MLISc-06

RESEARCH METHODOLOGY

Dr B R AMBEDKAR OPEN UNIVERSITY

Hyderabad

2004

COURSE TEAM:

Contributors:	Unit Nos
Dr K K Naidu NIRD, Hyderabad	1-4
Dr B C Muttaiah NIRD, Hyderabad	6,7
Dr V Chandrasekhara Rao Dr B R A O U, Hyderabad	5,8,11
Dr K Hanumantha Rao NIRD, Hyderabad	9,10,12
Dr K Papamma NIRD, Hyderabad	13
Dr Md Taher ASRC, Hyderabad	14
Dr R S R Varalakshmi Andhra University, Visakhapatnam	15,16
Prof K A Raju NIRD, Hyderabad	*Editor*
Dr V Chandrasekhara Rao Dr BRAOU, Hyderabad	*Associate Editor & Course Coordinator*

Cover Design: CHANDRA

Dr B R Ambedkar Open University, Hyderabad

First Published 1998,2001

This text forms part of an Open University Course.

Further information on Open University programmes may be obtained from the Director (Academic), Dr B.R.Ambedkar Open University, Prof.G.Ram Reddy Marg, Road No.46, Jubilee Hills, Hyderbad-500033 (A P), India.

Lr. No. 1029/Dr. BRAOU/DMP/PTG/F. No.01/Jo.No.16/2004-2005/Dt. 11.11.2004/No of Copies - 150

Printed at : CREATIVE OFFSET PRINTERS, Moosarambagh, Hyderabad-500 036. Ph : 27642538

COURSE-06 : RESEARCH METHODOLOGY

The course 'Research Methodology' conforms to the syllabus of the Master of Library and Information Science (MLISc) offered by Dr B R Ambedkar Open University. As you have done your BLISc we understand that you are familiar with the structure of the programmes and the course material developed by open universities. For the sake of convenience, the syllabus is divided into blocks, each of which comprises a number of units. Each unit generally covers a specific area of the subject. The units are prepared by specialists in accordance with the format designed to enable you to read and understand them without much difficulty. Each unit begins with the contents list and a statement of its aims and objectives, followed by an introduction to the content of the unit. The content of the units are divided into sub-themes and are numbered up to three levels for easy reference. Each unit ends up with Let Us Sum Up, Assignments, References, Recommended Books, and Model Examination Questions.

Research is an inseparable part of human knowledge. Whatever may be the branch of knowledge, research has a social bearing. Spread of knowledge has led to specialisation and that in turn has resulted in the emergence of distinct disciplines of learning. Social sciences are more complex and these are not comparable with other sciences. To understand the social phenomena and the knowledge associated with it one needs to have inquisitiveness and systematic approach. The students of social sciences at the postgraduate level need to have an understanding of various research methods in social sciences. The research process consists of several stages from selection of a problem to reporting of the results. The syllabus of the course has been structured on the basis of sequential steps involved in research process. Various research methods used in social sciences have been described and discussed. Wherever possible examples from library and information science field are included. To provide an understanding about the current trends in LIS research in India and abroad two units have been included at the end.

The specific objectives of the Course are

- to introduce you the concept of Research and its need and purpose;
- to explain you how to select a research problem and formulate the research design;
- to describe the importance of literature survey and review before undertaking a research programme;
- to explain you the method of formulating hypothesis
- to familiarize you with the basic research designs such as historical, experimental and survey; and various other methods - Case Study, Evaluation Research, Delphi method,Content Analysis, Comparative Study, Documentary Research, Sociometric techniques, Construction Typology, etc.
- to provide an overview of data analysis using statistical techniques and computer processing.
- to describe the methods of interpretation and presentation of results and how to prepare a research report;
- to appraise you about LIS research in India and abroad.

CONTENTS

BLOCK - I : INTRODUCTION TO RESEARCH METHODOLOGY

A successful research cannot be planned and conducted in a vacuum. Researchers need thorough understanding of the problem they selected and evaluating it before they venture into it. The recognition of a research problem for study is often a byproduct of careful and extensive reading of appropriate research reports, articles, or other literature. Literature survey provides the researchers to gain insights into the best procedures to solve the problems associated with a research design.

