the study, type of information needed, the degree of desired accuracy and such other factors. Data

collection procedures should incorporate safeguards against bias and unreliability. In other words, the whole effort must be directed to achieve a near true state of objectivity during the collection of data.

Most of the descriptive and diagnostic studies are conducted to develop an understanding about a group, community or some social phenomenon. Because of financial limitations and constraints of time it is not possible to approach each and every person of the population under study. To overcome this difficulty, a representative sample of the population is drawn so that relevant information is gathered with research economy. In other words, research findings based on a representative sample must correspond closely to those which would have been obtained if all the members of the community were studied. Researchers have developed an elaborate range of sampling techniques which must be made use of while designing a sample.

At the level of data analysis care must be taken in coding responses i.e. placing each item in the appropriate category, tabulating of data, performing statistical computations and so on. Steps must be taken to safeguard against potential errors. The considerations of economy demand that an outline of data analysis be prepared in advance. Advance decisions must be made whether the tabulation is to be done by hand or by machine. Accuracy of tabulating procedures, statistical operations etc. must be checked. Sufficient care must be taken so that wrong conclusions are not reached simply because of faulty procedures of data analysis. In sum, although basic outlines of data analysis must be worked out in advance, complete and thorough planning at the level of analysis is not always possible nor perhaps desirable while preparing a research design.

1.7.4 Experimental Study

We are used to associate experiments with the study of biological and physical sciences. The experimental method of the natural scientist has long been the envy of social scientists. Some maintain that no experimentation is possible in social sciences because of the complexity and dynamic nature of the subject matter, i.e. human beings. Human beings cannot be treated as guinea pigs in a sociological laboratory. On the other hand, it is felt that by reasoning alone it is rarely possible to reach at a scientific understanding of the phenomena, and the experimental method of modern science can be used in social research.

The central feature of the experimental method is to draw inferences by observation of human relations under controlled conditions. Now, in every complex social situation many factors operate. If the social scientist seeks to describe the single causal relation of factor A to factor B, then he must strive to set up an artificial situation in which all other factors such as C,D,E etc. are held constant or controlled. This is achieved by selecting two groups so that they do not differ from each other in significant respects (except by chance). Any one of these groups may be chosen as the 'experimental' group and the other chosen as the 'control' group. The 'experimental' group is exposed to the assumed causal (or independent) variable while the 'control' group is not. The two groups are then compared in terms of the assumed effect (dependent variable). This structure makes it possible to generate the required evidence for testing causal relations among variables, which is already stated in the form of a hypothesis.

In experimental research the demonstration of causal relationship among variables involves three distinct operations; (a) demonstrating covariation (b) eliminating spurious relationships and (c) establishing the time order of occurrence. The first two have already been discussed in the earlier section of this unit. The third operation, time order, requires the researcher to demonstrate that one phenomenon occurs first or changes prior to another phenomenon. For example, the status of parents influences the educational achievements of their children and not *vice versa*. The implicit assumption here is that phenomenon which is yet to occur in future cannot determine here the phenomena in the present or past.

The significance of experimental research is that it allows the investigator to draw causal inferences and observe with relatively little difficulty whether or not the independent variable caused change in the dependent variable. With other research

designs this cannot be easily determined. Thus, in the case of experimental research design utmost care is required while taking decisions about the selection of research topic, units of analysis, variables, operationlization of concepts, the precise formulation of hypothesis, tools of data collection, and data analysis.

Notwithstanding other aspects, experimental research design consists of three components — comparison, manipulation and control.

a) Comparison: Comparison allows us to demonstrate covariation, i.e., association between two or more variables. In other words, through comparison we can know that two variables are correlated. A common example is the correlation between cigarette smoking and lung cancer. If the initial hypothesis is that smoking is associated with greater risk of getting lung cancer then one may test it by comparing the distribution of cancer cases among smokers and non-smokers.

b) Manipulation: Manipulation helps in establishing the time order of events. The notion of causality implies that if Y is caused by X, then an induced change in X will be followed by a change in Y. In other words, one variable is the determining force and the other is a determined response. To establish this fact, the induced change in X would have to be prior to the change in Y. For example, if a particular programme (say X) is to influence the living conditions of the rural poor, then it has to be demonstrated that improvement in their condition takes place only after exposure to such a programme. This can be accomplished by measuring the extent of poverty before and after the introduction of the programme. The major evidence required to determine the sequence of events i.e. the independent variable precedes the dependent variable, is that a change occurred only after the activation of the independent variable.

c) Control: Control enables us to determine that the observed covariation is nonspurious. In other words, it means that all other factors must be ruled out as rival explanations of the observed association between the independent and dependent variables. The process of control is related to the internal validity of experimental research design. Factors that may jeopardise internal validity are of two types, i.e., intrinsic and extrinsic. Extrinsic factors are called selection effects which are generally in the form of biases resulting from the differential recruitment of respondents to the experimental and control groups. Intrinsic factors refer to changes in the individuals or the units studied that occur during the study period. For example, death of a person may prevent the researcher from obtaining complete information on all cases. Again, individuals may drop out from the experimental or control group and the final information on such a sample may be biased. We will not discuss here the nature of all the intrinsic factors which may influence an experimental study. But you must note a significant aspect that the probability of an experimental study to be affected by changes due to intrinsic factors remains the same for both the experimental and control groups. As regarding the extrinsic factors, two methods, viz., matching and randomization are employed to counter their adverse influence.

i) Matching: Matching is a way of equating the experimental and control groups. There are two methods to match the groups — precision matching and frequency distribution. With the former method (also known as pairwise matching), for each case in the experimental group, another one with identical characteristics is selected for the control group. For example, if you want to control the effect of age in a study, then against each individual of specific age category selected for one group you must select another individual of the same age category for the second group. The main drawback of this method is the difficulty in matching when a large number of variables are involved. For example, if you are interested in simultaneously controlling variables such as age, sex, caste, and education, then you are supposed to find matching pairs having the same combination of characteristics for both the experimental and control groups. But such as exercise is cumbersome and, in practice, it is generally difficult to find such matching pairs.

An efficient method of matching is frequency distribution. The experimental and control groups are made similar for each of the relevant factors separately. Instead of one-to-one matching, the two groups are matched on central characteristics. For

example, if age is the variable to be controlled, then the average age of the group should be equivalent to that of the other. Similarly, if the variable of sex is to be controlled then care should be taken that both the groups have the same proportion of males and females. The frequency distribution method is somewhat less precise but is easier to execute than precision matching. It enables the investigator to control for several factors without having to discard a large number of cases. It facilitates the process of recruiting the required individuals in an experimental study.

ii) Randomization: We have seen that matching is a method of control for a limited number of predefined factors. But one can never be sure that all factors which may be associated with variables under investigation have been identified. There may still be other factors beyond the knowledge of the investigator which may influence an experimental study leading to erroneous causal interpretations. This problem can be avoided by resorting to randomization. *Randomization* is a process which ensures that any of the given cases has an equal probability of being assigned to either the experimental group or the control group. Such a random selection of the cases can be accomplished by flipping a coin, by using a table of random digits and so on.

For instance, a researcher wants to examine the hypothesis that adequate facilities (improved housing, medical care etc.) are conducive to the performance of agricultural labour. Suppose workers are simply divided into experimental and control groups. The experimental group is furnished with adequate facilities. The production level of both the groups is measured at the beginning and at the end of the experiment. In pursuance of the main objective of the study, the researcher can know whether workers having adequate facilities have performed better than those in the control group. However, there may be a basic flaw in such a procedure. It may quite be possible that all workers belonging to the experimental group are highly motivated, more intelligent, physically fit or younger in age. One can never be sure whether personal factors such as age, fitness, intelligence and motivation rather than adequate facilities (independent variable) have contributed towards the better performance of the experimental group. Now randomization can counteract the effect of all the extraneous factors. It will assure similar distribution of all prior characteristics of the workers in both the groups. One can expect that all factors will be almost same in the two groups. Consequently, any difference in performance between the control and experimental groups can be safely interpreted as the effect of adequate facilities. Thus the advantage of randomization is that it helps to control several factors simultaneously. The researchers can equalize the experimental and control groups on all initial differences between them.

An additional significant question concerns the extent to which the research findings can be made applicable to broader social situations. This aspect is ensured by the representativeness of the sample. If the sample of a study is drawn from a single village, the findings of research may not be valid in the context of other rural areas. However, if the sample of the study covers a large area, there is a strong possibility that the results of the study will represent the true rural situation. There exists a whole range of sampling techniques with which the researcher must acquaint himself thoroughly while preparing an experimental research design. Besides this, there is much more involved in the designing of an experimental study and also there are many types of experimental research designs. We will not discuss all those aspects since our aim has been to acquaint you with the basic elementary aspects of research design.

Check Your Progress IV

Notes : a) Use the space provided below for your answer.
b) Compare your answer with the model answer given at the end of the Unit.

1) Complete the following statements:

a) The research design in exploratory studies follows a ..... and ..... approach.

b) An exploratory study does not ..... hypotheses. It collects facts which help in the subsequent ..... of hypothesis.

c) Diagnostics studies follow a more ..... design.

d) Experimental research demonstrates ..... relationships.

e) ..... enables us to determine that the observed covariation is nonspurious.

f) Equating of experimental and control groups is done through .....

g) The process that ensures that any of the given cases has an equal probability of being assigned to either the experimental group or control group is called .....

h) The most important characteristic of a sample is its ..... character.

i) In a simple random sample, we ensure that each unit has an ..... chance of selection.

2) When is an exploratory research design considered necessary? (Hint: See the text.)

.....

.....

.....

.....

3) What are the distinguishing features of a diagnostic study? (Hint: See the text.)

.....

.....

.....

.....

4) Name three salient features of an experimental research design. (Hint: See the text.)

.....

.....

.....

.....

5) Name the methods which are employed to encounter the adverse influence of extrinsic factors in an experimental study. (Hint: See the text.)

.....

.....

.....

.....

1.8 SAMPLING

We have earlier mentioned about the use of samples in the different types of research (exploratory, descriptive, diagnostic etc). Here we will give you a very preliminary idea about sampling. Most social science research is based on the study of some and not all the units. This is called a sample study. When each and every unit of the total population (or universe for the study) is covered it is called a census study. Please note that the unit may be a person, a household, a village or any other entity depending on the design of the research study.

A census study covering all units requires enormous resources — money, time and manpower. Moreover, when a very large population is involved, quality is difficult to control. The items on which information is collected have also to be limited.

Sometimes, the nature of the study may be such that it is impossible for the researcher to study all the units. A sample study does not have these disadvantages. Since the time taken in sample studies is short, the usefulness of the study increases. Social science research, therefore, depends heavily on sample studies.

The essence of sampling is that the selection of a part (i.e. sample) from the whole (i.e. the total population) is made in such a way as to permit inferences in regard to the whole i.e. generalise for the whole population. The important thing to ensure is that the sample is representative, i.e., it accurately reflects the population from which it is drawn.

Another important consideration is the size of the sample. There are scientific criteria on the basis of which a sample size is decided. It is not arbitrarily fixed at 5 per cent or 10 per cent size, for instance. Two important factors are — the degree of accuracy and the extent of tolerance of sampling error in the characteristics being studied.

There are two broad types of sampling — probability and non-probability. Let us briefly describe these.

Probability Sampling:
In probability sampling, each unit of the population (universe) has an equal chance (probability) of being selected. There are, however, chance factors due to which the sample and the total population (universe for the study) may show some differences. This is due to what is called as sampling error and is measured by a statistic called standard error. Probability sampling is sought after because it provides a more representative sample and enables estimates to be made for the population by calculating the standard error. The probability samples more frequently used in social science research are simple random sampling, stratified sampling, systematic sampling and cluster sampling.

Simple Random Sampling: In this type of sample, each unit in the population (universe) has as equal chance of being represented in the sample. This equality of chance is the special hallmark of simple random sample. Now while designing a simple random sample two aspects must be taken care of:

- a precise definition of the units.
- a complete enumeration of the units so as to give the complete sampling frame.

To satisfy the condition of randomness, the researcher cannot simply pick up a case in any hit or miss fashion as that could introduce bias since all units would not have had an equal chance of being selected in the sample. Mechanical or electronic aids are often used to ensure equal chance of all the units in a population to be selected. In practice, social scientists achieve randomness by using tables of random numbers which are readily available.

Stratified Sampling: This type of sample involves breaking up of the population into various sub-groups or strata which are relevant for ensuring better representation in the sample. A sample is taken from each strata according to the overall sampling design. Although stratified sampling follows a more complex procedure than simple random sampling, it ensures a more representative sample since it ensures that the different strata considered important for the study get represented in the sample.

Systematic Sampling: In this technique, the total population is first divided by the size of the sample population. For instance, if the total population is 1000 and the sample size is 10 per cent i.e. 100, then 1000 divided by 100 gives 10. For drawing the sample, every tenth person is selected from the list enumerating the population. The problem of starting from 1st, 3rd, or 4th or any other serial order is resolved by drawing the first number randomly. Suppose this is 6; then the units which will be selected are those at serial number 6, 16, 26, 36, and so on (i.e. at an interval of 10). Before drawing a systematic sample, it is necessary to check that the list of the total population has not been drawn in such a manner that there is a particular periodicity of occurrence in the listing of the persons as thereby the danger is that certain types of persons will be systematically included or excluded.

Cluster Sampling: In such a sampling, the population is broken down into groups of cases called clusters. A sample of clusters is then selected at random. Such clusters

usually consist of natural groupings such as, for instance, geographic units i.e., development blocks or tehsils. While in stratified sampling the cases are drawn from each stratum, in cluster sampling, cases are drawn only from those clusters which are selected for the sample.

Non-Probability Sampling: Non-probability samples are resorted to in situations where it is not possible to follow the probability sampling techniques because the universe for the sampling frame is not available and it would be very time consuming and difficult to list the total population from which the sample is to be drawn. For instance, if you were making a study of alcoholics, beggars etc. it is impossible to have a complete enumeration of the entire population to be able to draw the sample. The limitation in such samples is that the probability of the members being included is not known. There is, therefore, no way of calculating the amount of error inherent in the estimate made for the total population from the sample. As we read earlier, in the case of probability sampling this is possible through the computation of standard error.

The main types of non-probability sampling are Purposive Sampling and Quota Sampling.

Purposive Sampling: In purposive sampling, the researcher relies on his expert judgement to select units that are representative or 'typical' of the population. Generally, the important sources of variation are identified in the population and then a sample which reflects that variation is selected. For instance, you may select a single unit or subgroups of the population which you think to be typical of the whole population in important respects.

Quota Sampling: Quota sample begins by dividing the population into relevant strata such as age, gender, geographic region, or some other attribute. To obtain a correct proportion of the sample, interviewers are asked to gather information from a fixed quota of respondents in each stratum. For instance, you are to interview a certain number of men, women, persons of a given age or income. In order to fill the quota you are free to choose anyone who meets the quota requirements.

1.9 LET US SUM UP

In this unit we have discussed some of the basic elements related to the designing of a research study. However, before discussing the research design, we have explained to you the nature of scientific knowledge and the main characteristics of scientific method such as objectivity, empiricism, relativism, ethical neutrality, skepticism, and simplicity. Only that knowledge which is acquired with the help of scientific method attains scientific status. Scientific knowledge is united by the scientific method and is independent of the nature of content of any given discipline. Research is the process by means of which specialized knowledge is generated in various disciplines.

Later we have explained the meaning of research design followed by a discussion of its main elements. We have considered such basic aspects as the selection of a topic for research, units of analysis, variables, causal relationships, meaning of direction of influence, and non-spuriousness in research. The meaning and function of concepts and the importance of clarity and precision in the use of concepts was also discussed. We mentioned that operationalization of concepts and conceptual definitions are important exercises to bridge the gap between pure theoretical constructs and observable concrete reality. The meaning of hypothesis, its functions and its use in research studies was discussed. We noted some of the major criteria of a good research hypothesis.

We stressed that the designing of research study requires careful foreplanning. Decisions concerning different issues and anticipated problems should be made sufficiently in advance. An efficient research design helps to avoid wastage of time, energy and labour hence contributing towards the successful completion of a research study.

We have discussed the main types of research design. The character of a research design varies with the nature and purpose of a research investigation. We have also discussed the general nature of exploratory, descriptive, diagnostic and experimental research designs and briefly described their main features.

Lastly, we have given you a general idea of sampling and described very briefly the meaning and types of probability and non-probability sampling.

1.10 KEY WORDS

Concept: An abstraction communicated by words or other signs that refer to common properties among phenomena.

Control Group: The group under study in which potentially confounding variables are kept under control by holding them constant or by preventing them from varying during the course of observation or statistical analysis. Controls refer to procedures that eliminate, as far as possible, unwanted variation, such as sources of bias and error that may distort study results.

Empiricism: The theory that all concepts are derived from experience and that all statements claiming to express knowledge depend for their justification on experience.

Experimental Group: The group under study which is subjected to manipulation and control. Greater control is exercised over conditions of observation to effectively eliminate the possibility of extraneous factors.

Fact: An empirically verifiable observation.

Hypothesis: An expected but unconfirmed relationship among two or more variables.

Matching: A technique for assigning subjects to experimental and control groups so that the composition of each group matches the other on one or more characteristics thought to be related to the dependent variable.

Objectivity: In the usual sense of the term it refers to such observation that is free from emotion, conjecture or personal bias of the researcher.

Operational definition: The detailed description of the research operations or procedures necessary to assign units of analysis to the categories of a variable in order to represent conceptual properties.

Primary data: Data obtained directly from the original source through a specific investigation rather than from a published source. Social science research depends heavily on primary data collected through observation, schedules, interviews, etc.

Relativism: The view that beliefs and principles do not have universal or timeless validity.

Sample: A portion of the total number of units (referred to as universe).

Secondary data: Data which are at least one stage away from the original source and are normally available in some processed and structured form. Examples of secondary data are official statistics, census reports, statistical abstracts, etc.

Variable: Characteristics of units of analysis that vary i.e., they take on different values or attributes for different observations. Variables may vary over cases, over time or over both cases and time.

1.11 SUGGESTED READINGS

Goode, W.J. and Hatt, P.K., 1952 *Methods in Social Research,* McGraw Hill: London

Jahoda, M., Morton D. and Stuart, W.C., 1951 *Research Methods in Social Relations,* Dryden: New York

Royce, Singleton, Jr. *et al.* 1988 *Approaches to Social Research,* Oxford University Press: Oxford.

Young, P.V. and Schmid, C.F., 1977 *Scientific Social Surveys and Research,* Prentice-Hall: New Delhi.

1.12 MODEL ANSWERS

Check Your Progress I

1) (a) True (b) False (c) True (d) False (e) True (f) True (g) True (h) True (i) True (j) True

2)
- Objectivity
- Empiricism
- Relativism
- Ethical Neutrality
- Skepticism
- Simplicity

3) To produce knowledge or extend/modify/verify existing claims of knowledge.