Unit-1 provides definitions to the term 'Research' and explains the need and purpose of research in the context of social sciences.

Unit-2 explains the methods of selecting a research problem and its scrutiny and evaluation. The unit also discusses about formulating a research design and steps involved therein.

Unit-3 deals with survey and review of literature. It explains the need for literature survey in research, locating and searching of various sources of data, including reference sources, archival data and database searching. It also touches on how to review and record the surveyed literature.

Unit-4 provides definition of the term 'Hypothesis' and also lists the sources, qualities and functions of a hypothesis. It also provides an overview of the approaches to formulating research hypothesis.

UNIT - 1: RESEARCH - DEFINITION, NEED AND PURPOSE

Structure

1.0 AIMS AND OBJECTIVES

This unit aims to introduce you the basic aspects of research. It deals with the concept, definitions, need and purpose of research.

After going through this unit, you should be able to

- explain the concept and meaning of research
- describe the features and elements of research
- discuss the relationship between research and library and information services
- explain the influence of universities and government programmes on research
- discuss about the importance of ethics, confidentiality and anonymity in research.

1.1 INTRODUCTION

Research is an inseparable part of human knowledge. Whatever may be the branch of knowledge research has a social bearing. Spread of knowledge has led to specialisation and that in turn has resulted in the emergence of distinct disciplines of learning. Social sciences are more complex and more mysterious and cannot be treated as precisely as other sciences. Even then social scientists have gone a long way in giving as precise a scientific content to social research as possible within the existing frontiers of human knowledge. Since all social sciences are intimate with varied sociological behaviour, a hundred per cent accuracy may be difficult to achieve in social research.

Research in social sciences is the direct outcome of man's urge to understand his society, its nature and working. Social phenomena thus play a crucial role in the direction and depth of social research. Because of the inherent ambiguities and uniqueness of social phenomena, difficulties often arise as to the precise measure of social investigation. Human beings do not operate under controlled conditions; on the contrary they are always under the diverse influences such as environmental, psychological and social, and these influences freely interact with each other and seldom operate in isolation. This interplay of diverse influences makes social phenomena complex.

Social science research cannot be as precise as research in physical sciences owing to greater heterogeneity of social data. However the problem of heterogeneity is sought to be solved by proper stratification and classification of social phenomena. Within each situation, greater homogeneity is ensured.

1.2 RESEARCH - CONCEPT, MEANING AND DEFINITION

Since information is power, people possessing up-to-date information and technical knowledge trend to dominate the people who do not have access to information. In order to fulfill the need for latest information, research activity is not only promoted but also pursued in all fields of endeavour. The increasing research activity is evident from the number of research journals published in the world through which nascent information is disseminated.

The *Ulrich's International Periodicals Directory* gives details of more than 1,20,000 serials over 1,800 abstracting and indexing periodicals and 2,500 electronic publications on-line and over 500 serials available on CD-ROM. Even at the university level many universities in India offer M.Phil and Ph.D. programmes. Many industrial, scientific and technological research institutions like CSIR, ICAR, ICMR conduct regular researches for inventing or discovering new processes, formulations, designs and products leading to new information explosion. All these research activities have implications for libraries and librarians. The researchers visit the library for their information needs and the libraries will have to provide the needed information as quickly and exhaustively as possible. In order to know what exactly research is, how it is conducted the techniques and tools used, how new knowledge is created, it is necessary for the student of Library Science to know the process of research.

In the course of research methodology the student of Library Science may learn many things. Among them are:

1) Elements of careful and systematic enquiry including identification of problem, formulation of hypothesis, collection of data and analysis.

2) Nature and techniques of social research, diverse data needs apart from the relations between different techniques and how these should be employed in a given situation.