Check Your Progress III

a) Research design
b) Expenditure of time, labour and money
c) Structure and state of a discipline, social problems, practical considerations
d) Theoretical; practical
e) Units of analysis
f) Explanatory variable; extraneous variables
g) Relationships
h) Causal relationships
i) Hypothesis
j) Constant
k) Quantitative

Check Your Progress IV

1) a) Flexible; Open
b) Test; formulation
c) Structured
d) Causal
e) Control
f) Matching
g) Randomisation
h) Representative
i) Equal

UNIT 2 TOOLS OF DATA COLLECTION

Contents

2.0 Aims and Objectives
2.1 Introduction
2.2 Research Strategies and Nature of Data in Social Sciences
2.3 Main Tools of Data Collection
2.4 Observation
2.5 Interview
2.6 Schedule
2.7 Mailed Questionnaire
2.8 Let Us Sum Up
2.9 Key Words
2.10 Suggested Readings
2.11 Model Answers

2.0 AIMS AND OBJECTIVES

The aim of this unit is to acquaint you with a very important stage in a research study namely, the collection of data. Various tools have been developed over the years to enable a researcher to collect data in a manner which meets the requirements of a scientific inquiry. After studying this unit you should be able to:

- assess the strengths and weaknesses of different tools;
- decide which tools are most appropriate for a given situation;
- develop the different tools with guidance from a senior researcher;
- appreciate how the tools relate to the objectives of the study and the tabulation and analysis of data.

2.1 INTRODUCTION

In the previous unit you have studied some of the fundamental issues concerning the designing of a research study. After preparing the research design, the next task facing the researcher is implementation of the research plan into action. This involves the process of gathering data and deciding on the means by which such data are to be collected. The means or techniques which help in the procurement of data are often referred to as the tools of data collection. Obviously then, the foremost requirement at this stage is the selection of appropriate tools of data collection. Broadly speaking, the nature of the research problem, the type of research design, the type of respondents and their geographical distribution, the professional competence of investigators available for collecting data, the period within which the study is to be completed and the funds available etc. determine the selection of tools for the collection of data.

Let us now acquaint you with the different tools, their main features and their strengths and weaknesses. But before that a brief exposition of the nature of data in social sciences will be helpful.

2.2 RESEARCH STRATEGIES AND NATURE OF DATA IN SOICAL SCIENCES

Social scientists generally employ four principal research strategies for understanding the social world, viz., field research, surveys, experiments and the use of available data. Each discipline tends to favour a particular strategy. For instance, anthropologists characteristically conduct field research, sociologists most often do survey research, psychologists typically conduct experiments and historians tend to make use of available data. However, all the four strategies are important to the world of social research and one or more of the four can be used to study most social

phenomena. But the nature of data and requirements of data collection do vary from one strategy to another.

Field research is essentially a matter of studying a naturally occurring set of events so as to gain first hand knowledge of the situation. For instance, an anthropologist lives in a remote community for long periods in order to study the culture of its inhabitants. Sometimes, sociologists too take up work in an organisation without disclosing their true purpose, i.e. to gather inside information about the functioning of the organisation and to understand the organisation in a way its members see it. The most important tool for this kind of field research is observation. There are other tools also that are used in field research such as the schedule, the interview and the mailed questionnaire. We shall read about them later in this unit.

The second approach employed to study a social phenomenon is the survey research. Survey research usually covers big groups of people or large sections of the population. The term is used to describe a wide range of investigations, the objectives of which may cover an equally wide spectrum. The nature of data is mainly determined by the type of the research problem and the research objectives.

Surveys aim to find facts. They could be primarily **descriptive,** i.e. they seek to describe the distribution within a population of certain characteristics, attitudes or experiences, or **explanatory,** i.e. they investigate and attempt to explain the relationship between two or more characteristics or events. For example, you may be interested to determine the relationship between the incidence of a particular disease and the socio-economic characteristics of the people in a rural set up. You will, in that case, gather information about the salient socio-economic characteristics like age, gender, family, caste, social status, income, occupation, etc. of all those people who are suffering from that disease and then study the relationships.

Surveys can further be characterized as cross-sectional or longitudinal studies. In the former type, data on a sample constituting a cross-section of the people are gathered at essentially one point in time. In longitudinal studies, however, the individuals in the sample are observed at predetermined stages over an extended period of time in order to ascertain the changes in a particular characteristic or attitude. The number of respondents here is usually smaller and, at times, less representative than the cross-sectional survey, but the nature and direction of change can be more readily ascertained from a longitudinal survey. Generally speaking, longitudinal studies are of long duration, require more funds and are comparatively difficult to conduct. There is also the problem of attrition in the sample and finding replacements. However, since such studies yield valuable information on changes over time, in some situations they are preferred.

The term 'experiment' in social research denotes studies with certain distinctive features. The main reason for doing experiments is to test a causal hypothesis so that the relationship between variables becomes evident. All experiments possess certain basic requirements that permit strong inferences in terms of cause and effect, i.e., one variable causes changes in another variable. The two key features of the experimental approach are manipulation and control. In order to test a hypothesis, the researcher deliberately manipulates or introduces changes in the environment of the subjects and observes or measures the effects of the changes. Also, greater control is exercised over conditions of observation than in any other research strategy. Efforts are made to more effectively eliminate the possibility of extraneous variables influencing the research findings.

For example, you may be interested to study the impact of some new teaching method on the performance of students belonging to a school in a village. This can be done by following a 'before after' exprimental design or by having an experimental group — control group, experimental design. In the case of the former, the performance (dependent variable) **before** the introduction of the new teaching method (independent variable) is compared with the performance of the students **after** the introduction of the new teaching method. The difference is attributable to the independent variable, i.e. the new teaching method.

In the experimental group-control group, design, two groups of students belonging to the same age group or class could be selected with the help of random sampling.

Random assignment of students to the two groups will ensure that pre-experimental differences of age, sex, intelligence, social background and other relevant attributes are neutralised or distributed approximately evenly among both the groups — i.e. to the experimental group and to the control group. Further, the experimental and control group students must experience the same events or conditions (e.g. same class room environment, boarding or lodging facilities etc.) during the period of experiment except for the manipulation through the independent variable (in this case the new teaching method). In other words, all other factors being equal, the experimental group is subjected to the treatment of the new teaching method while the control group is not. In the main, the pre and post experimental observations constitute the data. The performance of the experimental and control groups is systematically observed and recorded before and after the experiment. The change in performance, if any, can be construed as the cause-effect of the independent variable (i.e. new teaching method) on the dependent variable (i.e. performance of students). The data in experimental research are precise, compact and to the point. Data are also limited in nature i.e. they are mainly constituted by the observations made on the variables involved.

The fourth general strategy for doing social research is to make use of available data such as that produced by the Census or National Sample Survey Organisation. Sometimes, the research utilises data that were originally produced for a completely different purpose. Such data are often subjected to some form of scrutiny and processing before analysis, keeping in view the objective of the study. Data also exist in the form of public documents, official statistics, court records, private documents (such as autobiographies, diaries, personal letters), mass media such as newspapers, magazines, journals, films etc. In spite of the extraordinary diversity of available data, a major problem with this approach is of locating the relevant information as these are often not kept in a systematic form. This can be a time consuming affair. Also, the researcher should have adequate knowledge about the functioning of libraries, archives, data storing facilities, etc. A major limitation is that the quality of such data shows considerable variation. Quite often, the definitions and the context are not known.

Check Your Progress I

Notes: a) Space is given below for your answers.
b) Check your answers from the text.

1) What is the basic nature of a longitudinal study?

.....

.....

.....

.....

2) How does a before-after experimental design proposes to measure impact?

.....

.....

.....

.....

2.3 MAIN TOOLS OF DATA COLLECTION

Most research studies collect fresh data from the respondents even though already existing data are utilised for developing the research design or supplementing the data to be collected. We will discuss here three important and widely used tools of data collection used by social scientists, viz.,

- Observation
- Interview
- Schedule
- Mailed Questionnaire.

Data are also collected through projective techniques, but we shall not cover them here. Nor shall also we be covering the different techniques for developing scales

which measure attitudes, opinions, etc. for reasons of space and also because they require a more advanced level of learning.

2.4 OBSERVATION

Observation is a basic method of getting information about the world around us. As long as we are awake, we are almost constantly engaged in observation. For instance, we get up in the morning and observe whether it is a sunny or a rainy day and make our decision about carrying an umbrella accordingly. Not only is observation part and parcel of our daily life but many types of data required as evidence to support social research are also obtained through the observational method. The greatest asset of observational technique is that it is possible to record the actual occurrence of social events. While many research techniques depend mainly if not entirely on recalling the past events, observational method yields such data as are related to real life situations. A trained researcher can even observe and record all the minor details of a community with the help of this technique which to others might seem insignificant.

It is important to note that all social phenomena are not open to observation. For example, a researcher cannot have access to the private life of the people. Again, all the events which can be studied through observation, may not necessarily find a ready observer. Notwithstanding these aspects, empirical research always involves some kind of observation.

Observation may serve a variety of research purposes. It can be used in exploratory research to develop a preliminary understanding of social phenomena. It can be applied to study real life situations as well as to conduct experimental research. Again, it can simply be used to collect supplementary data in support of other tools of data collection. Observation includes the most casual and uncontrolled experiences as well as exact recording as is done in experimentation. In fact, observation is useful for studying simpler as well as complex research problems.

Irrespective of the purpose of study, four important dimensions of observational technique are:

1. What is to be observed?
2. How to record observations?
3. How to ensure the accuracy of observations? and
4. What should be the relation between the observer and observed?

Observation becomes a scientific technique when it:

- is planned and recorded systematically;
- serves a formulated research purpose;
- is related to the more general propositions of scientific method; and
- is subjected to checks and controls on validity and reliability.

Observation technique can broadly be characterized as **uncontrolled** or **controlled.**

2.4.1 Uncontrolled Observation

Most of the knowledge which people have about their social environment is derived from uncontrolled observation. By uncontrolled is meant that observations are made without the help of any standardized techniques or rigid guidelines. The researcher observes the social phenomenon by looking at or listening to the people. He learns about the social situation simply by witnessing or participating in the activities of the people. No checks or controls are exercised on the observer and the observed. In other words, the validity of observations made by an observer cannot be cross checked by others.

Uncontrolled observation is of two types. Participant observation and nonparticipant observation.

i) **Participant Observation:** In this method the investigator conceals his real purpose in order to be accepted as a member of a group or community. The underlying assumption is that if the members of a community remain unaware of the researcher's

purpose, their behaviour is least likely to be affected. Participation ensures that the emotional reactions of researcher to all the events could be more or less similar as that of the actual members of a community. In fact, the researcher acquires the true feeling of all the activities through participation. The process helps to record the natural behaviour of the group. It makes possible the gathering of all such information as could not easily be obtained by merely looking in a disinterested fashion. The information collected has greater depth of experience. Furthermore, since participation can be for a short or a longer period the range of material collected is usually much wider. No other method (for instance, the interview or the schedule) can match participant observation on this count. Lastly, the researcher can also know and record the context of information thereby imparting richness and depth to data.

Participant observation has also some obvious disadvantages. There are chances that by becoming a participant, the researcher narrows his range of experience. Sometimes, he may take on a particular position within the group, associate with a friendship circle or develop an attitude of sympathy towards certain situations. This way he may acquire a partial or a partisan view, engage in ego satisfaction, or may react in anger or sympathy rather than really observing others. Due to such emotional participation the researcher comes to lose objectivity which is his greatest asset. In case there is a hierarchy of power or prestige in an organisation the researcher, while occupying one position, may foreclose to himself many other avenues of information. Lastly, the researcher may become so much acquainted with a situation that he may take many aspects as granted or perfectly natural and hence, frequently fail to note certain important details. In participant observation another problem relates to the standardisation of the technique due to which another researcher may not be able to repeat the procedure. Testing for reliability therefore becomes difficult.

ii) **Non-Participant Observation:** We have noted above some of the difficulties encountered within the method of participant observation. Notwithstanding these limitations, in many situations it may not at all be possible for the researcher to become a true participant. For instance, a researcher cannot embark on criminal behaviour in order to study the functioning of a criminal gang. Some of these problems can be overcome by the method of non-participant observation. By non-participant observation is meant that the researcher need not actually participate in the people's activities. He can live among them and can make his observations without performing any particular role in the community or organisation. However, purely non-participant observation is difficult. A community or group may develop apprehensions about the researcher as a non-member and outsider who is always present but never participating. In such a situation both the researcher as an outsider and the group as a whole may feel uncomfortable.

Non-participant observation at best, then, is quasi-participant observation. In this method, the researcher actually shifts from one role to another in the field. It is possible to take part in a number of activities of the group. Also, carrying out different roles is simpler than attempting to conceal one's identity and purpose completely. For example, while studying a rural society, a researcher may join a marriage ceremony as a participant but during harvesting he may simply sit and record the details of work sharing etc. Such a shift is made easier by the fact that patterns of social behaviour are not likely to be changed in important ways by the presence of the researcher once his role is properly defined and accepted by the community.

2.4.2 Controlled Observation

The main considerations in observational method are: i) Whether the situation is to be a natural one or a contrived one and ii) whether those observed are to be aware or unaware of the observer. In most uncontrolled observation, the situation is natural and those observed may or may not be aware of the researcher's presence. This may also be true of controlled observation. What is important in controlled observation, however, is that the researcher is ordinarily not limiting the activities of the observed individuals. Rather he tries to systematize the process of observation. The underlying assumption is that the researcher always affects the process of observation in some fashion; and such effects need to be reduced, limited to minor areas or at least made measurable with the help of some standardized checks or controls.

For observational method to attain a scientific character, there is need of some objective standards against which observations need to be evaluated. In other words, there must be certain controls or checks on the researcher's bias, selective perception and vagueness of his observations. Such controls on the observer can be exercised with the help of mechanical devices, team observation, films and recordings, standardized schedules with elaborate categories, etc. For example, instead of vague guesses about distances, correct and exact maps can be used; casual hunches about the nature of social interaction between two persons can be replaced by definite counts, cross tabulation or diagrammatic presentation. Through controls the research acquires greater depth, sharpness and the researcher relies less and less on uncontrolled observation.

The controls help in the subsequent verification of data. The researcher must report on how he made his observations, under what conditions, and when? This way, other researchers are enabled to repeat the observations. It also becomes possible for him to know the limitations of his data. While uncontrolled observation is most useful in exploratory studies the extreme form of controlled observation is to be found in experimental research in which research controls are applied to both the observer and the observed. The situation is a contrived or manipulated one, in which definite stimuli are introduced. The observations are made with the help of standardized scientific techniques so as to determine the causal relationships among variables. Most of the extraneous variables are controlled. Both the observer and observed are subjected to checks and controls. Exact state of laboratory experimentation can hardly be achieved in social sciences. Most of the controlled observations in social research at best can take the form of quasi-experiments, i.e., the situation in which observations are made is both natural as well as contrived one.

2.4.3 Research Aids in Observation

By now you have realized that very few controls are exercised both in participant and non-participant observation. Such uncontrolled observation is most useful in exploratory studies whereby the researcher's sole aim is to develop preliminary insights about a social situation. But even in such field research, the observer's task is facilitated if he applies certain research aids. First of all, what is necessary is a good plan for entering the group or community. It is the experience of most field researchers that after the initial period of introduction and explanation, the members of the community or group accept the presence of the field worker as legitimate. If the first contacts are satisfactory, the succeeding contacts are facilitated. Further, a researcher must be a good listener — a pupil who is eager to learn. In his mind he must attach greater significance to the community he wishes to study.

The most important aid in this type of research is the maintenance of a field log (diary recording field events and observations) on a daily basis. The record should be maintained under appropriate headings and sub-headings. It is also profitable to keep running notes on cards or in a note book. In any event, records must be relatively complete and nothing should be left to memory. It is erroneous to believe that a particular event or occurrence is so striking that it will not be forgotten. At the end of each day, an attempt must be made to make a full log from the notes taken during the day. Striking details noted during the early phases must be carefully recorded because later on many such details which now seem obvious may be forgotten. Reading of notes at intervals helps to organise or reorganise data. It also helps to determine which data are more important and to correct any errors while still in the field. By continuing attempts at organizing of observations new categories will emerge thus providing leads and clues for further observations. The preliminary categorization of data is useful if it is done each day.

It is also useful in field work to supplement uncontrolled observations by schedules of basic information. This must be drawn in outline before beginning the research work and should be revised in the field. Such schedules often contain basic data such as age, sex, and number of individuals, occupational structure, religion, income, hierarchy of power, family patterns etc. Even when these items are not the principal focus of a study, they will be essential for any description of the group, community or organisation under study.

In the case of rural situations, maps must be used so as to report accurately the location of the village, arrangement of houses, their proximity to one another and such [illegible]. This helps to set the data in their physical context. Neighbourhood [illegible], people's likes and dislikes, structure of influence, modes of social interaction etc. can be given graphic or diagrammatic representation.

Modern technical aids such as tape recorder, camera, and audio-video equipment are very useful in recording information. These are used in the full knowledge of the persons or groups that are being observed.

Only relevant categories of facts must be given priority while making observations. Those which do not fit into the research scheme need to be ignored. The important aspect is that data in such research cannot be collected merely with the help of observational method. The data needs to be supplemented with other means such as interviews, schedules and questionnaires.

2.4.4 Strengths and Weaknesses

Since observation is one of the basic tools of data collection it certainly has some special advantages in social research. It enables a researcher to gather information about naturally occurring events or social phenomena. Since the data are collected by direct observation, chances of distortion due to loss of memory, false reporting or communication are minimised. The process of data collection does not depend on the subject's ability or willingness to report. Some data can be collected even in situations where the subjects of study are unable to speak or if there exists a language barrier between the observer and the observed.

Observational method, as we have seen in the preceding discussion, is primarily directed toward describing and understanding the actual occurrence of a social phenomenon. It is less effective in giving information about a person's perceptions, beliefs, feelings, motivations, expectations or future plans. Observation requires a higher level of expertise and a greater demand on the time of the researcher — at times long durations and odd hours. He must wait for the specific occasion or event to happen. Observation as a tool is less standardised than the schedule. It also costs more and is suitable when the number of researchers is very small. Observation provides hardly any information about past events or private behaviour of persons, i.e., activities which by their very nature are not possible to observe. To obtain such information other tools of data collection, viz., the interview, the questionnaire and the projective method, have been devised.

Activity 1

Visit a village community to study the difference in the upbringing of boys and girls. Use observation as the tool. Record your observations.

2.5 INTERVIEW

All of you have some experience of interviews. You might have appeared for an interview at the time of admission to school or college or when you were looking for a job. You may have had an opportunity to interview other persons. Enquiring from a person about life in the village to which he belongs or his experiences with a rural bank involves the use of the interview method to gather information.

An interview is a face to face interaction between two individuals in which a person asks questions from another person in order to gather information. Interview emerged as a tool of data collection by the turn of the last century and has by now become an integral part of social research. During earlier times, interviews were conducted more in the nature of probing conversation. Guided by a careful observer this method was used as a powerful instrument for obtaining information. Later on, however, the tool of interview acquired greater sophistication in terms of scientific standardization and reliability.

In the interview method the researcher prepares an interview guide listing the topics or questions to be covered. It is generally in the form of broad guidelines and allows

considerable flexibility as to the order and wording in which the researcher asks the questions. He approaches the people he wishes to study to seek answers to such questions. Heavy reliance is placed on a person's verbal report. Such information may be related to an individual's experiences to which he is exposed or to his knowledge of certain events. A fundamental aspect about the interview method is that the starting point is always a person's self-report and this method can obtain only such information as the concerned person is willing or able to furnish. On the whole, the method is quite flexible and adaptable to specific situations. The main benefit is that face-to-face interaction helps to clarify any doubts or queries on the part of respondents while in the field. The interviewer can provide emotional support to enable the person to answer the questions in difficult situations. Interviewing also permits probing at depth and helps in understanding the context of data. But the main drawback is that the process is time consuming and involves more investment of money and skilled manpower.

Interviews can be used with almost all segments of the population. This tool can be useful in all the three major types of research, viz., exploratory, survey and experimental research. It can be used as an exploratory device so as to identify certain variables or to develop a hypothesis. Though the interview stands on its own as a useful tool of data collection, it is used more frequently to supplement other methods of data collection.

2.5.1 Interview Guide

A researcher develops an interview guide before he beigns the actual task of interviewing. An interview guide contains a set of questions related to the central research problem. All the questions obviously do not have the same form. They may be categorized by the degree to which they are structured, i.e. the way the question is asked and the attendant flexibility in the variety of responses a question permits. For example, a structured question would indicate same few alternative ways of answering it. On the other hand, an unstructured question would allow a wider variety of responses. The interview guide uses a greater proportion of unstructured or open-ended questions.

The unstructured interview is open and flexible in nature. It is adaptable to particular situations in the field. The character or exact form of questions is not determined in advance. The researcher approaches the field with certain general ideas or options in mind. He engages in prolonged conversations with different persons. His main purpose usually is to focus attention on specific social situations or events. Depth is the main consideration. In other words, the researcher is interested to make queries as deeply as possible so as to thoroughly understand the given social phenomenon.

In the case of unstructured interviews, even though there is an interview guide and the researcher broadly knows what will be the principal areas on which he will seek information, it is during the course of the interview that the researcher frames the questions. Usually the answer to a particular question determines the requirement and character of the next question. A series of questions are formulated and asked during the course of interview so as to probe deeply the given social phenomena. No doubt this type of unstandardized interview helps to gain indepth knowledge of an event, but the information gathered cannot usually be checked and verified by other researchers.

To overcome this drawback the researchers often strike a middle path. An interview guide is prepared in a manner that seeks certain items of information about each respondent but allows sufficient scope for the interviewer to rephrase questions in keeping with his understanding of the situation. This permits the interviewer to ask the question in such a manner that the respondent can understand more easily. It also allows the interviewer to probe more deeply when the occasion demands. The process facilitates a more adequate interpretation of the answers and to a certain degree helps towards their standardization. However, this method requires a higher level of interviewing quality. The greater the amount of discretion and flexibility in an interview, the higher the demands in terms of competence on the part of an interviewer.

2.5.2 Fundamental Aspects of an Interview

To the interviewer the primary purpose of an interview may be research but to the respondent its meaning may be different. On the whole, it must be kept clearly in mind that notwithstanding the requirements of research, interviewing is also a process of social interaction. The first and foremost aspect is how to establish rapport with the respondent. The initial contact and acceptance may be achieved through a warm and sympathetic approach. You should never adopt a pompous, patronising or overbearing attitude or dress or talk in a manner which tends to show that you consider yourself as belonging to a superior class. Some type of friendly conversation is useful in establishing good relations with the respondent.

Though a friendly atmosphere is necessary for conducting an interview, mere friendliness is not sufficient. The goal is to obtain the facts to the extent the respondent is capable of presenting them. Thus, from the very beginning, the interviewer must maintain the additional element of professional respect and competence. He should convey the feeling that his sole interest lies in the collection of objective facts. He must give assurance where necessary that the conversation will be kept confidential. In fact, the researcher must establish complete rapport with the interviewee. He must clear all his doubts or apprehensions. Most interviewers have found that after initial contact, few respondents require any further identification. However, it is advisable for the interviewer to carry some identity card or official letter which can help to allay any apprehensions. The experience has been that most respondents are willing to talk if the interviewer is able to establish rapport and play the role of a good listener.

The next step involved in an interview is how to carry it forward. The most important criterion which determines the character of an interview is depth i.e. how deeply the interview attempts to probe a situation. On the one hand, there are extremely lengthy and intensive interviews which probe into the most intimate aspects of a person's life. On the other hand, there are short interviews which merely obtain information about some social aspect. For example, the interviewer may simply want to know about the respondent's views on panchayati raj.

Lengthy and intensive interviews do utilise unstructured or open ended questions. The answers to such questions cannot be easily classified in advance. The researcher can formulate and ask the next question only after knowing the respondent's answer to his earlier question. This involves instant and on the spot decisions about what is to be asked or what is to be ignored. Hence, carrying forward such an interview requires greater skill and competence.

On the whole, the process of carrying an interview forward is greatly aided by asking questions in a fashion which most closely approximates a conversation. While encouraging the interview to grow in an easy conversational style the interviewer must probe those items for which research is being conducted.

It is of particular importance that questions are not asked in quick succession. The interviewer must pause in between questions. Answers should not be cut short as soon as the required information is obtained. When questions related to a different subject matter are being asked during the course of an interview, the transition should be made as smoothly as possible. This can be attained by making appropriate explanatory comments. Silence may crop up during an interview which need not be quickly filled by a hurried question. Such silent moments may be allowed to pass by doing some other activity, e.g., one may pretend to sort papers etc. The respondent should be given adequate time to revert to normal conversation. The interviewer should not take on the role of an educator. This may antagonize the respondent. He should not enter into arguments. His main aim is to gather knowledge and information from the respondent.

The interviewer should not stick to any rigid position. Rather the interview should be conducted in a relaxed atmosphere. It should not be a forced, unreal and emotionally unsatisfactory experience for the interviewer or the respondent. The interview which is developed in an easy and natural fashion approximating a conversation yields good results. Please remember that the respondent does not cease to exist after the interview is over. He will narrate his experience to others who may turn out to be

respondents. A good impression left behind by the interviewer will help future researchers.

The next task in an interview pertains to its recording. The interview must be recorded adequately as the interview proceeds. In the case of intensive and qualitative type of interview which may be having many unstructured probe questions, the problem of recording can become even quite complex. Efforts must be made to record the exact words of the respondent. The statements should not be edited for grammar or meaning. It is necessary to develop skill in writing fast and legibly. However, even the best interviewer will face difficulty in copying all that is said at conversational speed. It is often useful to interrupt the conversation by some adequate comment so as to buy time for recording.

The interview at best can extract information only about a small part of the respondent's experience. Many other aspects which may otherwise seem obvious at the time of interview must also be given due consideration. For instance, the interviewer may gather the impression that the respondent is trying to hide the truth under some stress or fear. Again, the respondent may consistently answer all the questions except making some contradictory statement towards the end. It becomes necessary to record some additional comments which will be useful at the time of analysis. Such aspects should not be left to memory.

After each interview, the notes should be expanded to write down the details while the materials are still fresh in the mind. Delays may distort or blur the details. The comments which are scribbled in haste during an interview may not later on convey all the richness of data. In sum, careful noting and expanding the notes without delay is extremely helpful for adequate recording.

The last stage in an interview pertains to its closing. In the case of a brief interview the closing aspect of an interview is relatively simpler. A normal thanks with a friendly smile may be adequate to take leave of the respondent. However, in the intensive interview, a common experience of interviewers is that closing is more difficult than opening. Generally, the respondent gets emotionally involved if the interview touches upon certain personal aspects. Thus, when the interview is of longer duration and greater intensity, the interviewer should select the occasion for departure more carefully and extend an expression of warm thanks for the time given by the respondent.

2.5.3 Strengths and Weaknesses

By now you must have got an idea about the advantages of interview as a tool of data collection. Let us recapituate:

- The interview method can be employed in almost all the major types of social research and can be used to study almost all segments of the population. The language of questions can be adapted to the ability or educational level of the respondent. It can cover even the non-literate population.
- The interview is particularly suitable for revealing information about complex, emotionally laden subjects or for probing the sentiments that may underlie an expressed opinion.
- Interview is a flexible method. The advantage of flexibility makes it a far superior technique of data collection. The flexibility allows for the rephrasing of questions or making on the spot adjustments in such a way that the social situation can be probed deeply.
- Since the interviewer and the respondent are both present there is always scope for on the spot clarification (or verification). The interviewer can collect supplementary information by observing personal behaviour, facial expressions etc. of the respondent. The interviewer can also assess whether the respondent is under any fear or stress or is trying to hide the truth due to some social pressures. Also, he can record details of the overall environment and context of the interview. Such additional data is very useful at the time of analysis.
- The presence of the interviewer ensures the validity and reliability of reports. Information secured through interviews is likely to be qualitatively superior on some aspects, particularly those which are of a complex nature as compared to other techniques.

The limitations of the interview as a tool of data collection are:

- The main disadvantage is the difficulty of covering larger areas. The recurring uncertainty of finding people at home (or some other place where the interview can be conducted) makes extra demands on time and money, specially when the respondents are at a distance and spread over a large geographical area.
- The method requires skilled and trained interviewers who are often hard to find. The establishment of rapport and carrying the interview forward are delicate tasks. Slight errors on these counts may hamper the prospects of an interview. An unskilled interviewer may collect erroneous data. In a face-to-face situation the interviewer may antagonize the respondent or may get emotionally involved with him. Consequently, it may introduce bias and the real purpose of the interview is defeated.
- The recording of the interview and later on sifting of the information for facts require a high degree of expertise.

Check Your Progress II

Note : a) Use the space below for your answer.

1) List the points to be kept in view for conducting an interview with a respondent.

(Hint: See the Text)

.....

.....

.....

.....

.....

2) In which situations is the interview a preferred tool over the observation method for the collection of data?
(Hint: Read the text carefully and discuss your answer with the Counsellor).

.....

.....

.....

.....

Activity II

Suppose you want to make a study on why people are lukewarm to participate in development programmes, and wish to collect information through the interview. Develop an interview guide to help you. When you next visit a village, use the interview guide and based on your experience, modify it. Record your interview with three respondents using the hints given in the text.

2.6 SCHEDULE

A schedule refers to a set of questions related to a subject, printed or typed in a definite order. It is a device for securing information whereby a person is asked to answer the given set of questions.

The schedule is an important research tool which facilitates the collection of data from large, diverse and widely scattered groups of people. It can be used to collect quantitative data as well as for securing information of qualitative nature. In most empirical studies the primary research tool is the schedule.

The schedule is administered by the research investigator. It is not self-administered. In the case of the schedule, the answers are obtained from the respondent in a face-to-face situation. The responses or answers are noted down by the interviewer. Moreover, the interviewer can act as a stimulus or can furnish on the spot clarifications to the respondent if required. Thus, a schedule presupposes a face-to-face interaction between the interviewer and the respondent. When the

questions are self-administered and require the respondent to answer all questions by himself it is called the mailed questionnaire. Such questionnaires are often sent by mail/post to the respondent. Sometimes they are distributed to a group of people who may have come to attend a conference, with the request that they may fill it up and return the same.

Usually a schedule contains structured items. By structured is meant that questions have fixed wording. They are also typed or printed in a definite order. Thus a particular set of questions having exactly the same wording and sequence is administered to all the respondents.

2.6.1 Open Ended and Closed Ended Questions

The questions in a questionnaire can be open ended or closed ended. Let us take as a very simple example the question, what is your marital status? If the responses are given as unmarried/married/widowed/divorced/separated and the respondent is asked to indicate which of these is applicable to him, then it becomes a closed ended question. Thus a closed ended question offers a largely pre-determined choice of answers to it — the responses are in fixed alternatives. The question is designed in such a way that it permits choice only from the responses mentioned. If the choices are not mentioned and it is left to the respondent to indicate his opinion then it becomes an open ended question.

Sometimes, the answer can be in terms of yes or no, or agree/disagree/no opinion. You may ask the question about the chances of the candidate, belonging to a deprived section, of being elected in a Panchayat election. The options to answer such a question may be listed as Poor/Fair/Good/Do not know, for instance. Thus from different alternatives provided, the respondent may pick up one approximating to his own situation or viewpoint. If the alternatives were not provided, the respondent's answer would be recorded.

Closed ended questions are useful when data are to be presented in a categorised form following a definite system of classification. It ensures standardization and greater uniformity in presentation of results. The data can be easily coded as the responses fall neatly into one of the given alternatives. It considerably facilitates the tabulation of the responses. The provision of alternative answers helps to make clear the nature of the question and the dimensions along which an answer is sought. The closed ended schedule ensures that responses are pointedly directed towards the requirements of a research problem. It can even focus the respondent's attention on a particular dimension of the problem. Since the responses are in the form of clear cut alternatives, it enhances the efficiency and reliability of research. Such a schedule is simple, quick to administer and is relatively inexpensive.

The major drawback of a closed ended schedule is that it may furnish a superficial understanding of a phenomenon. The respondent is left with little scope for reflecting or relating his own experiences. By providing fixed alternatives, sometimes an omission of an alternative while designing the schedule may introduce a bias in the study. The structuring of the different possible responses requires considerable preparatory work and expertise. Again, the researcher may force a statement or opinion on an issue about which a respondent may have a different viewpoint. It is quite possible that none of the given alternatives may, in fact, truly reflect the respondent's thinking. It may irritable or antagonize the respondent and he may give false information.

Open ended questions are flexible and are usually designed to gather qualitative information. The questions are formulated to elicit free response about a particular subject matter rather than limiting the respondent's options to particular alternatives. The distinguishing characteristic of an open ended question is that it raises an issue but does not provide or suggest any structure for the respondent's reply. The respondent is supposed to answer a question on his own terms and in the context of his own experience. For example, one may ask the question what are your views about dowry? To answer such a question the respondent is expected to express his personal opinion based upon his observation and past experience.

Thus open ended questions help in the understanding of a respondent's true response in terms of personal feelings, beliefs or attitudes about a particular research problem and is useful when a particular aspect is to be probed more deeply. Open ended questions have been successfully employed to gather qualitative information. However, the coding of such responses is often a difficult task. The analysis is generally complex and requires more skill. Another drawback is that all the respondents are not equally efficient in expressing themselves.

2.6.2 Important Considerations in the Construction of a Schedule

A schedule should contain a limited number of questions. Only such questions as are extremely important to fulfil the requirements of a study should be included. Questions which may elicit some already known or obvious information should be deleted. Generally, repondents do not like to devote sufficient time and energy for a long schedule.

A schedule is usually divided into several sections. Each section must contain a set of questions related to a particular item or theme. In the light of the requirements of the research problem, some rationale must be developed to include each item or theme. Questions relating to the same general theme must be placed together. Questions placed at the beginning of a schedule should be such that they are able to draw the respondent's complete attention. They should be able to evoke his interest. However, the opening question should be neutral . In other words, the beginning should not contain any controversial issue as the respondents may develop a negative orientation towards the whole questionnaire. This may lead to tardy responses or even outright refusals.

The researcher should carefully determine the sequence of the different themes covered in a schedule. A gradation of themes starting from simpler ones and gradually leading to complex ones should be made. There should also be a logical sequence of themes as well as questions. The transition from one theme to the other should not be abrupt. Complex questions requiring serious thinking are preferably placed somewhere in the middle of the schedule as the respondent may develop fatigue towards the end and may not give to such questions the attention they deserve.

A crucial aspect of a schedule relates to the formulation of questions. These should be framed in such a way that the researcher may logically expect the answers to be significant for his research problem. A good schedule grows from sound study of the problem and review of literature on the subject. Every item included in the schedule (around which a number of questions are to be asked) must be related to the objectives of the study. A tentative list of areas on which questions are to be asked is drawn up. It is useful to consult knowledgeable people and conduct interviews to arrive at greater clarity. In the beginning efforts should be made to cover as many items as possible. Gradually the researcher may detect omissions, gaps or ambiguities. Also, he may determine which items are most important and which are not. Any research venture can adequately cover only a limited number of themes directly related to the research objectives. Most of the questions should focus sharply on such important themes. The actual process of formulating questions requires great skill and expertise. A detailed discussion on this aspect runs beyond the scope of this unit.

The amount of space needed for answering open ended questions has to be determined. For example, suddenly the researcher may find that an open ended question needs more space for the answer than has been provided in the printed schedule. Some sort of preliminary exercise helps in assessing this requirement as well.

The paper used for the printing of a schedule should be of good quality. It should allow the use of both ink and pencil. The questions should be spaced adequately. Liberal spacing acts as a stimulus for the investigator to write freely. A larger schedule should be divided into a number of sections whereby each section is related to different aspects of the research problem. Sometimes, the schedule is kept anonymous — the name and address of the respondent is not written on the form. This facilitates free and frank response from the respondent on some sensitive or delicate issues since the person concerned need not worry about the consequences of his views.

The physical layout of a schedule has to be planned. The name of the person who has filled the schedule, date and place of filling the schedule, date of scrutiny of entries and name of the person who has scrutinised the filled up schedule should be given at the end. It is not necessary to get the signature of the respondent as it produces negative results. However, the name and address of the respondent has to be entered in case a check becomes necessary. Confidentiality of the information provided must be ensured. Space should be left for any comments/observations of the investigator.

The framing of a schedule is facilitated if there has been pilot study as this often provides the logical basis upon which the schedule for the main study can be constructed. Such a study may also help to determine whether there is need for a highly structured set of questions or whether the questions should remain relatively open ended. A pilot study not only helps to save time and money but also puts the main study on a sounder footing. However, pilot studies are usually not done when the surveys or research studies are small.

In order to enhance the quality of a schedule a 'pretest' should be conducted. In the 'pretest' every part of the procedure for administration of the schedule is laid out in the same manner as intended for the final study. The schedule is given the final form complete with covering letter, instructions etc.. Purposive selection of sample respondents is preferable to have feedback on as wide a range of responses as possible. The data collected from pretest should be carefully tabulated. The scrutiny of responses may reveal shortcomings or ambiguities in the questions. It may point out that some vital information is lacking in the schedule. It may indicate the need to reformulate or add a question which may clarify a number of unexplained facts. The pretest also helps in arriving at the full range of responses for coding from which the respondent is asked to choose one or more alternatives.

Activity III
Suppose you wish to study the extent and causes of rural indebtedness. Before you actually frame the questions, do the preparatory work mentioned in the text. In the set of questions, have a mix of open ended and closed ended ones. Do the pretesting on five respondents. Then modify the schedule in the light of the feedback you get. Discuss your schedule with your counsellor.