1.2.1 Concept of Research

Research is as old as the academic consciousness of human kind. Man has all along tried to look back at his history for better understanding of the evolutionary process leading to the present stage of mankind. He has also endeavoured to reexamine and reinterpret the things he has already gone through. The present has its roots in the past and the seeds of future are sown in the present itself. The urge of human brain to reexamine and to reunderstand things may rightly be called research, at least to begin with.

Research, therefore, has been a integral part of academic pursuits in the past. It has served the two-fold purpose of intellectual sharpening and evolving new theories to explain diverse phenomena through which the mankind has survived in its present form. As man marched ahead, his awareness grew and his curiosity led him to understand his environment through different angles. Slowly and steadily this awareness of man manifested itself in a process, now rightly known as research.

Lack of scientific knowledge in the past led mankind to believe in the existence of supernatural powers and their manipulative capabilities of stars and other planets. Priesthood and authority of religion became supreme. Everything that happened or was thought to happen would solely be attributed to fate and operation of supernatural powers. This blind faith in nature's operation gave a serious setback to the process of research. The process of thinking and understanding of universal phenomena slackened and the urge for research also weakened. Historical records suggest that logic when and wherever it came into conflict with the religious authority or belief, was outrightly rejected and the dogma accepted as the only possible explanation of what had happened or was to happen. Those persisting on scientific analysis and logical deductions would be severely punished for not believing in what the religion or religious men had already said.

With the passage of time, the resistance to orthodox explanations mounted and people with scientific bent of mind started putting greater reliance on facts. The concept of research thus developed into a scientific investigation. The lack of systematic scientific observations and objective methodology greatly restricted the progress of research on the lines of scientific enquiry. Thinking itself was not systematic and many a time the researcher would either over or under generalise the observations and sometimes make the interpretations biased. The first step, therefore, in the development of the concept of research was, to think systematically about thinking itself. With this, the era of logic began.

The early beginnings of logic were in the form of deductive doctrines and this process is attributed to the Greeks. Deductive method derives conclusions from generalizations. Generalizations are results which are universally accepted as facts. Therefore, if something happens to the whole it must also happen to an individual. This is the principal logic of a deductive doctrine. Plato's and Aristotle's political philosophy tended to give more specific shape to human imagination by analysing thoughts based on facts that were inductively ascertained. The interruptions in the process of intellectual activity were followed by a period of renaissance in which the disintegration of feudalism was witnessed.

The cultural awakening emerged and the trade and commerce of Europe also expanded and led to the discovery of new parts of the world such as America etc. The theory of scientific analysis and research spread rather quickly with these discoveries and the human mind witnessed a process of liberation from the dogmas and the ancient pronouncements. Nicholas Copernicus changed the world by doing extensive work on the phenomenon of universe and brought about considerable impact on social thought and human behaviour. In the age of Galileo, Newton and other contemporaries, scientific enquiry strengthened and the intellectual process developed more on the lines which may rightly be called as research.

The concept of research took the shape of social scientific investigation during the medieval times and developed into a full body of intellectual exercises only in the modern age. Research is not the invention of modernists or their monopoly but a culmination of human quest for a better socio-economic and political order.

When we look at research in its modern setting, two things have to be taken note of viz., (a) that it must take stock of changed economic, political, social and cultural institutions; (b) that it must substitute inductive reasoning for deductive one leading to more meaningful and acceptable academic contribution. The concept of research consists:

(1) Avenue of quest, inquisition or investigation;

(2) Specific or well defined domain;

(3) A well laid hypothesis;

(4) Capable of being based upon observable data;

(5) Scientific technique of analysis; and

(6) Logical interpretation of results.

Research should never be treated as a piece of compilation work. Reading a number of books and compiling their material in yet another book is no research. Research is always expected to be something original or a piece of work that advances human knowledge. This may be done in several ways such as re-interpretation of an existing theory, investigations of an unexplored area or development of a new theory etc. The main thing, therefore, is that a researcher should select a field of operation that offers possibility of inquisition or quest. Researcher must be clear in his mind as to what he is looking for.

The domain of research problems must be well specified and accurately defined. It is always useful to select a smaller area of investigation and study it more intensively.