2.7 MAILED QUESTIONNAIRE

In the previous section you have been given guidelines on the construction of a schedule. When the set of questions is mailed to the respondents or distributed to them for being filled by them without the help of an investigator or researcher, it is called a mailed questionnaire. In other words, the mailed questionnaire is self-administered. The fact that it is self-administered makes it incumbent that some precautions are taken. The questionnaire to be mailed should:

- have a covering letter appealing to the respondent to fill up the questionnaire and return the same. The letter should explain the nature and the purpose of the study, how the respondent has been selected and why his cooperation is necessary.
- give clear instructions to the respondent on how to fill the questionnaire. It is very desirable to pretest these instructions on respondents of the same educational level as would be ultimately of those filling the questionnaire. The questionnaire should have a proper layout and easy to read print. Some extra space may be given at the end for any observation which the respondent wishes to make.
- A self addressed and duly stamped envelope should be enclosed for returning the questionnaire so that it incures no cost to the respondent.

In certain situations, the mailed questionnaire is a useful tool; it is so when:

- the respondents are spread over a large geographical area and live in difficult to reach places;
- the type of information needed is relatively simpler and is limited to relatively fewer items of information;
- not much depth of information is aimed at and the subject is not complex;
- the respondents are a rather select group, likely to be motivated enough and educated so as to be able to and interested in filling the questionnaire and mailing it back.

An absolutely essential requirement for the use of mailed questionnaire relates to the maintaining and keeping of records of questionnaires mailed and received back. The researcher must record dates of transmission and receipt of questionnaires, scrutinise the forms properly and make certain that all questions are answered adequately. A flowchart or control sheet should be maintained for checking each questionnaire as it goes through the different stages.

Strengths and Weaknesses

The main advantage of the mailed questionnaire method is that it is less expensive and more standardised than the interview method. It requires much less skill to administer such a questionnaire once it has been prepared. Since the responses are usually pre-coded, tabulation is easy. The questionnaires can either be mailed or simply handed over to the respondents with a minimum of explanation. They can also be administered to a large number of individuals simultaneously as, for instance, when they come for a conference. Given a particular amount of funds available for a project, it is possible to cover a wider geographical area or larger sections of a population.

Another advantage stems from the impersonal character of the questionnaire. The respondent has greater confidence in his anonymity and feels more free to express his views. If the questionnaire is presented as anonymous without any identifying information, the respondent feels assured that his replies will not be associated with his name. Thus, he need not fear that his views will invite trouble in future. Finally, the respondent can fill the questionnaire according to his own convenience and liberty. Since the method places no immediate pressure on the person for quick response, he can consider each question carefully before answering it.

The main limitation is that the response rate is usually low unless the respondents are highly motivated. The nature of bias introduced as a result of this is likely to affect the representative character of the sample. The mailed questionnaire is not suitable for the study of complex problems. The tool does not also permit supplementation or clarification. Nor can it be used if the population is illiterate or have a very low level of literacy.

Check Your Porgress III

Notes : a) Following statements are based on text that you have already read. State whether they are true or false by putting a tick-mark ✓ in the relevant box.

b) Check your answer with the model answer given at the end of this Unit.

	True	False
a) Schedule can be used for the collection of qualitative as well as quantitative data	[ ]	[ ]
b) Schedule is self-administered and does not require the medium of research-investigator.	[ ]	[ ]
c) Each section of schedule should contain questions relating to the same general theme.	[ ]	[ ]
d) A pretest of schedule enhances the quality of a schedule.	[ ]	[ ]
e) Mailed questionnaire is not self-administered.	[ ]	[ ]
f) Clear instructions for answering the questions must accompany a mailed questionnaire	[ ]	[ ]
g) Mailed questionnaire is used to collect data from educated respondents.	[ ]	[ ]
h) Response rate in the case of Mailed questionnaire is usually low	[ ]	[ ]

2.8 LET US SUM UP

The choice of tools to be used in the collection of data is a crucial decision. Broadly speaking, the nature of the research problem, the type of research design, the type of

respondents and their geographical distribution, the professional competence of investigators available for collecting data, the period within which the study is to be completed and the funds available determine the tools to be selected for the collection of data.

We considered the features of each of the four tools most frequently used for data collection, viz., Observation, Interview, Schedule and the Mailed Questionnaire. We discussed their respective strengths and weaknesses. We found that even though the schedule is the most widely used tool for data collection, a judicious use of other tools to supplement the data collected through the schedule enriches the quality of the information that is obtained. In some types of research studies, however, the primary dependence has to be on one or the other methods of data collection. The use of tools requires much expertise which can be acquired not just by reading books but by actual experience acquired through the use of these tools.

2.9 KEY WORDS

Coding: The categorizing and sorting of raw data into groups, such as responses to open ended questions or field data.

Pretest: A trial run of an experiment or survey tool with a small number of preliminary subjects or respondents to evaluate and rehearse the study procedures and personnel.

Rapport: A relationship in which the respondent identifies himself with the purpose of the study and extends full and willing cooperation for furnishing the information.

Reliability: The stability or consistency of an operational definition.

Scale: A composite measure of a concept constructed by continuing separate indicators according to procedures designed to ensure unidimensionality or other desirable qualities.

Validity: The congruence or 'goodness of fit' between an operational definition and the concept it is purported to measure. In logical analysis validity pertains to the adequacy of reasoning (premises must be properly related to the conclusion so that the argument is logically correct) as opposed to the truth of the premises.

2.10 SUGGESTED READINGS

De Vaus, D.A. 1986 — *Surveys in Social Research,* George Allen and Unwin: London.

Goode, W.J. and Hatt, P.K. 1981 — *Methods in Social Science,* McGraw-Hill: London.

Madge, J. 1965 — *The Tools of Social Science,* Longmans: London.

Moser, C.A. and Kalton, G. 1973 — *Survey Methods in Social Investigation,* Heinemann Educational Books: London.

2.11 MODEL ANSWERS

Check Your Progress III

a) True b) False c) True d) True e) False f) True g) True h) True.

UNIT 3 ANALYSIS OF DATA AND THEIR PRESENTATION

Contents

3.0 Aims and Objectives
3.1 Introduction
3.2 Meaning of Data Analysis
3.3 Elementary Analysis of Data
3.4 Coding and Tabulation
3.5 Statistical Analysis of Data
3.6 Presentation of Data
3.7 Let Us Sum Up
3.8 Key Words
3.9 Suggested Readings
3.10 Model Answers

3.0 AIMS AND OBJECTIVES

The aim of this unit is to acquaint you with the basic aspects of analysis of data and their presentation. Analysis of data in social sciences is not a simple exercise rather it involves complex and cumbersome techniques and procedures. Therefore when you actually take up a research project you may require further readings in this area. However after having learned this unit you should be able to:

- describe the fundamental aspects pertaining to the analysis of data
- identify the major steps involved in the processing and interpretation of data
- describe some of the important characteristics of tabulation and statistical analysis of data
- state some of the basic elements related to presentation of data.

3.1 INTRODUCTION

In Unit 1 of this block, we have discussed some of the important aspects concerning the designing of a research study. Unit 2 has taken care of the nature and selection of research tools and their application in the process of data collection. By now you know that preparation of an adequate research design requires a lot of advance thinking and planning. The execution of a research design with the help of appropriate research tools culminates into the collection of data. Significant meanings can be drawn through the analysis and interpretation of such data.

The data collected do not automatically provide answers to research questions. It is the systematic analysis of data that helps to build an intellectual edifice in which properly sorted and sifted observations and facts locate their true meaning. In fact, the function of analysis is to summarize the data in such a manner, that they yield answers to research questions. The process of data collection also helps to reveal the broader significance of a research study by linking it to other available knowledge. Hence, in order to draw meaningful conclusions from a research venture, the data are needed to be analysed and presented in an appropriate manner.

3.2 MEANING OF DATA ANALYSIS

Scientific knowledge is a means to understand the empirical world. Its progress involves a repetitive interplay between theoretical ideas and empirical evidence. In most of the research, theoretical formulations guide the process of data collection. Data and facts, in turn, help to generate tentative theories. Deductions from a

tentative theory are compared with new facts. The theory is upheld, modified or discarded in the light of such facts. This constant interplay between theory and data is a continuous process that helps in the progress of scientific knowledge.

Analysis and interpretation of data form a vital link between theory and empirical evidence. Data analysis takes place whenever theory and data are compared. This comparison begins in field research when an investigator struggles to bring order to or to make sense out of his observations. In surveys and experiments, the researcher typically begins to bring theory and data together when testing the validity of a hypothesis. On the whole, thus, demand for this continued interaction between theory and data sets the stage for data-analysis.

Data analysis involves processing and interpretation of data. Processing of data involves the transformation of categories into coding schemes that are amenable to quantitative treatment. It demands organization, reorganization and understanding of data so as to reach certain conclusions in the light of the theoretical concerns of the study. Interpretation of data means that certain inferences are drawn from the study and the findings of research are linked to the already existing state of knowledge.

The problems raised by analysis of data are directly related to the complexity of the hypothesis in studies where the purpose is hypothesis-testing. Under ideal conditions of precision and simplicity, data analysis presents very few problems. In such a situation, the statements of the hypothesis or research questions automatically provide for an adequate basis for the analysis of data. For instance, in an experimental study, the researcher is supposed to collect such data as fall within the experimental design. In other words, to depict a causal relationship the focus of data collection is usually on the variables involved in the study. Once the data about the independent and dependent variables become available, the process of data analysis turns out to be more or less, an operation which follows a pre-determined path.

Thus, when relevant factors are known in advance, no serious problems are encountered in data analysis. However, the occurrence of such an ideal situation is relatively infrequent in social research. Many a times, the situation demands that data analysis must cope with diverse factors not anticipated in the original research design. In such studies, the process of data analysis is cumbersome and the researcher must have thorough knowledge about the different steps involved in it.

Now, if data are analysed to answer specific questions, it is termed as primary analysis. On the other hand, if data collected for a particular purpose are later on examined in terms of some new objective or for resolving another problem, then it is called secondary analysis. When the object of data-analysis is to determine the causal relationship between two variables, this is done through bivariate analysis. However, when the relationship among more than two variables is investigated, the statistical tool of multivariate analysis is used.

Two aspects need special attention in the context of data analysis. These are measurement and evaluation. Measurement involves: 1) identification of the quality or attribute that is to be measured; 2) determination of a set of operations by which the attribute may be made manifest and perceivable; 3) establishment of certain procedures for translating observations into quantitative statements of some degree or amount. Evaluation concerns itself with the nature and quality of measurement and the determination of their reliability and validity.

Now, in everyday life, we all are engaged in measurement at various levels. We generally make statements like the soil is very fertile, the road is extremely bad, the room is very big and so on. While making such statements we are measuring some attribute or quality. What we actually do in this process is that we assign labels 'very fertile', 'extremely bad', 'very big' to objects such as soil, road, room so as to represent their conceptual properties. Thus, whenever we rate something, we are engaged in measurement.

There is difference, of course, between such every day examples of measurement and the process of measurement in social research. In the former cases, the rules for assigning labels or numbers to objects are more or less intuitive whereas in social

sciences these rules must be spelled out in detail. Scientific norms require that we specify the guidelines and follow procedures to obtain a measurement so that others can repeat our observations and judge the quality of information yielded by our measurement procedures.

In social sciences, the measurement process consists of moving from the abstract to the concrete. This process begins as soon as the researcher formulates the research problem. Every research problem or hypothesis contains concepts or variables that refer to aspects of reality in which the researcher is interested. Thus at the time of formulation of questions the researcher must think : which terms are to be used and what they mean in both an abstract and an empirical sense. In fact, the ultimate goal of measurement is to specify clearly the observable referents of terms contained in a research hypothesis.

The process of measurement can be broken into three steps—conceptualization, specification of variables and indicators, and operationalization. The development and clarification of concepts is called conceptualization. We can identify and observe the actual existence of a concept only if its meaning is made clear. The initial step is to clarify the mental construct conveyed by the concepts. In other words, we must arrive at precise definitions of concepts. Suppose you are interested in testing the hypothesis that "education reduces poverty", then you would begin by defining the meaning of 'education' and 'poverty'. The usual practice recommended for the beginning researcher is to rely on existing definitions in social science literature and carefully consider which one is suitable for his work. Investigators, however, sometimes formulate their own definitions, especially when disagreement exists in the literature over the exact meaning of the concept.

Measurement assumes the possibility of assigning values to concepts. But we measure concepts that vary, which we refer to as variables. Now most concepts in social sciences are not directly observable. For instance, we cannot see education in the same sense as we can see a table or fan. While we cannot see education, we can observe what knowledgeable people are like and how much formal schooling they have. Thus after conceptualization, the important aspect is to identify the varying manifestations of concepts in real life situations. It is at this point that we move from the language of concepts to the language of variables.

Researchers often use the terms 'concept' and 'variable' interchangeably. But these terms connote different levels of abstraction and different stages in the measurement process. While a concept is a 'mental construct', a variable usually signifies some observable event that represents the underlying concept. Thus, we must specify in which sense a given concept may vary or assume different sense or meaning in observable terms. For instance, the concept of education can have varying manifestations in real social life. Education can be in the field of medicine, engineering, social sciences and so on. For purposes of measurement, it is of paramount importance that we become fully aware of the variable character of concepts and specify their meaning and sense as employed in a study.

Measurement operations aim to estimate the existing values of variables. But to achieve this objective the main problem arises as to how to assign different values, numbers or labels to the variables. This problem is resolved by the operationalization of concepts or variables. The detailed description of the research operations or procedures necessary to assign various units (e.g. persons) to the categories of variables is called an operational definition. We must choose or develop such operational definitions which correspond reasonably close to the concept or variable in question. The gap between the concept (mental contruct) and the observable reality represented by an operational definition should be the minimum possible.

For example, we are interested in determining the 'educational status' of the people belonging to a particular area. After taking care of research design, sampling etc., we will investigate how many persons have acquired formal education. We will further investigate how many years they have spent at school, college etc. and what type of certificates, diplomas or degrees they have obtained. Thus by linking the concept of education with 'years of schooling' or degrees, we can know how many people have studied upto primary, middle, higher secondary, Bachelor's degree, Master's degree

and so on. We can further calculate the percentage in each category and can produce a quantitative statement about the 'educational status' of the population under study. Similarly, the concept of 'poverty' can be operationalized by linking it with 'income'. A particular household monthly income (say less than Rs 500/-) can be taken as representing the presence of poverty. Thus, all households whose monthly income falls below Rs 500 will be considered as poor.

No matter how simple or how elaborate an investigators' formal definitions of the concepts, he must find some way of translating them into observable events so as to carry out the process of measurement in quantitative terms. The investigator must devise some operations that will produce data as an indicator of the concept. Only then the measurement and consequently the analysis of data is possible. This requires considerable ingenuity especially if the theoretical constructs are far removed from everyday events and if little research using such constructs has been done in the past.

Operational definitions are adequate if they help in the gathering of such data as adequately represent the observable manifestations of the concept. However, to which extent this aim is achieved generally remains a matter of judgement. It must be borne in mind that operational definitions so essential to social research are necessarily somewhat arbitrary and restricted expressions of what a concept really means.

Now, for any concept a large number of operational definitions are possible. Creative insight, good judgement and relevant theory aid in the development of operational definitions. However, such aids are to some extent subjective in nature. For instance, an investigator may feel that his data provide reasonably good indicators of his concepts, but a critic of the study may feel they do not. It frequently happens that even a researcher himself becomes aware of the fact that his data constitute only a limited reflection of the concepts employed. In this context, an important question arises: how to evaluate the soundness of specific operational definitions in an objective manner?

Evaluation basically concerns with the nature and quality of measurements. Measurement in data analysis is greatly influenced by the quality of operational definition. Social scientists use the terms 'reliability' and 'validity' to describe issues involved in evaluating the quality of operational definitions. Reliability is concerned with questions of stability and consistency. Is the operational definition measuring an attribute or 'quality' consistent and dependable? Do repeated applications of the operational definition under similar conditions yield consistent results?

For example, a steel tape is a highly reliable measuring instrument. A piece of wood (say 20 cm long) will measure 20 cm every time a measurement is taken with a steel tape. But the measurement may vary if a cloth tape is used. Elasticity of the cloth used may affect the length of cloth tape. Variations may occur depending on how loosely or tightly the tape is stretched. In other words the cloth tape is less reliable as compared to the steel tape.

Validity in measurement refers to the extent of matching, congruence or 'goodness of fit' between an operational definition and the concept it is purported to measure. In other words, an operational definition must truly reflect what the concept means. The researcher must ensure that he is measuring something which he intends to measure. For instance, the concept of 'education' may be operationalized by linking it with the years of schooling. Now if the educational standards — syllabi, teaching methods, examination system etc. are uniform in different schools, then 'years of schooling' may be considered as a valid measure. Twelve years of schooling and the attendant percentage of marks in the school leaving certificate will reflect the performance and knowledge of the students. However, if the educational standards vary from region to region then years of schooling cannot be considered as a valid measure. The procedures for assessing validity and reliability are complex and cumbersome in social sciences. Many research investigators prefer to borrow from prior studies measures that have established records of validity and reliability.

Check Your Progress I

Indicate whether the following statements are true or false. Insert a tick mark (✓) in the relevant box.

	True	False
1) The continuous interplay between theory and data enhances the growth of scientific knowledge.	[ ]	[ ]
2) Data analysis forms a vital link between theory and empirical evidence	[ ]	[ ]
3) Social facts are usually simple and pose no difficulties in data analysis	[ ]	[ ]
4) Processing and interpretation of data are the two most important dimensions of data analysis.	[ ]	[ ]
5) Two aspects, viz.measurement and evaluation require special attention in the context of data analysis.	[ ]	[ ]
6) Measurement process means a movement from concrete reality to abstract ideas.	[ ]	[ ]
7) Measurement process involves conceptualization, specification of variables and indicators, and operationalization of concepts.	[ ]	[ ]
8) Operational definitions are not arbitrary expressions of concept.	[ ]	[ ]
9) Evaluation determines the nature and quality of measurement.	[ ]	[ ]
10) The procedure for assessing reliability and validity are very simple in social research.	[ ]	[ ]

3.3 ELEMENTARY ANALYSIS OF DATA

Data analysis is a fundamental requirement of social research. Its main function is to impart order to the whole body of information gathered during the course of study. This demands a systematic treatment of the data collected. In other words, it means grouping, regrouping, sorting, or sifting of data keeping in mind the overall needs of the research design.

The ultimate aim of data-analysis is to provide answers to the already formulated research questions. The information is systematically organized to perform this function. Particular type of information is grouped together. Data are presented in such a way that relationship among different types of information becomes clear and data become amenable to interpretation. This is then organised systematically to take the form of report, book or article which can be shared by others.

We will now discuss some of the major steps involved in the analysis of data.

3.3.1 Scrutiny of the Assembled Data

An early but important step in data analysis is a critical examination of assembled material. Research should make all efforts to search the contrasts or similarities contained in the data. Adequate attention should be given to all the little and big events which symbolize the field situation so as to reveal their comparative significance. While conducting this preliminary exercise, the researcher should keep in mind the key elements of the study, viz., nature of the research problem, hypothesis and research design.

In fact, the researcher needs to cultivate the habit of asking questions to himself, once he confronts the data. For instance, are the data complete enough to reveal patterns of behaviour, sequences, and relationships which might explain the research problem under investigation? Are the data objectively recorded and are they reproducible? Are the data susceptible to quantitative treatment? Are the data collected with the help of adequate sampling techniques? Many more such questions may be raised while scrutinizing the data.

Most of the questions should relate to the research project and the data collected. But the researcher should not refrain from asking even those questions which may otherwise look unimportant in the beginning. This procedure invigorates imagination and opens up new patterns for the resolution of problems. It also helps to develop a critical attitude which is most essential to know the real depths of data.

Data analysis requires total involvement in data. The researcher should make all out efforts to become aware of each and every facet of data. It is just like putting an object in front of one's eyes and closely looking at it until one becomes thoroughly acquainted with all its essential features. Reading, re-reading, examining and re-examining of data helps to reveal the true nature of data. This process also helps to eliminate a host of problems from data. The whole exercise of scrutinizing data accomplishes several purposes in the sense that the researcher:

- gets acquainted with the intrinsic complexity of data to be analysed;
- is able to perceive the essential relationships, similarities and differences contained in the data;
- is able to know about the internal consistency and completeness of various aspects of data;
- learns about the relative significance of each and every bit of recorded information.

3.3.2 Classification of Data

By now, it is obvious that data in social research are constituted through the collection of accurate social facts. But social facts become meaningful only when they are logically connected with other relevant facts and are sorted according to their essential nature. Similar type of facts must form a chain of evidence so that they are able to mutually explain each other. This process which is aimed at the orderly compilation of facts is termed as classification of data.

It should not be assumed that classification is undertaken only after the data have been collected. The careful researcher is always a classifier and organiser of data. As he observes and gathers his data, he inevitably adopts some scheme (howsoever rudimentary it may be) for classifying and coordinating social facts. For instance, the initial selection of variables (dependent, independent variable etc.) provides a broader scheme for classification even during the course of data collection.

Major work regarding classification is carried out only after the completion of data collection. To achieve this the researcher needs to select, adopt or develop some classification system. The suitability of this is determined by the nature of the study (hypothesis testing, diagnostic study etc.) and the completeness or accuracy of data. The classification further depends upon the researcher's insight into data and the extent of sophistication attained by him in the concerned discipline.

The main value of any classification system lies in its potential for grouping together masses of comparable data into relatively few classes or categories. For instance, a given bulk of information which may otherwise display a diverse range of characteristics to a casual observer, can be grouped under a suitably adopted or evolved category. This way, a decrease in the number of units, (classes, categories etc.) makes the research material more readily comprehensible. The essence of the data is also more easily grasped.

In addition to some common sense classification the researcher may begin to sort his data in terms of:

- essential similarities and dissimilarities in them;
- clusters of related facts which can be observed repeatedly with consistent regularity; and
- recurring sequence of events.

It should be clearly borne in mind that the aim of classification is to discover such series, sequences and relationships as may throw light on the uniformities in one group of data and on the differences in another.

3.3.3 Establishment of Categories

You have learned that data analysis involves classification of research materials. For the classification of data, a fundamental requirement is the establishment of categories. To reiterate, the data in social research are usually constituted by a variety of information gathered during the course of study. In order to provide answers to research questions, the loose data need to be organised and classified. This means that diverse kinds of information (field observations, responses through questionnaires/schedules and interviews) need to be grouped into a limited number of classes or categories. Similar types of responses or observations must be grouped together to form a particular category. But how to select, adopt or develop a category or categories?

It is important to note that decisions regarding the selection or adoption of categories cannot be arrived at in an arbitrary fashion. In order to decide about the relevant categories under which research data may be classified, some principle of classification must be adopted or evolved. Now, when a study is conducted within a particular theoretical framework, the categories usually get established prior to data collection. Furthermore, if structured tools of data collection are employed, then the appropriate principles required for classification of data are clearly prescribed by the nature of questions and their attendant responses. Firstly, in such cases, only those data are collected which are relevant to the hypothesis or the variables involved in the study. Secondly, structured questions themselves provide clear cut options for their answers. For instance, in order to determine the demographic profile of scheduled castes in a rural area, the researcher may ask the question. "Do you belong to a Scheduled Caste?" Now, the answer can be of yes/no type. All respondents will get automatically categorised as belonging to Scheduled Castes or not belonging to Scheduled Castes. Sometimes, a number of choices are given for answering a question. Barring some exceptions, almost all the responses usually fall under one or the other choice. This way the problem of developing categories becomes simplified. The information gets more or less precategorized during the process of data collection. Thus, in a study conducted within a particular theoretical premise and with the help of structured tools (schedule) each question provides for a natural unit for categorization.

When the researcher approaches the field situation without the help of existing theories the problem of developing categories becomes cumbersome. In such cases, the researcher usually makes efforts to move from data towards theory. His aim is to gather each and every bit of information during fieldwork which may later on be used to produce a theoretical statement or generate some hypotheses. Categories are mainly developed only after completing the process of data collection. The researcher may continue to sort or sift his data even during the course of data collection under tentative categories. But the main task of establishing categories is taken up only at the time of data analysis.

For instance, in exploratory studies, establishment of principles of classification is especially difficult. It is absolutely impossible to select categories in advance. Such studies by definition do not start with explicit hypotheses. At the time of data collection, the investigator is generally not clear which information might prove useful. He simply collects each and every bit of information about the community under study. The 'wording' and 'sequence' of questions usually varies from person to person. Consequently, the responses also do not display a uniform pattern. The data in such studies turn out to be complex. There are usually no convenient 'natural units' (as in the case of structured questions) under which data may be classified. Hence, the researcher faces crucial problems while selecting or developing categories at the time of data analysis. The researcher, in such a situation, is supposed to deal with a diverse range of unstructured information. But a substantial part of information may turn out to be irrelevant for the study. Therefore, the foremost requirement is to identify that part of data which is relevant to the research investigation. After having done this, the researcher should then move on to think about classificatory principles needed for the establishment of categories.

In an exploratory study, the initial concern should be to group similar information under some suitable heading. The process can even be started during the time of data collection. For example, all information regarding 'family structure' can be grouped

together. More observations can be grouped together under the headings of 'kinship structure', 'marriage' and so on. Similarly, observations relating to 'political institutions' and 'economic structure' can be grouped separately.

Now, if you as a researcher are interested in some particular aspect of a community, you need to develop particular categories for the classification of data. Suppose you are interested to study the problem of alcoholism among landless labour in a rural area. You may categorize your data as follows:

Alcoholism among landless labour; Present.... Absent.... In this classification, you have used only two categories, i.e., presence or absence. These two categories form a category set. However, a category set may also consist of more than two categories. Assuming that you wish to study degree of alcoholism among the workers then you may decide to use six categories: very high, high, medium, low, very low and absent. Each category set (irrespective of the number of categories involved) is subject to certain basic rules. It must be:

- derived from a single classification principle;
- sufficiently exhaustive to include all possible responses related to a certain item in one of the categories within a set;
- mutually exclusive i.e. responses must fit into one and only one category within the set;
- clearly defined, so that it is understood alike by all the investigators who would code the data.

Categorization of data is not as simple as might be inferred from the above discussion. In actual data analysis, establishment of categories requires considerable research experience, thorough knowledge in the concerned discipline and of the problem being studied.

Check Your Progress II

Note: a) Use the space provided below for your answer.

1) How does scrutiny of data help a researcher?
(Hint: see the text.)

.....

.....

.....

.....

2) Which principles need to be followed in establishing categories for the classification of data?
(Hint: see the text)

.....

.....

.....

.....

3.4 CODING AND TABULATION

In the following we will acquaint you with basic aspects of coding and tabulation.

3.4.1 Coding

Coding is an operation by which numbers or symbols are assigned to all items, according to the class or category in which they fall. The basic purpose is to facilitate greatly the task of tabulation of the responses. Assigning the symbol or number to a given datum is more or less a mechanical procedure. The basic operation is mainly that of classification. For instance, in surveys, the categories are usually answers to questions; and the common practice which simplifies data entry and analysis is to use numerical codes only.

Now if the answers to questions are already numbered, then there is no need to code the data further. But if the answers are not expressed numerically, then numbers must be assigned to each answer. For a moment, suppose that the variety of answers to a particular question is 'yes', 'no', 'don't know', 'no answer' respectively. The codes would be 1 for 'yes' 2 for 'no', 3 for 'don't know', and 4 for 'no answer'.

For structured questions coding is a straightforward operation. Generally, there are relatively few categories and you simply need to assign a different code to each category. However, for unstructured questions the number of responses may be large. In such a case you need to develop a coding scheme that does not require a separate code for each and every response. A particular code should adequately cover a full range of similar responses. The whole idea is to put data in manageable form.

The nature of coding operations depends mainly on three factors:
- the number of respondents or sources of data;
- the number of questions asked;
- the number and complexity of statistical operations planned for the study.

If the number of respondents is large then any kind of tabulation becomes difficult unless the data are coded. There is no easy way to carry out complex cross tabulation without some form of coding. Furthermore, a statistical operation involves manipulation of numbers, which in turn must represent the data from the schedule. The researcher should assign numbers to the answers and summarize them separately on sheets or cards. The operations become much simpler. The more complex the data, the more useful is some form of coding.

Depending upon the nature of study, coding can be done either at the stage of finalising the schedule, or at the data collection stage (where schedule/mailed questionnaire is not being used) or at a time just prior to tabulation of data. In a study employing closed ended questions, the numerals may be assigned to answers in advance. This operation is called precoding. For example take the question:

Do you belong to a Scheduled Caste?
(Circle the answer)

Yes 1
No 2

This way the data get automatically coded. These precoded responses can be tabulated very easily by hand or they may be transferred directly on cards for machine tabulation. Another procedure is that the researcher may do the coding work as soon as he hears answers. But it requires considerable research experience and should be done in a fairly cautious manner.

Precoding a schedule has some obvious advantages. For instance, it makes possible elimination of all or major part of the work required during office coding operation. However, one of the possible disadvantages is that categories selected (consequently coded) may prove to be too general for revealing inconsistencies in reporting. In fact, the tendency among inexperienced researchers is to use a few categories which generally gloss over potentially meaningful differences. Another limitation of precoding may stem from 'catch-all' categories like 'Others' or 'Miscellaneous' as one of the possible response to a question. Such categories may conceal meaningful information unless explained by the interviewer. If the number of responses in the 'Others' category is unusually large it implies that much useful information is being lost. Once data are coded and data analysis is under way, it is still possible to combine code categories for purpose of analysis; but the reverse operation cannot be done.

When coding is carried out after data collection (just before tabulation) coding operation in a central office may be organised. In such a situation, much depends upon the nature of data and the number of respondents. It is useful to handle relatively simple factual items separately. The more complicated and special information may be grouped into items subject-wise. The most important factor in any coding scheme is that it must cover each and every variable or question involved in the study.

There are some useful aids to coding. A heavy coloured pencil can be used to mark code symbols opposite the answers. If the number of respondents is small, then

summary sheets may be attached to each schedule. A large summary sheet can also be designed so as to take care of all the cases. This is quite useful for quick overview of data. If the number of cases goes beyond 150, the student may use some form of punch-cards. The data transferred to punch-cards facilitates analysis particularly when mechanical aids are employed.

While carrying out coding of schedules, instructions should be prepared in a written form. Such instructions should be followed during classification. Now, it is possible to make errors no matter how simple the problems of classification and coding may be. Therefore, it is advisable to make a definite sample of the coded data so as to detect errors. The procedure is called 'spot checking' and it acts as a good precautionary measure against errors in the mechanics of coding.

Once codes are established for all the items and the data are ready for tabulation, the researcher should prepare a code-book. A code-book is like a dictionary in the sense that it defines the meaning of the numerical codes used for each answer. The code-book should contain coding rules employed in the organization of data. A well-prepared code-book helps to locate each item in the data-file.

3.4.2 Tabulation

After having completed the process of classification and coding, the next step involved in the analysis is the tabulation of data. Tabulation is mainly a part of the technical process in the analysis of data. The essential operation in tabulation is counting so as to determine the number of cases that fall into various categories. In other words, the individual cases or items in the assembled data are sorted and counted according to various categories of classification. The whole aim of tabulation is to arrange research material into some kind of concise and logical order.

Tabulation may be done entirely by hand or it may be done by machine. In other words, sorting and counting process can be carried out by hand or with the help of mechanical aids depending upon the number of cases or items to be tabulated as well as upon the availability of staff and money. If the number of cases is small, the tabulating work can be completed by hand. In case the data has been transferred to cards or small sheets, they can be sorted into piles and counted directly. This system was generally used for large scale tabulation before the advent of mechanical tabulation.

A common method in hand tabulation is to enter each case in the appropriate compartment of a tally-sheet by means of a line, dot or some other mark. The tally sheet is a blank table with properly labelled columns and rows. Figure 3.1 below is an example of a tally-sheet for a frequency distribution. It shows the distribution of 65 executives according to their weight.

Table 3.1: Weight of executives

Weight (kg.)	Tabulation	Frequency
45-49	1	1
50-54	1111	4
55-59	𝍸 1111	9
60-64	𝍸 𝍸 𝍸 𝍸 11	22
65-69	𝍸 𝍸 𝍸 1	16
70-74	𝍸 𝍸	10
75-79	111	3
	Total	65

Four tallymarks which have been crossed denote a frequency count of 5. This way counting becomes simpler. In this case, the information from schedules in regard to the question on weight of the respondent has been tabulated. There are 65 schedules. The weight of each person could have ranged from 45 kg. to 79 kg. or a figure inbetween. If the data were recorded in kg. (i.e. decimal points are ignored), one

possibility could have been to tabulate the weight by 1 kg. interval. That would have made the table very long and not served any useful purpose from the point of tabulation. Similarly, class intervals of 15 (i.e. 45-59, 60-74, 75 and above) or 10 (i.e. 45-54, 55-64, 65-74, 75 and above) would be too large. Therefore, tabulation of the data by 5 kg intervals is far more satisfactory in this case. The coding plan would have recognised this and the codes put against each entry accordingly.

The term cross tabulation is employed to refer to the tabulation of the number of cases that occur jointly in two or more categories. Cross tabulation is an essential step in the discovery or testing of relationships among the variables in one's data. A cross tabulation requires a table with rows representing the categories of one variable and columns representing the categories of another. When a dependent variable can be identified, it is customary to make this the row variable and to treat the independent variable as the column variable, although this is only an arbitrary convention.

Let us consider the cross tabulation of two variables, i.e., 'attitude towards dowry' and 'gender'. In table below the row variable consists of 'Attitude towards dowry system'. In other words, it indicates whether the respondents favour or do not favour dowry system. The column variable is 'gender'. Gender is here the independent variable. Notice that the last column and bottom row, each labelled as 'Total' show the total number of respondents with each single characteristic. Because the four numbers (112, 88, 109 and 91) are along the right and bottom margins of the table, they are called **marginal frequencies.**

Table 3.2: Attitude towards dowry by gender

Attitude towards dowry system	Gender		Total
	Male	Female	
Favour	78	34	112
Do not favour	31	57	88
Total	109	91	200

Each intersection in the body of the table where the categories of two variables intersect is called a **cell.** The number in each cell is called the **cell frequency.** The cell frequencies in this table are 78, 34, 31 and 57. For example, 78 men in favour of the dowry system constitute one of the cell frequencies. Because this table has two rows and two columns, it is referred to as a 2×2 table.

Both machine and hand tabulation presuppose that data have been coded and that coding has been checked. Now hand tabulation can be quite fast and accurate if proper techniques are employed. One of the most efficient techniques is to use small code cards which can be easily sorted and counted. Further, with the use of such devices as colours, heavy lines etc., the codes can be easily distinguished and the cards efficiently sorted. Counting is also rapid if small cards are used. On the whole, manual tabulation is less expensive and less time consuming when data and category sets to be counted are small and also when not too many cross tabulations are involved.

In machine tabulation, scores for each case are usually transcribed on to a sheet from which these can be punched on to cards by a card-puncher. In fact, machine tabulation involves operations like card punching, the checking of machine tabulations, the transposition of the results from machine tabulation forms to tables etc. All these steps are not required in hand tabulation. If the data is large or many cross tabulations are involved, then the efficiency of machines compensates for the time consumed in the preparatory operations. Many types of machines are available which can be used for tabulation. Developments in this field have been extremely rapid in recent years. Some machines simply sort and count cards; others sort, count and print the results; still others are capable of performing the most complicated statistical operations. The machines of latter variety are extremely complex and they need to be "programmed" for a given operation by a specialist in the field. The cost of using machine tabulation is usually high. But as the number of cases or the number of cross-tabulations increases, the use of machine tabulators becomes progressively

more economical. The number of cross-tabulations is the most important factor in determining the relative preference for one rather than the other mechanical device in tabulation. With the advent of the computer, tabulation has become a quick exercise. The selection of appropriate software permits not only the data to be tabulated but also their statistical analysis. For instance, computation of the correlations among a large number of variables is a statistical job that can now be accomplished within a very short time.

Check You Progress III

Note: a) Use the space provided below for your answer.

1) What are the main factors which affect the nature of coding operations? (Hint: see the text.)

.....

.....

.....

.....

2) What is the main aim of tabulation of data? (Hint: see the text.)

.....

.....

.....

.....

Activity I

Frame a question to be asked from the head of a household whose child below 10 years age has dropped from school. Ask the question from some 25 respondents. Try the categorisation,coding and tabulation of this data. Show this to your academic counsellor.

3.5 STATISTICAL ANALYSIS OF DATA

In the previous sections, we have discussed some of the important steps involved in the analysis of data. In this section, we will give you some preliminary idea about the statistical analysis of data. We will also acquaint you with the elementary aspects involved in the presentation of data. Towards the end we will furnish a brief outline for the preparation of a research report.

After having taken care of coding, tabulation etc., the next crucial decision pertains to appropriate statistical analysis of data. But how to decide about the relevant kind of analysis? Obviously, such a decision depends on what you want to know. Let us suppose that you want to know about the causal relationship between independent and dependent variables involved in your study. Now in order to establish such a causal relationship, the foremost requirement would be to demonstrate that both the variables are associated. In other words, you must demonstrate that changes in one variable are accompanied by changes in the other.

At this stage you might be thinking that one may simply examine the raw data, respondent by respondent, to reveal whether changes in the independent variable are associated with changes in the dependent variable. But such an exercise is extremely cumbersome. When you have to examine a large number of respondents the task becomes incredibly tedious and also not very reliable. Many a times, it is not at all feasible or possible to carry out such an operation. Therefore, what the researcher actually needs is a fast and efficient means of summarizing the association between variables for the entire sample of respondents. This is precisely where statistics comes into picture.

Statistics as a discipline provides techniques of organizing and analysing data. Statistics serves the function of summarizing the available data to make them more

intelligible. For instance, with the help of statistics we can summarize very fast the association between variables in a given sample. Statistics facilitates the drawing of inference from a given set of data in the form of generalizations. It also makes possible the testing of the reliability of research findings. Based on probability theory, statistics can be used to estimate population characteristics from the sample survey and to test hypotheses. In sum, statistics are designed to summarise data and to reveal the extent to which one may generalize beyond the data at hand.