The formulation of hypothesis is equally important for meaningful research. Hypothesis is an assertion that is sought to be proved or disproved. If there is no hypothesis, what a researcher is to prove or disprove remains a big question. A well laid hypothesis also keeps the researcher on the right track and saves his time and effort by not going astray.

Research must be based on facts. Observable data form a sound basis of research. Inductive investigations lend greater support to research findings. For analysing facts a scientific methodology of analysis must be developed and results interpreted logically.

1.2.2 Meaning of Research

Research is composed of two words 're' and 'search' which means to search again or a careful investigation to understand or re-examine the facts or to search for new facts or to modify older ones in any branch of knowledge.

There research is an inquiry into the nature of the reasons for and the consequences of any particular set of circumstances whether these circumstances are experimentally controlled or recorded just as they occur. Further research is interested in the repeatability of the results and in the extension to more complicated and general situations.

Whenever traditional theory is found inadequate to explain the existing phenomena and a novel situation is faced, research originates. Research rejects either old theories or modifies them or suggests new theories.

In technical terms scientific research is systematic, controlled, empirical and critical investigation of hypothetical propositions about the presumed relations among natural and human phenomena. The purpose of research is to discover answers to questions through the application of scientific procedures.

1.2.3 Definitions of Research

Let us examine the definitions of research by well-known scholars.

Francis Rummel - "Research is a careful inquiry or examination to discover new information or relationships and to expand and to verify existing knowledge".

Lundberg - "Research is sufficiently objective and systematic to make possible classification, generalisation and verification of the data observed".

M. H. Gopal - "It is essentially systematic enquiry seeking facts through objective verifiable methods in order to discover the relationship among them and to deduce from them broad principles or laws".

John Best - "Research is a more systematic activity directed towards discovery and the development of an organised body of knowledge".

S. Slessinger and M. Stevenson - Social research is a systematic method of exploring, analysing and conceptualizing social life in order to "extend, correct or verify knowledge, whether that knowledge aids in the construction of a theory or in the practice of an art".

Herring - "Scientific research is a cumulative process; it is also a rejective process especially in the social sciences understanding can be (advanced) not only by gains in knowledge but also by discarding outworn assumptions".

Moser C. A. - "Systematized investigation to gain knowledge about social phenomena and problems, we call social research".

Bogardus, E. - "Social research is the investigation of the underlying process operative in the lives of persons who are in association".

Encyclopedia of Social Sciences - "It is the manipulation of things, concepts or symbols for the purpose of generalising and to extend, correct or verify knowledge, whether that knowledge aids in the construction of a theory or in practice of an art".

According to **Dr. S. R. Ranganathan**, the Father of Library Science in India, 'research' represents a critical and exhaustive investigation to discover new facts, to interpret them in the light of known ideas, theories and laws in order to apply the conclusions to practical purposes.

In simple words research can be defined as "any scholarly investigation in search of truth for facts and for certainties".

In other words, any systematized effort to gain new knowledge is also known as research.

1.3 THE SCIENTIFIC METHOD

Modern science is based on certain assumptions, certain concepts and when we make use of those assumptions and concepts in searching the unknown in a systematic manner this course of action is known as the scientific method. This means searching truth after accepting certain norms which have been given by the modern science. Research also aims at acquiring knowledge. However, when we conduct enquiry under scientific rules, our enquiry may be called the scientific work or activity.

1.3.1 What is Scientific Method ?

Broadly speaking, scientific method means scientific orientation, it is in fact a perspective to look at things or phenomena, as if, they are facts. Method also consists of plurality of components. The way we arrive at our goals becomes our method and its operationalization results in various techniques. Techniques include means and tools we use in collecting facts. In social sciences various techniques are being used in order to arrive at collecting facts.

1.3.2 Features of Scientific Method

1) Scientific method is used with a purpose such as to verify any old theory or to find a new area of knowledge.