There are various statistical operations which help to achieve the objectives of data analysis. Most of the procedures are described fully in statistics text books (or in books on research methodology). We shall give here only a preliminary introduction. Statistical analysis sometimes begins with the inspection of each variable separately. It is called univariate analysis, (i.e., single variable analysis). Univariate analysis is conducted as a prelude to more complex analyses or as a part of a basic descriptive study. In any case, the goal is to get a clear picture of the data by examining one variable at a time. Univariate analysis helps to reveal the nature of the variation in the variables under study. The purpose is to see if there is sufficient variation in responses to warrant the inclusion of the variable in the analysis.

The variations in responses are revealed by statistically organizing the data in terms of **frequency distribution** and **percentage distribution**. A frequency distribution is created by first listing all of the response categories and then adding up the number of cases that fall into each category. Frequency distribution presents a clearer picture than case by case listing of responses. However, it only serves the purpose of a preliminary organisation of data. To get a more clear picture of data, researchers often compute the percentage of respondents in each category. Percentage distributions provide a comparative framework for interpreting responses and depict more clearly the relative difference among responses. They also make visible the size of the particular category in relation to the size of the sample. To create percentage distribution, you divide the number of cases in each category by the total number of cases and multiply by 100.

Statistics helps to reveal various statistical properties of univariate distributions. The notable properties examined by statistical analysis are **central tendency, dispersion** and **shape**. The determination of central tendency characterizes what is 'typical' in the data. In the terminology of statistics three main measures of central tendency are the **mean,** the **median** and the **mode.**

The mean is the arithmatical average, calculated by adding up all the responses and dividing by the total number of respondents. It is the 'balancing' point in a distribution, because sum of the differences of all the values from the mean is exactly equal to zero. The median is the midpoint in a distribution — the value of the middle response — half of the responses are above it and half are below. You find the median by arranging the values of a variable from low to high (or high to low) and then counting up until you find the middle value. The mode is the value or category with the highest frequency.

Let us go back to Table 3.1 which is an univariate table. We have the frequency distributions by weight i.e., the number of executives whose weight falls in each category. Now we can calculate the percentage distribution in each category by dividing the frequency in each category by the total i.e. 65 and multiplying it with 100. The sum total of the percentage distribution will add up to 100.

The second property that can be summarized statistically is the degree of variability or dispersion among a set of values. The simplest dispersion measure is the **range**. Statistically, this is the difference between the lowest and the highest values and is usually reported by identifying the end points. For instance, in Table 3.1 if the least weight of a person is 47 kg. and the highest 78 kg. the range will be 78-47 or 31. The most commonly used measure of dispersion is the standard deviation. This is a measure of the 'average' spread of observations around the mean.

The third statistical property of univariate distribution is their shape. This property is made apparent through a graphic presentation called a frequency or percentage polygon. The combination of these three statistical properties i.e. central tendency, dispersion and shape provide a good picture of quantitative data

Bivariate analysis examines the nature of the relationship between two variables. Such analysis begins with the construction of cross-tabulations. There are rules for percentaging cross-tabulations and also for making necessary comparisons. Statistically, bivariate analysis amounts to calculating the degree of association between variables. Causal inferences are based not only on association but also on theoretical assumptions and empirical evidence about direction of influence and non-spuriousness.

However, let us briefly indicate what is involved in any statistical analysis. In giving an adequate description of the mass of data, we usually wish to do one or more of the following things:

- to characterize what is 'typical' in the group;
- to indicate how widely individuals in the group vary;
- to show other aspects of how individuals are distributed with respect to the variable being measured;
- to show the relationship of the different variables in the data to one another;
- to describe the differences between two or more groups of individuals.

We also wish to draw generalisations applicable to populations from which samples were taken for study and to infer causal relationship among variables.

Since statistical statements are statements of probability, we can never rely on statistical evidence alone for a judgement of whether or not we will accept a hypothesis as true. Confidence in the interpretation of a research result requires not only statistical confidence in the reliability of the findings but, in addition, it also requires evidence about the validity of the presuppositions of the research venture.

3.6 PRESENTATION OF DATA

For anyone seriously interested in social research, knowledge about the presentation of data is indispensable. In the following we will give you a preliminary idea of data presentation through statistical tables and diagrammatic presentation of data.

3.6.1 Statistical Tables

No matter what type of problem one is investigating, there will almost invariably be a need to arrange data (or at least part of them) in statistical tables. It is, therefore, important for a student of social research to understand the basic principles involved in the construction of tables. We state them below:

- A table must have internal logic and order. For instance, if you are to present a variable such as height of individuals in tabular form, then you must order them either from the tallest to the shortest or *vice versa*. Similarly, in the case of qualities where order may not be so obvious, the need for some kind of logical treatment is equally important.
- The units entered in the left hand column describing the qualities or values of the variable must be mutually exclusive. However, they must cover or include the vast majority of observations. For instance, if you are interested to present your data in terms of religion then mutually exclusive categories will be Hindus, Muslims, Jains, Sikhs, Christians, Buddhists and 'Others'.
- If the left hand coloumn of a table is a quantitative variable, then the class interval must be carefully and reasonably chosen. While this would depend to some extent on the range and the pattern of distribution of the data, the usual practice is that the number of categories should not be less than 5 and not more than 12. There would of course be exceptions. Care must be taken to construct intervals of uniform size. After having decided about the size of class intervals, it is important to clearly designate them. Each interval must have definite lower and upper limits.

In any discussion of the rules and practices which should be followed in the construction of statistical tables, it is important to make a distinction between the general purpose table and the special purpose table. The former is also known as an original, primary or reference table and the latter has been often referred to as an analytical, summary, interpretative or secondary table.

The general purpose table is constructed to include a large amount of source data in convenient and accessible form. The special purpose table is designed to illustrate or demonstrate certain aspects of statistical analysis or to emphasize a significant relationship pertaining to data. For example, tables in the reports on population, or vital statistics are of the general purpose type. They are usually extensive repositories of statistical information. Examples of the special purpose tables may be found in monographs and articles. In such cases, a table is fundamentally an instrument of analysis.

Now the procedures for constructing statistical tables are not entirely standardized, but there are certain generally accepted practices. We state below usages which need to be followed:

- Every table should have a title. The title should indicate very briefly a description of the contents of the table and give an idea about the type of data it will contain. This is important specially when there is a list of tables which the reader goes through before actually looking at the table;
- Every table should be identified by a number to facilitate easy reference;
- The headings of columns and rows of the table should be clear and brief;
- Any explanatory footnotes concerning the table must be placed beneath the table itself;
- The actual arrangement of the major classes in the table should depend on the facts and relationships that are to be emphasized;
- The arrangement of the categories in a table may be chronological, geographical, alphabetical or according to magnitude. In other words, such an arrangement should fulfil the demands related to the nature of data;
- Abbreviations should be avoided whenever possible and ditto marks should not be used;
- The unit of reporting (where applicable) should be stated. For instance, if the population given in the table is in million then this should be stated at the top of the table. Sometimes presentation of numbers in such summary form becomes necessary;
- If the table is from some other study or statistical compilation, the source should invariably be given.

The main advantages of presenting data in the form of statistical tables are:

- Tables conserve space and reduce explanatory and descriptive statements to a minimum;
- The visualization of relations among variables and the process of comparison is greatly enhanced by tables;
- Data presented in tables can be more easily remembered;
- Tables facilitate the summation of items and the detection of errors and omissions; and
- Statistical tables provide basis for computations.

Activity II

Select two tables and see if they satisfy the criteria we have mentioned in the preceding pages. Discuss with your academic counsellor.

3.6.2 Diagrammatic Presentation

Now it is quite possible that the meaning of a series of figures in a textual or tabular form may be difficult for the mind to grasp or retain. In order to overcome such a limitation, the data is usually presented in the form of diagrams. Diagrams are basic aids in research and help in the translation of complex statistical data into a form which can be easily grasped by the mind. Diagrammatic presentations make large masses of data clear and comprehensible.

Diagrams help in the depiction of facts concisely, logically and in a simple manner. For instance, properly constructed graphs relieve the mind of burdensome details and serve to indicate the conclusions of a study. In fact, the graph is one of the most commonly used devices for proving a trend or relationship between variables.

In social research different types of diagrams are used to present data. There are dimensional diagrams which may further be characterized as one dimensional, two dimensional and three dimensional diagrams. One dimensional diagrams are in the

shape of vertical or horizontal lines or bars. The length of lines or bars is made in proportion to the different figures which they are supposed to represent. Such diagrams are also in the form of bar diagrams and the width of bars varies in accordance with the nature and value of data. These bar diagrams can further be simple bar diagrams, multiple bar diagrams and sub-divided bar diagrams.

Two dimensional diagrams are in the shape of rectangles, squares or circles wherein a specific area represents the size of an item or the characteristics of a particular data. Three dimensional diagrams are in the form of cubes, blocks or cylinders. Again the cubes or blocks are constructed in such a way that their volume is in proportion to the given value of data.

Diagrammatic presentation of data can also be in the form of graphs or curves. Graphs are used as a good aid in almost all types of studies. For instance, analysis of frequency distributions generally requires the use of some kind of graphic presentation. Three main types of graphs which are used to represent frequency distribution are called histogram, frequency polygon and smoothened curve. Graphs depict the data in a comprehensible form and help the reader to understand them in a simpler manner.

Now, certain basic considerations must be kept in mind while presenting the data in the form of diagrams. The researcher, should pose the following questions to himself: what is the purpose of the diagram? What facts are to be emphasized? What is the educational level of the audience? How much time is available for the preparation of the diagram? What type of diagram can have the greatest appeal in a given situation? What type of diagram will portray the data most clearly and accurately?

Like all social research techniques, diagrammatic presentation must be used with care and discrimination. Poorly constructed diagrams may seriously damage the prospects of an otherwise good research work since they may actually misrepresent or distort the facts. Since the main objective of diagrammatic presentation is to clarify data, it is important not to make them complicated. Diagrams crowded with too many facts may cause confusion. Extreme care must be exercised while selecting a particular type of diagram. It should be most appropriate to the problem at hand. For instance, there must be some sound reason for preferring a graph to some other type of diagram.

Accuracy is just as important in the construction of diagrams as it is in other aspects related to social research. It is most desirable to be extremely meticulous in the actual construction of diagrams.

In our discussion we have just furnished a preliminary view of diagrammatic presentation of data. However, much more is involved in the actual construction of diagrams. Before you actually enter into any research venture, you should acquire thorough knowledge about the techniques involved in the construction of diagrams. For more knowledge in this area you should consult some textbook on statistics and a manual on diagrammatic presentation.

3.6.3 Outline of a Research Report

The preparation of a report is the final stage of the research. Its purpose is to convey the results of study in sufficient detail. It enables others to comprehend the data. It is also a statement about the validity of conclusions.

Now an obvious question that may arise in your mind is: What should be the final form of a research report? In the following, we will suggest some of the major features which a good research report should contain.

A Introduction
This should give a concise and clear cut statement as to the nature of the study, aims and objectives and scope of the study.

B Design of the Study
This should cover: Concepts and definitions; Character of the research design; Sampling design; Tools of data collection employed in the study and the

conditions under which they were used; The period of data collection; Limitations of the data (if any).

C Statistical procedures used in the analysis of data

D Findings of the study

E Conclusions

F Bibliographical references

G Appendices

In general, this is a broad outline for the preparation of a research report. However, this is subject to change and any additions or deletions can be made depending upon the nature of a particular research study.

Check Your Progress IV

Note: a) After carefully reading the preceding text complete the following statements.

b) Compare your answer with the model answer given at the end of the Unit.

a) ..... provides techniques for the organization and analysis of **data.**

b) With the help of statistics we can summarize very fast the ..... between variables in a given sample.

c) Statistical organization of data reveals variations in responses in terms of ..... and ..... .

d) The notable properties of univariate distributions revealed by statistical analysis are 1..... 2..... 3..... .

e) Three main measures of central tendency in the language of statistics are 1..... 2..... 3..... .

f) The general purpose table is constructed to include ..... in convenient and accessible form.

g) The special purpose table is designed to illustrate a ..... pertaining to data.

h) ..... help in the depiction of facts concisely, logically and in a simple manner.

i) The different types of diagrams used to present data in social research are 1..... 2..... 3..... .

3.7 LET US SUM UP

In this Unit we have discussed some of the fundamental aspects related to the analysis of data and their presentation. At the very outset we have explained that analysis and interpretation of data form a vital link between theory and empirical evidence. We have explained that processing and interpretation of data demand organization, reorganization and understanding of data so as to reach certain conclusions in the light of theoretical concerns of the study. We have also noted that measurement and evaluation and the attendant issues of reliability and validity are important aspects which need special attention in the context of data analysis.

We have given considerable attention to various steps involved in the systematic treatment of data. We have explained to you what is involved in the scrutiny and classification of data; establishment of categories; coding operation; and basic elements of Tabulation. Towards the end of this Unit we have discussed preliminary aspects related to statistical analysis of data. However, for elaborate statistical operations you will require further reading in this area. Discussed also is the form in which data are to be presented alongwith some idea about the outline of a research report.

3.8 KEY WORDS

Univariate : involving use of one variable.

Bivariate : involving use of two variable.

Multivariate : involving use of more than two variables.

Mean : A measure of central tendency that indicates the average value of a univariate distribution of interval or ratio scale data. The mean or arithmatic average is calculated by adding up the individual values and dividing by the total number of cases.

Median : A measure of central tendency indicating the midpoint in a univariate distribution of interval or ratio scale data. The median indicates the point below and above which 50 per cent of the values fall.

Mode : A measure of central tendency representing the value or category of a frequency distribution having the highest frequency; the most typical value.

Standard Deviation : A univariate measure of variability or dispersion that indicates the average "spread" of observations about the mean.

Frequency Distribution : A tabulation of the number of cases falling into each category of a variable.

Percentage Distribution : A norming operation that facilitates interpreting and comparing frequency distributions by transforming each to a common yardstick of 100 units (percentage points in length), the number of cases in each category is divided by the total and multiplied by 100.

3.9 SUGGESTED READINGS

Davis, J.A., 1971 — *Elementary Survey Analysis,* Prentice Hall: Englewood Cliffs, New Jersey.

Goode, W.J. and Hatt, P.K. 1952 — *Methods in Social Research,* McGraw-Hill: London.

Hymn, H.H. 1955 — *Survey Research and Analysis,* The Free Press: Glencoe Illinois.

Jahoda, M., Morton D, and Stuart, WC. 1951 — *Research Methods in Social Relations,* Dryden: New York.

3.10 MODEL ANSWERS

Check Your Progress I

1) True 2) True 3) False 4) True
5) True 6) True 7) True 8) False
9) True 10) False

Check Your Progress IV

(a) statistics (b) association (c) frequency distribution and percentage distribution (d) (1) central tendency (2) dispersion (3) shape (e) (1) mean (2) median (3) mode (f) large amount of source data (g) significant relationship (h) diagrams (i) (1) one dimensional (2) two dimensional (3) Three dimensional.

UNIT 4 STATISTICS ON RURAL DEVELOPMENT

Contents

4.0 Aims and Objectives
4.1 Introduction
4.2 Basic Rural Statistics
4.3 Sources of Rural Statistics
4.4 Rural Development Statistics
4.5 Computerisation
4.6 Let Us Sum Up
4.7 Key Words
4.8 Suggested Readings
4.9 Model Answers

4.0 AIMS AND OBJECTIVES

This unit is aimed at familiarising you with the statistics required for planning, implementation, monitoring and evaluation of rural development programmes at the national, state, district and block levels. After learning this unit you should be able to:

- identify statistical requirements for proper planning and efficient implementation of rural development programmes at various levels;
- obtain the requisite statistics from the appropriate sources;
- use the data for assessing the situation; and
- appreciate the scope and potential of the computer in rural development.

4.1 INTRODUCTION

Different aspects of rural development planning and management have been learned by you in the preceding blocks. You might have realised that data or statistics are key inputs for the different stages of planning and for various management functions. In this unit we will cover basic statistics on rural India and statistics related to rural development programmes. We will also briefly touch upon the role of computer in rural development in the context of an experiment in Computerisation known as Computerised Rural Information Systems Project (CRISP).

4.2 BASIC RURAL STATISTICS

Statistics are the building blocks for development planning. This is true in the case of planning for rural development as well. To remove the abstractness of the concept of basic rural statistics we can begin with the listing of the items normally included under the following heads:

- Area and Population — national, state, district, block and village; also classification of population into rural/urban, Scheduled Castes/Scheduled Tribes population, population of disadvantaged groups like landless labour, small and marginal farmers etc.
- Demographic characteristics of the population — age and sex distribution, birth rate, death rate, and natural growth rate, infant mortality rate, expectation of life, migration etc.
- Labour force and employment — Estimates of labour force; employment, unemployment and under-employment; sector-wise and enterprise-wise details of employment; self-employment and wage-employment; growth of labour and employment; occupational distribution; age and sex distribution of the labour force.

- Agriculture — Pattern of land utilisation; number, size of agricultural holdings and their distribution; area production and productivity of principal crops; key indicators of agricultural progress such as irrigation, area under high yielding varieties, use of fertilisers and other modern agricultural inputs, technologies etc.
- Livestock — Details of livestock population; milk production poultry production, meat production, fish production etc.
- Rural credit — Financial institutions including credit cooperatives; short-term, medium-term and long-term loans advanced to agriculture and other rural enterprises; non-institutional rural credit etc.
- Prices and cost of living — Prices of inputs and outputs; wholesale prices and retail prices, cost of living index etc.
- Basic amenities — availability of educational institutions, health and medical facilities, safe drinking water, road, transport, communication facilities, markets, electricity and other basic social and economic infrastructure facilities.
- Social and economic siutation of the population, with special reference to the disadvantaged; social and economic problems facing the population.
- Organisational and institutional infrastructure, both state sponsored and non-governmental.
- Social and economic development programmes in an area; extent to which benefits have flowed to different segments of the population.

The list above is long, but is certainly not exhaustive.

4.3 SOURCES OF RURAL STATISTICS

Depending upon the nature, extent of sophistication and level of disaggregation of planning, the statistical requirements vary considerably. For example, to prepare a proper IRDP plan at the block level, one requires all the statistics listed above plus a lot more detailed information about the poverty profile, ownership of income generating assets, income level and main sources of income, resource potential in the region which may include as diverse information as the ground water potential of the area, the skills and educational levels of the population, the health status and morbidity pattern of the population and so on. Some of the data may not be readily available and may have to be collected through special surveys which may be costly and time consuming. We shall devote the following paragraphs to briefly discuss the sources and methods of collection of basic rural statistics, periodicity and timeliness, reliability and validity of the data and estimates, and uses to which statistics are put.

4.3.1 Administrative Statistics

Broadly, the sources of basic rural statistics can be divided into three types:
(i) administrative records,
(ii) censuses and
(iii) sample surveys.