2) Scientific method is related to a theory in two ways:

   i) Every scientific method is operationalized on some theoretical base.

   ii) Every scientific method through its research activity gives birth to a new theory or modifies the old one. At least, it may come out with a new hypothesis rejecting the old one.

   iii) Every scientific activity aims at validating or revalidating the established truth. However, it must be noted that all research activities do not lead to generating new knowledge, sometimes a research activity does not even modify the established truth.

1.3.3 Elements/Components of Scientific Method

By elements of scientific method we mean those aspects of method which stand as pillars of study on which practical design of enquiry is prepared. Generally we call these handy aspects as the design of study.

1) **Empirical Investigation Plan:** Empirical research plan is based on collection of data. However, if we say that all researches are always based on a set of empirical data, it is a misleading statement. For example, (bibliographical) library research is not an empirical one. Say you have examined some two dozen papers on academic libraries and have drawn a trend, your observation in the true sense of the term is not empirical. It is an examination of already existing material and not based on presently collected data.

2) **Conceptualization:** Conceptualization is a fundamental element of scientific method. The raw data are useless unless we give meaning to them by finding their causal relations. We observe data, make selective collection, process them and then make abstraction. It may also take some form of codification. So conceptualization actually is a process of making concrete to abstract and abstract to concrete, the latter is a process of decoding. We observe data, establish their cause - effect or causal relationship, note their regularity or frequency of occurrence and then we reflect in terms of inference and conclusion.

3) **Hypothesis:** Hypothesis is an important element of scientific method. In order to find out facts we have to start with some assumptions or propositions. This initial guess is subjected to rigorous testing, examination and verification. If this hypothesis is proved by the collected facts, such proved hypothesis after repeated approval in the long run may suggest regularity and pattern of events and may be accepted as theory and finally as established principle

4) Statistical Precision and Accuracy: The use of statistics in social sciences makes the presentation precise and concrete. By using statistics/mathematics several statements can be put into a few words or figures. However, statistics should be used only when it enhances the quality of work.

5) Objectivity: Objectivity is another basic characteristic of scientific method. If the object is one, even two observers will arrive at the same conclusion provided the observers are working with common goals, methods and tools used by them individually. The extent to which an observer remains objective depends on the nature of the object being studied. Now a days several tools and techniques are available which can help us in maintaining objectivity.

6) Verifiability: A truth is not to be accepted at first encounter. But after repeated observations, it may warrant pause and attention if the same facts occur again and again, whether under experimental condition or in a natural situation. Truth has to be tested and retested. This retest is called verifiability.

7) Expertise: Skill, expertise and efficiency are essential for any method. Expertise is a net result of several components which make out a case for a scientific method.

8) Impartiality: The science, it is said, searches truth without partiality. In the course of our search for truth, if we are not impartial it means, we are deceiving science. If we ignore personalised experience, prejudices and convictions, the goal of impartiality could be attained to a greater extent.

1.4 GENERAL MODEL OF SCIENTIFIC ENQUIRY

You have learned how the scientific method helps in conducting a systematic and valuable research. Busha and Harter (1980) explained a general model of scientific inquiry. The steps illustrated are to be followed in the order shown in the model. Let us briefly describe it. (For more detailed discussion read Unit-2).

State General Problem: The research problem may originate as a result of intuition, or as an inference from previously derived theory. Researchers need to understand the problem and its context before proceeding further.

Conduct of Literature Search: For understanding of the research problem and its facets, scanning of the relevant professional and scholarly literature is necessary. This helps to delineate the problem as well as to formulate the methodologies to be followed.

State Specific Problem: Literature survey lead us to identify the limitations of the problem, to recognise pertinent independent and dependent variables associated with it and also help us to outline the specific problem clearly and precisely.

Design of Methodology: There are many methods which can be employed in conducting research. Experimental, historical, and survey methods are the basic experimental formulations. Depending on the problem investigators have choose an appropriate method.

Data Gathering and Analysis: It relates to proper design and use of instruments to gather data. The data gathered is analysed and then interpreted.

Report Results of Study: The results of the study are recorded and reported.

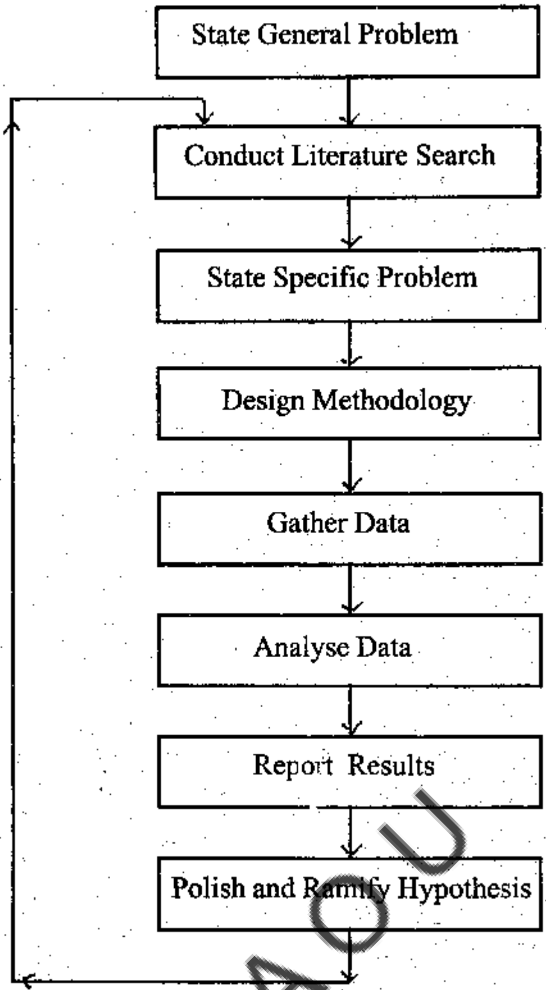


Figure-1 : General model of the scientific method of enquiry
Source : Busha and Harter. *Research methods in librarianship*, p.18

1.5 TYPES OF SOCIAL RESEARCH

Any attempt to classify types of social research is merely a mechanical exercise. Different researches are conducted with different objectives and definitely with different methods, tools and techniques. However, we should note here that one type of research does not stand opposite to the other. Moreover, one research type is not entirely different from the other. Rather features of different types of researches are intermixed or we can say that they have transcendental characteristics. The main types of research are as follows:

(1) *Pure/Basic/Fundamental/Theoretical Research*: Here development of new knowledge, theories are aimed at.

(2) *Applied Research*: As the name suggests, application of known theories and knowledge to test and validate, the operational fields or populations.

(3) *Action Research*: It aims at providing workable solutions to the given problems by direct action.

(4) *Empirical Research*: It consists of collection of new data with the involvement of researcher personally.

It is a very difficult job to classify social research in a watertight compartment. We can say that classification of research work is based on quality of knowledge itself. Every research work is related to a new knowledge. The difference in evidence may be in terms of novelty and that too relatively. Therefore, the above mentioned types of research work are to be understood in terms of relative difference in perspective and newness.

1.5 RESEARCH *VIS-A-VIS* LIBRARY AND INFORMATION SERVICES

Library being a store-house of knowledge, it is an essential part of academic life. Any researcher who has to identify the problem of research has to undertake literature search as a first step. He is compelled to go to library to read books and other material relevant to his field of interest to identify the research problem. Therefore, the implications of effects of research will influence the development of libraries.

Thus libraries become an integral part of research thereby implying that the Librarian has to give utmost attention to the type of material he or she should acquire, the organisation to be adopted and the dissemination to be pursued. It is also essential for the librarian to know in detail, the information requirements of the researchers and what exactly research is, how it is conducted, what techniques and tools they use for conducting research and how new knowledge is generated. Such a knowledge will enable the librarian to serve the researcher in a much better fashion.

Similarly, knowledge about the numerous facets of library and information science can be obtained by identifying the gaps and problems, thinking of possible solutions and testing the solutions by means of accepted scientific methodologies. In a library situation a researcher should have a clear grasp of the behavioural aspects of the user, nature of his demands with respect to the focus of the library, types of collection and above all the type of information products and services the library provides for the benefit of its users.