By and large, a significant amount of basic rural statistics are available in the administrative records of various government offices at the district, block/**taluka** and village/**panchayat** levels. Some of these are generated as byproducts of administration often as a result of a delibrate action, and others as products of certain government regulations. For example, as a byproduct of revenue administration, land ownership statistics and land utilisation statistics are maintained by the village **patwari** (revenue official). Similarly, the gram panchayat **adhikari** (official) maintains essential details of the inhabitants within the **panchayat** area for various official purposes. The primary school records contain details regarding school enrolment and dropout, while vital events like births and deaths are recorded in the registration office for the purpose. A major merit of such statistics is that they are collected/generated regularly without any additional explicit expenditure for data collection. Because of the legal authority of the administrative machinery, there is little difficulty in getting cooperation from respondents to collect the requisite statistics. Often, the respondents have to furnish the requisite information to the authority as a matter of routine to comply with certain administrative and regulatory stipulations.

A major problem with the administrative statistics, however, is that they often lie scattered in various administrative records of the offices of various functional departments. From the user's point of view, the first task will be to systematically compile these statistics from the records, files and progress reports maintained in the field offices as well as higher level offices. Quite often this may be a difficult task as the records may not be properly maintained or uptodate. Further, explanatory notes/definitions in regard to the data may be missing. For example, the records of ownership of land holding maintained by the **patwari** may not indicate the current ownership status as mutations due to partitions, sales etc., are not normally carried out promptly.

4.3.2 Census Statistics

Censuses are the next major source of basic statistics including rural statistics. Census enumeration is conducted periodically to collect specific informaion relating to the entire population. The population census is the most important of them. It is conducted once in ten years in India. The last census was in 1991. The first population census in our country was conducted in 1871 and since then we have been having regular censuses. A variety of information relating to the entire population is collected during the census. These include statistics on demographic characteristics, housing and other infrastructure facilities, economic status, occupational details, employment status, literacy level, and other social statistics. Almost all census data provide a rural urban break-up. The Census Commissioner of India is responsible for the conduct of population census throughout the country. At the state level, there are state census commissioners. The field work or actual data collection is normally got done through school teachers and other government functionaries who are given special training and remuneration for this purpose.

Though the actual field work for census is only a one-month operation (usually in Feburary of the census year), the preparations for census take a few years and the compilation and analysis of census data and preparation of various census reports at the national, state and district levels take several years. Normally, the total population figures and the important characteristics of the population at the national and state level are made available soon after the census. More detailed and disaggregated figures are normally available only with considerable time lag. The delays in the processing of census data are a matter of concern, and the expectation is that with the aid of computerisation this can be considerably reduced. Apart from national and state level aggregate estimates, the population census provides detailed disaggregated data at the district, block and village levels which are extremely useful for decentralised planning.

In addition to the population census, the other important censuses conducted in our country on a regular basis are Agricultural Census, Economic Census and Livestock Census. Agricultural census has been conducted every five years since 1970-71. It throws up detailed statistics on land holdings, distribution of operational holdings by size, area of operational holdings in different size, classes etc. Economic Census collects data on household and unregistered economic enterprises. Besides, data on various village amenities — social and economic infrastructure facilities — are also collected under Economic Census which are extremely useful for planning of basic amenities in rural areas. So far, three economic censuses have been conducted, the first one in 1977. Livestock census collects detailed statistics on the livestock population in the country which is quite useful in planning for rural development.

4.3.3 Sample Survey Statistics

The third major source of basic rural statistics is sample surveys. Unlike the census where information is collected from all the individuals (units) in the population, in sample survey; information is collected only from a representative sample of individuals from the population. On the basis of statistics collected through sample surveys, reliable estimates about the characteristics can be made for the population. As compared to census, a sample survey is less time consuming, less costly and often more reliable as non-sampling errors can be minimised by having better trained professional enumerators. A major limitation of sample survey is that it will not provide reliable estimates of population characteristics at disaggregated levels (beyond what was decided at the time the sample design was prepared) due to

limitations of sample size. Of course, sample surveys will not provide individual data relating to all the units of the population either.

The most important and best known sample surveys are those conducted by the National Sample Survey Organisation (NSSO). NSSO has been conducting socio-economic surveys on a regular basis since 1951. These surveys are of national coverage and are conducted in rounds of one year duration, after the initial few years when the duration of a round was less than one year. Each round covers selected facets of the economy and society. Over a period of ten years the subjects are covered in rotation. For instance, the 43rd round of NSSO survey was conducted during 1987-88 and covered employment and unemployment. It may be noted in this connection that the estimates of poverty line and the population below the poverty line are based on the consumer expenditure surveys conducted by the NSSO periodically.

For the collection of vital statistics, the Sample Registration System makes available every year for the country as a whole data on birth rates, death rates, infant mortality rates, age specific death rates etc. Sometimes special surveys are carried out by the Sample Registration System along with the regular surveys. The SRS was initiated in 1964-65 on a pilot basis in a few selected states but now covers the entire country.

4.3.4 Statistics from Research Studies

Empirical research — exploratory, diagnostic, and evaluative — on different facets of rural society and economy are sponsored by organisations like the Indian Council of Social Science Research, Planning Commission and different Departments of the Government of India. These micro studies conducted by research institutes/university departments with the help of grants given by the sponsoring organisation provide useful feedback on the development scenario. Sometimes these studies are conducted throughout the country with the participation of different university department/research institutes following a common research design and a common core tabulation plan. The state and other state agencies occasionally conduct sample surveys to collect specific basic rural statistics for planning, monitoring and evaluation of various rural development programmes. Another agency which has nationwide field officers and survey teams for conducting socio-economic sample surveys is the Programme Evaluation Organisation (PEO) of the Planning Commission. The PEO conducts evaluation studies of important national programmes. In recent years, it has conducted evaluation studies of IRDP and NREP. The states, too, have their own evaluation wing, usually within the Directorate of Economics and Statistics.

Check Your Progress I

Notes: a) Use the space given below for answers.
b) Check your answers with those given at the end of the unit.

1) Indicate whether the following statements are True or False. Put a tick (✓) mark in the relevant box.

		True	False
a)	Statistics are the building blocks for rural development planning	☐	☐
b)	Special Surveys are the only source of statistics for rural development	☐	☐
c)	Administrative statistics are specially collected for IRDP planning	☐	☐
d)	Population Census is conducted every five years	☐	☐
e)	Sample surveys will not provide data relating to each and every unit in the population	☐	☐

2) Name three censuses regularly conducted in the country.

.....

.....

.....

.....

3) Briefly describe the best known sample surveys in the country.

.....

.....

.....

.....

4.4 RURAL DEVELOPMENT STATISTICS

Rural development programmes in India are of vast dimensions requiring massive data for planning and implementation. Also, in the wake of their implementation and for the purpose of progress reporting, monitoring and concurrent evaluation, large volumes of data and information are being generated. In the following paragraphs we shall briefly discuss statistics that are being collected for monitoring and concurrent evaluation of the major anti-poverty programmes.

4.4.1 Monitoring of Anti-Poverty Programmes

Over the years, a fairly comprehensive system of monitoring of anti-poverty programmes has been developed. This is based on a regular system of progress reports and feedback. For this purpose statistical proformae have been prescribed. The monthly progress reports are brief and contain key information relating to physical and financial progress. Quarterly reports are more detailed which provide data and information relating to progress achieved, both quantitative and qualitative. Annual reports provide even more detailed statistics on progress achieved, problems faced and remedial corrective measures taken, and quantitative/qualitative information on the effects and impact of the programme.

The primary reporting agency is the block office which maintains the basic details regarding the schemes and projects being implemented within the jurisdiction of the block. The reports from different block offices are sent to the District Rural Development Agency (DRDA) where they are consolidated. The reports received from different DRDAs are consolidated at the state headquarters. National level progress reports are prepared by the Department of Rural Development of the Government of India on the basis of reports received from each state.

The Department of Rural Development at the Centre has prescribed that the monthly progress reports from the states should reach the Centre by the 10th of the following month, the quarterly reports by the 25th of the month following the quarter and the annual report before the end of the first quarter following the year. Though monthly reports are, by and large, regularly received in time from most of the states, the same is not true in the case of quarterly and annual reports. This is mainly due to the time consuming process of manual tabulation and record maintenance. A major constraint in this connection is the shortage of trained manpower for compilation of data which might affect the quality of reporting also.

Whereas the Central and State Governments monitor the implementation of these programmes on the basis of certain key indicators on a monthly, quarterly and annual basis, closer day-to-day monitoring is the shared responsibility of the DRDAs and block offices. As such, the requirements at the lower levels of management are much higher. In practice, however, the collection and compilation of data are more often dictated by the requirements of the higher authorities.

In the case of IRDP, the basic data collected on a monthly basis for monitoring purpose include the number of families assisted, the sectoral distribution of the beneficiaries, the social class and gender distribution of the beneficiaries etc. On the financial side, the total investment, the per capita investment — by sector and social class, the share of subsidy and bank credit, the institutional distribution of bank loans

etc. are collected. These are compared with the targets and norms provided. Significant deviations from the targets and norms are matters of concern needing corrective supportive action. The quarterly reports, apart from the above details, provide information on qualitative aspects of the implementation of the programme — problems faced and how they have been overcome. The annual reports provide, in addition, details of income generation from the assets, and changes in the overall income of the assisted families as well as details of loan repayments.

The progress reports of Jawahar Rozgar Yojana contain details regarding the number of mandays of employment generation, the social occupational class and gender of those who get work, the types of work taken under the programme, details of foodgrain distribution and so on. On the financial side, details of outlays, the share of wage and material component, average wage rate and average cost of generation of mandays of employment etc. are reported. Again, achievements are compared with the targets and norms and wherever major departures occur, these are indicated for follow up action and corrective (or supportive) measures.

4.4.2 Concurrent Evaluation of Anti-Poverty Programmes

A few national level evaluation studies of IRDP during the Sixth Five Year Plan brought out a number of drawbacks in the implementation of the programme. Some of the major shortcomings noted were: (i) assistance going to ineligible beneficiaries; (ii) inadequate assistance, especially to the poorest of the poor; (iii) weak planning and implementing machinery; (iv) inadequate training and frequent transfers of the staff; (v) short and unrealistic repayment periods, arbitrary recovery schedules, rejection of applications on flimsy grounds, insisting on securities even if it is not required and, in general, an absence of development banking culture on the part of banks; (vi) bias towards the distribution of a few assets which are easy to procure, irrespective of their viability — milch animals and sewing machines being the best examples; and (vii) bureaucratic approach with very limited involvement of the people.

On the basis of the findings of the evaluation studies and the fresh data on rural poverty furnished by the latest round of consumer expenditure survey of NSSO, the IRDP was restructured in a number of important ways during the Seventh Plan. Correct feedback on the implementation of the programme with the shortest possible delay is an important management tool. In that respect the experience of monitoring based on official progress reporting was not encouraging. As a result, an alternate system known as 'Concurrent Evaluation' was introduced in 1985. Basically, it is a combination of monitoring and evaluation, and the results have been quite encouraging.

The data for concurrent evaluation of IRDP are being collected by reputed research institutions in the country which are independent of the government machinery implementing the anti-poverty programmes. The survey is conducted on the basis of well-defined multi stage random sampling scheme. Every month, 36 districts are selected at random and from each district 2 blocks are selected at random. From each of the selected blocks, 20 IRDP beneficiaries are selected for detailed study, of whom 10 are new beneficiaries and the remaining 10 are old beneficiaries who were assisted two years before the survey period. Thus the total sample size for the monthly survey is 1440 beneficiaries. Over a period of 12 months, all the districts are covered by the survey and a total of about 17,280 beneficiaries are covered by the study.

The data are collected by well-trained field investigators of the research institutions using a structured and pre-coded household schedule. Interviews are backed by observation, physical verification of assets and checking of records for filling up the schedules. Most of the information collected from new beneficiaries relates to selection procedures, delays, quality of assets, loan procedures, infrastructural facilities etc. In the case of old beneficiaries the stress is more on the incremental incomes and the repayment of loans, overdues etc. In other words, information from new beneficiaries is essentially for monitoring and from old beneficiaries for assessing the effects of the programme.

Twentynine research institutions are participating in the scheme. Except for a few, the others conduct the survey in a district each, every month. The filled up schedules,

after proper scrutiny, are sent to the Monitoring Division of the Department of Rural Development of the Government of India so as to reach by the middle of the month following the survey. A copy is sent to the concerned State government and another is retained by the institution for its own research and analysis purposes. The schedules received at the Centre are fed into a computer and based on a set of tabulations obtained from the computer, a monthly summary report is prepared with the least time lag. Copies of these reports are made available to all the concerned parties, including state governments, DRDAs and banks. Apart from the monthly reports, quarterly, six monthly and annual reports using cumulated sample data are also prepared to get a more stable picture. Wherever significant departures from targets/norms are indicated, the concerned state government/DRDA/Bank is requested to examine the matter and take necessary corrective/supportive follow up action.

The first round of IRDP concurrent evaluation survey was completed by September, 1986. The second round was conducted during the calendar year 1987. The third round had begun in January 1989. Apart from generating a large volume of micro level data relating to the IRDP beneficiaries, the concurrent evaluation has helped to improve the official system of progress reporting. Also, the fact that independent survey teams may visit the beneficiaries on a random basis and check the records has led to an overall improvement in the quality of implementation of the programme.

On the basis of the experience gained from concurrent evaluation of IRDP, a similar scheme was started for the erstwhile NREP. The first round of NREP concurrent evaluation survey was conducted during November 1987 to October 1988. The findings of the survey enabled the agencies concerned to pinpoint and correct a number of drawbacks in the implementation of the programme which were not brought to notice by the normal system of progress reporting.

Check Your Progress II

Notes: a) Use the space given below for answers.
b) Check your answers with those given at the end of the unit.

1) Indicate whether the following statements are True or False. Put a tick-mark (✓) in the relevant box.

	True	False
a) Most of the rural development administrative statistics are highly aggregated in nature.	☐	☐
b) Information requirements at the lower management level are much more detailed than at the higher levels of management.	☐	☐
c) Concurrent evaluation is a combination of monitoring and evaluation.	☐	☐
d) Information is obtained from new beneficiaries of IRDP for evaluation.	☐	☐
e) Concurrent evaluation has adversely affected the quality of official progress reporting.	☐	☐

2) In what important ways is the monitoring system of rural development project different from the concurrent evaluation system.

(Hint : See the text carefully — the different purposes and different methods employed.)

.....

.....

.....

.....

.....

.....

4.5 COMPUTERISATION

A major problem with the manual system of data compilation and storage is that data and information maintained in the records and files of block offices, DRDAs and other offices are not easily retrievable and, as such, cannot be used for further analysis for planning purposes. In the existing system of progress reports, the data mobility is invariably upwards. There is hardly any horizontal or downward mobility of data. Because of this inflexibility of the manual system, often basic data needed for planning are not available at the DRDA and other offices. Also, elected representatives of the people, non-governmental agencies and the public at large are not adequately kept informed about the details and progress of various rural development programmes. As a result, there is limited opportunity for an interface between the progress as reported upwards and the progress as it is witnessed in reality.

The approach to the Seventh Plan stated that a massive volume of data had already been collected in connection with the planning and implementation of anti-poverty programmes. However, not much of these data had been analysed and effectively used for plan formulation, implementation and management control. The manual system of data management was found wanting to meet these requirements. In this connection, the Planning Commission recommended that the Department of Rural Development may initiate steps to develop a computerised management information system (MIS) at the DRDA level.

4.5.1 The Karwar Experiment

Computerisation of MIS for rural development programmes had already been taken up on an experimental basis in a few districts in the country. One such rather promising experiment was in Karwar in Karnataka State. The Karwar experiment which was started in mid 1984 had started showing significant positive results. The experimental computer at Karwar consisted of a micro-processor and the requisite peripherals. This was installed at the office of the Project Director, DRDA. A system design had been built up for data management for all the programmes implemented by the DRDA. Information on about 25,000 IRDP beneficiaries had already been filed on individual records. Daily transactions relating to IRDP of every bank branch in the district was being input for undating all beneficiary records. A separate module was prepared for monitoring loan repayment. A village database for all planning activities in the district was being fed into the system. Area planning for the district relating to NREP and RLEGP for 1985-86 had been done with the help of the DRDA computer.

Some of the tangible results of the Karwar experiment are worth noting. To start with, all basic data and key information relating to various rural development programmes implemented by DRDAs were available at the headquarters and they were easily accessible. This had given better management control for the implementing agency. Further, updated information regarding the progress of various projects and activities available at the headquarters enabled the project director and other officials to conduct effective field inspections which kept the field functionaries vigilant in respect to the correctness of the information furnished, the quality of the assets provided to the individual beneficiaries and the works undertaken. In other words, computerised MIS ensured that the quality of reporting, the quality of implementation itself and beneficiary satisfaction have simultaneously improved.

An added advantage of Karwar experiment was that it had relieved the office staff from much of the routinised and repetitive time consuming work. A happy feature of the experiment was that from the beginning, the entire office was involved in it and virtually everybody in the office had operational access to the system and drew benefit from it. Thus, it had demystified a science and an art which was for long totally esoteric in character, handled by highly skilled professionals.

4.5.2 Crisp

Based on the experience gained in Karwar district, the Department of Rural Development, Government of India decided to introduce a pilot project on Computerisation of Rural Information System in 10 districts in different States.

A detailed project report was prepared. The focus of the project was to develop a functional software to plan, monitor and evaluate various anti-poverty programmes sponsored by the Government of India and implemented by the DRDAs in all the States of the country. The project was named Computerised Rural Information Systems Project and came to be known as CRISP.

CRISP was formally launched in May, 1986 and a series of training programmes and field visits were organised by the project teams. A Project Review Committee with Secretary, Department of Rural Development as Chairman and experts from computer software and hardware, Management Information System, District Planning and Rural Development as Members was constituted to oversee the implementation of CRISP in the pilot districts. The National Informatics Centre of the Department of Electronics was made the nodal agency for implementation of the pilot Project. Selection of the pilot districts was done with care to ensure that the 10 selected districts represent a true sample of all the districts in the country in terms of level of development, diversity of agro-climatic conditions, administrative cultures etc.

Given the various diversities in the different districts, slowly a standardization of input and output formats was evolved. In almost all the pilot districts the computers started functioning in the first half of 1987, after the necessary ground work was done. Once the computers in the pilot districts had become functional, an evaluation process started, culminating in a meeting of Secretaries of Rural Development from all the States.

Based on the experience gained in the pilot districts and the rapid success of the software developed, in October 1987 the Department of Rural Development sanctioned funds to all the DRDAs in the country for purchase of computers and the requisite peripherals. Training of DRDA staff was also organised.

The CRISP team was given the task of preparing a detailed Users' Handbook as well as imparting training to the trainers selected by the different State Governments. A comprehensive Handbook for Users was prepared and distributed to all the training institutions and DRDAs by January, 1988. By the end of 1988, training was imparted to staff in about 350 districts. Four staff members of each DRDA were trained in handling the computer. In a number of states installation of computer hardware has been completed. Overall, in about 300 DRDAs the computers have been installed. By the end of March, 1989, training as well as installation of computers were expected to be completed in almost all the DRDAs.

The software prepared under CRISP is user-friendly. The level of technical competence needed for learning and using it is not high. This ensures easy acceptance of CRISP by administrators and managers. Further, arrangements have been made for imparting computer training using CRISP software to probationers of Indian Administrative Service at the National Academy of Administration in Mussoorie. CRISP software has been carefully prepared for interfacing with any future district computerisation to be taken up by the Department of Electronics or other agencies of the Government of India or State Government. In a few cases, DRDAs have been already been linked with the State headquarters.

The available software is CRISP Version-1. This is essentially for planning, monitoring and concurrent evaluation of the major anti-poverty programmes. CRISP Version-2 which is under preparation, is essentially for database management CRISP Version-3, which is yet to be prepared will meet the requirements of District Planning. Graphics form an important part of CRISP package. It would be possible with the help of graphics to show in a map the areas of concentration and areas which require greater attention. The software will enable one to develop various planning models which will provide rational answers to the questions on priority for locating various infrastructural facilities. It would also be possible to determine in

advance the schemes to be taken up in different villages. Availability and dissemination of such information will be able to procure for such programmes public participation and involvement which has not been available till now.

Check Your Progress III

Notes : a) Use the space provided below for the answers.
b) Check your answers with those given at the end of this unit.

1) List reasons for computerisation of rural development statistics.

.....

.....

.....

.....

2) What were the tangible results of Karwar experiment?
(Hint : See the text of subsection 1.5.1)

.....

.....

.....

.....

3) 'CRISP is a highly decentralised project.' Comment?
(Hint : See the text for indications in regard to decentralisation of the functioning of the project.)

.....

.....

.....

.....

4) Graphics form an important part of CRISP package. Why?

.....

.....

.....

.....

4.6 LET US SUM UP

In this unit we have discussed the types of statistics which have a bearing on planning, implementation, monitoring and concurrent evaluation of rural development programmes at national, state, district and block levels.

We have discussed basic rural statistics which are essential pre-requisite for proper planning of rural development at different levels. We have identified the various types of basic rural statistics and the main sources of such statistics, viz. administrative records, censuses and sample surveys. The merits and demerits of the different sources have also been indicated.

We have noted the statistics relating to rural Development Programmes being implemented in the country. This was followed by discussion of monitoring of anti-poverty programmes and their concurrent evaluation. Under monitoring we have dealt with statistics obtained through periodic progress reports whereas under concurrent evaluation we have discussed the system of data collection through sample surveys conducted by independent research institutions.

In the final section we have introduced the issue of computerised management information system for rural development at the district level. This was followed by a discussion of the Computerised Rural Information Systems Project (CRISP) which is being implemented to cover all the districts in the country.

4.7 KEY WORDS

Demography: The science of population statistics.

Evaluation: A detailed feedback on the extent of attainment of the objectives of the programme, the delivery system, cost effectiveness, impact, etc.

Infrastructure: The basic facilities like roads, communications, electricity, schools, hospitals etc.

Management Information System: A management device that enables managers to get feedback through timely, accurate, and relevant information on the programme.

User-Friendly: Easy to use.

4.8 SUGGESTED READINGS

Government of India, 1984. *Report of the Working Group on District Planning* Volume I & II, Planning Commission, New Delhi.

Government of India, 1985. *The Seventh Five Year Plan* Volume II (Chapter 2) Planning Commission, New Delhi.

Government of India, 1985. *Report on Concurrent Evaluation of IRDP,* Department of Rural Development, New Delhi. (Different Issues since December, 1985).

National Institute of Rural Development 1985. *Rural Development Statistics-1985* National Institute of Rural Development, Rajendranagar, Hyderabad.

4.9 MODEL ANSWERS

Check Your Progress I

1) a) True b) False c) False d) False e) True

2) i) Population census conducted every 10 years. The last Census was conducted in 1991.
   ii) Agricultural Census conducted every five years.
   iii) Economic Census.

3) 1) The National Sample Survey Organisation which conducts socio-economic surveys on a regular basis covering the entire country. The surveys are conducted in rounds of one year duration. Each round cover selected facets of the economy and society.
   2) The Sample Registration System which is conducted every year and furnishes data on vital events.
   3) The Programme Evaluation Organisation of the Planning Commission which evaluates different programmes on a regular basis.

Check Your Progress II

1. a) False b) True c) True d) False e) False

Check Your Progress III

1. • Retrieval
   • Speed in compilation and dissemination (vertical, lateral, diagonal).
   • Standardisation

4. It shows very clearly the pattern of distribution.

ignou
THE PEOPLE'S
UNIVERSITY

GUIDELINES FOR PREPARATION OF A PROJECT PROPOSAL

ignou
THE PEOPLE'S
UNIVERSITY

GUIDELINES FOR PREPARATION OF A PROJECT PROPOSAL

You may prepare a draft proposal and discuss the same with your supervisor. The following steps are provided to you for preparing the project proposal (see the example given in Part IV of this Project Work Handbook).

1. **Project Title**

   The title of the project should clearly specify the type and nature of the project. It should be clear, brief and specific. It is suggested that the topic should be related to your field of work.

2. **Introduction**

   In the introduction, you may provide a description of the topic, broad area of the study; why the topic is important and clarification of the important concepts.

3. **Statement of the Problem**

   The statement of problem should briefly contain an analysis and relevance of the problem. To be exacts, this is a rationale for carrying out the study. Existing literature is reviewed and gaps are brought out so as to provide justification for choosing specific topic of the study.

4. **Objectives**

   The objectives of the proposed project work should state your intention for carrying out this project. Usually a topic contains three to four objectives. These objectives can be given in a serial form, pinpointing your approach. For example, you want to study the problem of drop-out among girl children in a rural area. You may like to study the reasons for high drop-outs rate among various socio-economic classes of people. Such objectives thus, will spell out the scope of the study.

5. **Hypothesis**

   Hypothesis is the probable answer to the problem you have undertaken, and the project tests the hypothesis. But not all studies involve testing of hypothesis (mostly experiment-based studies have hypotheses). You can discuss with your supervisor as to type of study you want to undertake. **You may even avoid hypotheses for your study since you are expected to make only a small- survey of about 100 respondents.**

6. **Universe of the Study**

   If the project involves field work, you should define the universe of the study. Universe is defined as the entire area or population taken for the particular study. This will depend on the geographical limits of the study and the unit of the study. For instance, if you are studying the characteristics of households in a village, then all the households will

comprise the universe of the study. You will draw the sample from this universe.

7. **Sample**

   If your project proposal is based on field work you will have to choose a sample from the universe. A sample is representative of the whole population. Sampling can be done in many ways such as random sampling, cluster sampling etc. To know more about the different types of sampling, please refer to on any standard book on statistics which you might have consulted while studies on preparing for Research Methods is Rural Development.

8. **Tools for Data Collection**

   You should suggest what kinds of tools you intend to use for the collection of data from various sources. You may have to choose more than one tool for a particular study since social reality is always complex and multi-faceted.

9. **Data Analysis**

   The raw data has to be scrutinized and coded and then data analysis can be done, using statistical methods. It is desirable that the techniques, which you intend to use are mentioned in the proposal.

10. **Tables**

    Tables will have relevant statistical calculations such as percentage, mean, median standard deviation, co-relation etc.

    Tables will have proper numbers (i.e. Chapter I will have Table 1.1, Table 1.2 etc; Chapter 2 will have Table 2.1, Table 2.2 etc.). Table number will be followed by Table title which should be as brief as possible while conveying the matter contained in the table clearly. If a table is lifted from some other sources, it should be indicated below the table. Table interpretation may be given in three paragraphs: (i) an introduction to the parameter or topic, (ii) interpretation, (iii) major findings-indications.

11. **Chapter Plan**

    Chapter plan or chapterization will give a tentative plan for writing the report. This exercise will help you in completing your dissertation smoothly and in a systematic way.

12. **Report Writing**

    For writing report, you must follow the following guidelines:

    - Make a chapter plan (see the example in Part IV)
    - The length of each chapter should be more or less the same. For example, it should not be that one chapter is of 15 pages and another of 40 pages. The ideal chapter length should be of around 20 pages.

- Each chapter will bear
  - Chapter number
  - Chapter title
  - Introduction
  - Main titles, subtitles and sub-sub titles
  - Conclusions
  - Proper references
  - Graphics/illustrations etc.
  - Well presented with proper layout.
- Do not fill up chapters with too may tables. You may have about 5 to 7 tables in a chapter.
- Choose only the most important areas for presentation in tabular form.
- Rest of the findings can be given in a narrative form.

ignou
THE PEOPLE'S
UNIVERSITY

BROAD AREAS AND TOPICS FOR RESEARCH AND PROJECT WORK

ignou
THE PEOPLE'S
UNIVERSITY

ignou
THE PEOPLE'S
UNIVERSITY

BROAD AREAS AND TOPICS FOR RESEARCH

In order to enable you to choose a topic for your empirical study, we have listed below some broad areas. This does not mean that you cannot choose another topic of your interest pertaining to rural setting. The areas or topics listed below are only to give you some idea about the numerous titles which will guide you in selecting a topic relevant to your state, district or block which falls nearer to your place of stay. You must discuss your topic and take necessary approval from your guide before going ahead with the collection of data. Your guide will perhaps help you in the preparation of the project proposal as well as in selecting appropriate tools for the collection of data. Kindly ensure that you take prior approval for your proposal. Original copy of the approved proposal must be attached with the thesis before binding it. The suggested topics are as under:

13. A study of the factors affecting rural under development
14. A study of the effectiveness of the strategies of rural development planning
15. Impact of Five Year Plans on rural development
16. Reasons for progress in rural development programmes
17. Importance of evolving new strategies for rural development
18. Analysis of the extent of grassroots level participation in rural development
19. Rural Development policies and grassroots level realities
20. Exploration of the welfare aspects in rural development
21. Impact of IRDP on Tribes
22. Impact of IRDP on Scheduled Castes
23. A study of the approaches of NGOs to rural development
24. Salient features of rural development schemes after independence.
25. Comparative study of the characteristics of rural and urban communities in India
26. Comparative study of the characteristics of tribal and peasant communities
27. A study of traditional practices in rural cultivation
28. A study of the patterns of migration in rural India
29. An analysis of the extent to which caste system is practiced in rural India
30. A study of the role of voluntary agencies for rural development
31. A study of the relevance of *Jajmani system*
32. An analysis of rural power structure in India

33. A study of the trends in agricultural growth (in your state)
34. Importance of animal husbandry for rural development
35. A study of non-governmental activities in rural India
36. An analysis of the contribution of rural credit markets
37. A study of the importance of rural employment
38. A study of the characteristics of rural workforce
39. An analysis of wage system in rural India
40. Agriculture as a source sustaining economic growth
41. A study of the extent of equality/inequality in land distribution
42. A study of the problems faced by rural credit markets
43. Extent of rural poverty in India
44. A description of the incidence of rural poverty
45. A critical analysis of poverty alleviation programmes
46. A study of the agrarian system
47. A study of the agrarian system and classes
48. A study of the feudalistic agrarian system
49. A study of capitalistic agrarian system
50. A study of agrarian movements in India
51. A study of the approaches to land reforms
52. Impact of abolition of intermediarytenures
53. A study of the shift from cooperative to contract farming
54. A study of the problems associated with land reforms
55. A study of the strategies for green revolution
56. A study of the impact of green revolution on income andemployment
57. A study of the problems and their consequence on Green Revolution
58. A study of challenges to agricultural development due to globalization
59. Impact of agriculture extension services after independence
60. A study of the extent and nature of privatisation of rural extension services
61. Problems and challenges in rural extension services
62. A study of the organizational and administrative aspects of extension services
63. A study of the problem faced in administrative linkages from national level to village level.

64. A study of the functioning of Panchayati Raj Institutions

65. An analysis of the problems faced by rural development functionaries

66. A study of the effectiveness of the Panchayati Raj Institution for Rural Development

67. An analysis of the impact of Panchayati Raj Institutions

68. A study of the contribution of cooperatives towards rural development

69. An analysis of the functioning of credit cooperatives

70. A comparative study of the contribution of credit and non credit cooperatives in rural India

71. A comparative study of the features of cooperative and other enterprises

72. A study of the problems faced by cooperatives in rural India

73. An analysis of the relevance of cooperative organizations in a mixed economy

74. A study of the role of credit societies in rural development

75. A study of the impact of non institutional credit in India

76. Analysis of the functioning of credit cooperatives and micro-credit societies

77. A study of the extent of social changes in rural India after independence

78. An analysis of the role of self help groups in women's empowerment

79. A study of the extent of economic independence of rural women

80. A study of the extent of social empowerment of rural women

81. A study of the extent of economic empowerment of rural women

82. A study of the political management by rural women

83. Involvement of rural women in agriculture

84. Role of voluntary agencies in the empowerment of rural women

85. A study of the role of mass media for rural development

86. A study of the role of print media for rural development

87. A study of the impact of television in rural development

88. A study of the role of information, education and communication (IEC) in rural development

89. A study of the relevance of communication technology for rural development

90. A study of the innovative uses of information technology for rural development

91. A study of resources available for rural development and their utilization

ignou THE PEOPLE'S UNIVERSITY

92. A study of the extent of people's participation in rural development planning
93. Analysis of the strengths and weaknesses of decentralised planning
94. A study of the nature and process of district level planning for rural development
95. An evaluation of block level planning for rural development
96. A study of grassroots level planning for rural development
97. A study of the effectiveness of rural development management
98. An evaluation of management and administration of rural development projects
99. A study of the effectiveness of rural development projects for socio-economic development
100. An evaluation of the effectiveness of rural development projects in social development
101. A study of the extent of effectiveness of rural development projects
102. A study of the costs and benefits involved in rural development projects
103. Need-based study of rural development project proposals
104. A study of the effectiveness of the monitoring system for rural development projects
105. A study of the dimensions of rural development project evaluation
106. A study of the role of Government in promoting voluntary action for rural development
107. An analysis of the factors responsible for people's participation in rural development
108. A study of the problems of NGOs working in rural sector
109. An analysis of the strengths and limitations of NGOs in rural sector
110. An analysis of voluntary agency administration in rural sector
111. A study of the effectiveness of the management of projects by NGOs
112. An analysis of the effectiveness of community-based programmes
113. A study of programme-formulation strategies in community-based programmes
114. A study of social action strategies in rural development
115. A study of the impact of grassroot-level organizations in rural sector
116. A study of the social status of rural women
117. A study of the political status of rural women
118. A study of economic status of rural women
119. A study of educational status of rural women

120. A study of the social and educational status of girls in rural India
121. An analysis of the impact of various schemes for rural girls'education
122. A study of the factors contributing to the educational backwardness of rural women
123. An analysis of the key indicators of rural women's health and nutrition status
124. A study of the policies and programmes for health and nutrition of rural women
125. A study of the initiatives for the empowerment of rural women
126. A study of the factors influencing rural women's empowerment
127. A study of the extent of child labour in rural India
128. A study of the educational status of rural children
129. A study of the health status of rural children
130. A study of the policies and programmes for health and nutrition for rural children
131. A study of the status of non formal education of rural children
132. An evaluation of ICDS programmes in rural India
133. A study of the success of empowerment programmes of Schedule Castes
134. A study of the success of empowerment programmes of Scheduled Tribes
135. A study of the impact of various schemes for the development of Scheduled Tribes
136. A study of the socio-economic development programmes for Schedule Castes
137. A study of socio-economic development programmes of Schedule Tribes
138. A study of rural indebtedness
139. A study of the child bonded labour
140. A study of bonded women labour
141. Socio-economic status of artisans and craftsmen
142. A study of the state of rural industries
143. A study of the status of agricultural labourers
144. Contribution of Khadi and Village industries to rural development
145. A study of the extent of legal knowledge among rural masses
146. A study of the rural health care facilities in India
147. Problem faced in rural health care

148. A study of the indicators of health in rural area
149. A study of the determinants of rural health
150. A study of different diseases common in rural areas
151. A study of the linkages between health and development
152. A study of the role of NGOs in rural health care
153. A study of rural health services after independence
154. Role of primary health care centers in rural health services
155. A study of health and nutrition status of young girls in rural India
156. A study of health and nutrition status of children in rural India
157. A study of health and nutrition status of pregnant women in rural India
158. A study of the factors influencing health and nutrition status in rural India
159. A study of the problems relating to malnutrition in rural India
160. A study of different models of health care delivery in rural India
161. A study of communicable diseases in rural India
162. A study of the problems associated with communicable diseases in India
163. A study of the interventions in the area of controlling communicable diseases
164. A study of the management of health care system in rural India
165. An evaluation of national immunization programme in rural context
166. A study of the national malaria control programme
167. A study of the national tuberculosis control programme
168. A study of illiteracy-eradication programme
169. A study of the extent of the awareness about HIV/AIDS in rural India
170. An evaluation of guinea worm eradication programme
171. A study of rural sanitation programme
172. A study of the factors influencing environmental sanitation and hygiene in rural India
173. A study of the status of maternal and child health in rural India
174. An evaluation of reproductive and child health programmes in Rural India
175. An analysis of policies of schemes for rural health care
176. A critical analysis of national family welfare programme
177. A study of rural health care planning
178. People's participation in rural health care programme

179. An analysis of monitoring and evaluation system of rural health care
180. A study of rural health care education in India
181. A study of communication in rural health care
182. A study of innovative health programmes by NGOs for rural India
183. A study of communication models for rural development
184. A study of various communication channels in rural development
185. A study of various methods of communication for rural development
186. An evaluation of the contribution of folk media in rural development
187. A study of the emerging electronic information technology for rural development
188. Impact of television in rural development
189. A study of rural extension programme
190. An analysis of the performance of rural extension programmes
191. A study of the government efforts in extension services in rural development
192. A study of community initiatives for extension programme in rural India
193. A study of Panchyati Raj Institutions in rural India
194. An analysis of IRDP programmes
195. A study of the role of CAPART in rural development
196. An analysis of animal husbandary in rural India
197. An analysis of village industries in rural India
198. An analysis of rural health extension
199. An analysis of rural education extension
200. A study of various extension methods
201. A study of different kinds of communication methods used in extension work
202. A study of communication problems in rural development
203. A study of extension management
204. A study of strategies for extension

ignou THE PEOPLE'S UNIVERSITY

ignou
THE PEOPLE'S
UNIVERSITY

Annexure V

DECLARATION

I hereby declare that the Dissertation entitled .....

.....

(write the title in Block letters) submitted by me for the partial fulfilment of the PGDRD to Indira Gandhi National Open University, (IGNOU) New Delhi is my own original work and has not been submitted earlier either to IGNOU or to any other institution for the fulfilment of the requirement for any course of study. I also declare that no chapter of this manuscript in whole or in part is lifted and incorporated in this report from any earlier work done by me or others.

Place :

Date :

Signature

Enrolment No.

Name

Address

ignou
THE PEOPLE'S
UNIVERSITY

ignou
THE PEOPLE'S
UNIVERSITY

Annexure VI

CERTIFICATE

.This is to certify that Mr./Miss/Mrs.....

student of PGDRD from Indira Gandhi National Open University, New Delhi was working under my supervision and guidance for his/her Project Work for the Course MRDP-205. His/Her Project Work entitled

.....

.....

which he/she is submitting, is his/her genuine and original work.

Place :

Date :

Signature

Enrolment No.

Name

Address

ignou
THE PEOPLE'S
UNIVERSITY