In the light of the continuous advancements taking place in information technology, the researcher in the information science should be dynamic enough to keep track of the fast changing developments so that his or her research would not become outdated. Thus a modern researcher who is concerned with the marketing of library and information services to the benefit of library users should be aware of the latest trends and developments taking place in the field of his inquiry like his users, his library collection and his services thereby enhancing his effectivity.

1.6 INFLUENCE OF UNIVERSITIES AND GOVERNMENT ON RESEARCH

1.6.1 Influence of Universities

Why is social science research important today in India? To ask this question is almost to answer it. A few points may, however, be noted. First, the current annual production of Ph.D theses in India is about 1200. Increased interest in post-graduate study is evident in all the Indian Universities today. The number of social science research theses and monographs will obviously grow with the general increase and may very well, form a higher ratio of the total in the years immediately ahead. The expansion of university research - by staff as well as research scholars places a broader obligation upon all research libraries (not merely university libraries) to increase their resources.

1.6.2 Influence on Government Programmes

A second reason why social science research is important to India today lies in the economic and social policies and programmes of the government. The research programme of the Planning Commission is an adequate demonstration of this point. The Planning Commission is indeed utilising university personnel for many of its studies, as well as conducting much research on its own account. Many other government agencies, Centre and State, are also engaged in a wide variety of social science research. It is unlikely that this emphasis on research in many areas of the social sciences will decline in the foreseeable future. It becomes important, therefore, to equip the relevant libraries, universities, governmental and specialised research organisations, with the necessary research resources to make the results of future research as effective as possible.

1.7 ETHICS AND RESEARCH

Professional organisations in the social sciences have developed guidelines for ethical conduct of research with human participants. American Psychological Association and American Anthrpological Association have come out with a code of ethics for researchers in their fields. The ethical code consists of defining certain responsibilities for preservation and protection of fundamental rights. While undertaking these responsibilities every effort is made to protect the welfare of research participants and do not permit misuse of results by others. While demanding, for themselves freedom of inquiry and communication, the researchers should accept the responsibilities such as, competence, objectivity in the application of skills and concern for the best interests of the clients, colleagues, students, research participants and society.

1.8 CONFIDENTIALITY AND ANONYMITY

Two methods used by researchers to protect participants in behavioural studies are confidentiality and anonymity.

1.8.1 Confidentiality

'Confidentiality' means that the respondents' identity is known to the investigator but protected from public exposure. The researcher keeps any identifying information out of published reports. Confidentiality is particularly important when people's statements or actions would cause them some embarrassment if they become known. It is best for the researcher to be on the safe side regarding confidentiality, since it is difficult to predict how people's answers might be interpreted or used by others. Instead of identifying organizations or cities by name, a general description, for example: a middle sized industrial city in the Northeast or a fictitious name such as Blue city or work town is used instead.

The best way to ensure that the people you have interviewed or observed will not be embarrassed by your data is to remove identifying information, such as names and addresses, as soon as the data are tabulated. Some researchers use a special code at the time of the interview so that no names or other identifying information are in the researcher's files.

1.8.2 Anonymity

'Anonymity' means that the researcher does not know the identity of the participants in the study. The best protection for people in an observational study is anonymity. It makes more sense to observe people in a public park or courtyard without knowing their names than to ask each person to sign a written consent form. Anonymity is also a useful safeguard in questionnaire and public opinion research. Most questionnaire studies do not ask people to sign their names. Often this is emphasized by telling respondents **not** to sign their names.

1.9 LET US SUM UP

Let us recapitulate what has been discussed so far in this unit.

* Research has been defined by various scientists in different ways. Research is essentially a systematic enquiry to discover new information or to discover the relationships and to expand and verify the existing knowledge.

* Research is a careful and systematic enquiry including identification of problem, formulation of hypothesis, collection of data and analysis.

* Scientific method aims at validating or revalidating the established truth. The components of scientific method are empirical investigation plan, conceptualisation, hypothesis, statistical precision and accuracy, objectivity, verifiability, expertise and impartiality.

* The main types of research are i) Pure/ Basic/ Fundamental/ Theoretical research, ii) Applied research, iii) Action research, and iv) Empirical research.

* Research in library and information science need to concentrate on the problems and find out solutions to them, especially in identifying the nature of user's needs and shaping the information products and services to suit them.

* Researchers have to follow the guidelines for ethical conduct of research developed by the professional organisations. The two methods used by researchers to protect the participants in behavioural studies are confidentiality and anonymity

1.9 REFERENCES

BEST, John, W. *Research in Education*. New Delhi: Prentice-Hall, 1977-78.

MOSER, C.A. *Methods in Social Investigation*. London: Heinemann, 1961.

RANGANATHAN, S.R. and Girja Kumar. *Social Science Research and Libraries*. Bombay: Asia Publishing House, 1960.

REDDY, C.R. *Research Methodology in Social Sciences*. Delhi: Daya Publishing House, 1987.

SADHU, A.N. and Amarjit Singh. *Research Methodology in Social Sciences*. Bombay: Himalaya Publishing House, 1983.

SINGH, R.A.P. *Methods in Social Research*. Jaipur: Printwell Publishers, 1990.

SOMMER, R. and Sommer B. *A Practical Guide to Behavioural Research - Tools and Techniques*. Oxford: University Press, 1986.

VERMA, R.K. and Gopal Verma. *Research Methodology*. New Delhi: Commonwealth Publishers, 1989.

WHITNEY, F.L. *The Elements of Research*. Bombay: Asia Publications.

YOUNG, P.V. *scientific Social Surveys and Research*. New York: Prentice-Hall, 1977.

1.10 RECOMMENDED BOOKS

BOGARDUS, E. *Sociology*. New York: MacMillan, 1954.

GOODE, W.J. and Hatt, P.K. *Methods in Social Research*. New York: McGraw Hill Book Company Inc., 1952.

KAPLAN, A. *The Conduct of Inquiry - Methodology for Behavioural Science*. Jaipur: Sachin Pub., 1980.

KERLINGER, F.N. *Foundations of Behavioural Research*. Japan: CBS Publishing, 1986.

MANN, H.P. *Methods of Sociological Enquiry*. Oxford: Basil Blackwell, 1968.

OGBURN, W.F. and Nimkoff, M.F. *A Hand Book of Sociology*. New Delhi: Eurasia Publishing House, 1966.

SOROKIN, P.A. *Sociological Theories of Today*. London: A Harper International Edition, 1966.

1.11 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) Define research. Discuss its objectives and importance to the society.

2) Describe briefly the concept of research.

3) What is the meaning of research ? Discuss its characteristics of research ?

4) What are the elements of scientific method? Discuss them in brief.

5) Explain various types of research methods ?

6) What is the relevance of research and benefits to LIS studies ?

7) What are the differences between social science research and physical science research ?

II. SHORT NOTES

a) Confidentiality in Research

b) Anonymity in Research

c) Ethics of Research

d) Empirical Research

UNIT - 2: RESEARCH PROBLEM AND RESEARCH DESIGN

Structure

2.0 AIMS AND OBJECTIVES

The unit aims to introduce you to the approaches of selecting a research problem and its scrutiny and evaluation. The unit also deals with the ways of finding an appropriate research design to work on the selected problem.

After studying the unit, you should be in a position to

- explain how to select a problem for research
- describe the methods of scrutinizing and evaluating the research problem before starting the work
- discuss about formulating a research design.

2.1 INTRODUCTION

The term problem is derived from the Greek word "**proballein**" which means anything thrown forward; a question proposed for solution; a matter stated for examination. R.W. Woodworth defines problem as "a situation for which we have no ready and successful response by instinct or by previously acquired habit". In brief, problem means the state of situation one could not find a solution. Thus, necessity is the mother of innovation. In other words, without a problem, research cannot proceed because there is nothing to proceed from and proceed forward.

For a new research worker one of the arduous tasks is to decide upon a suitable problem wherein his most valuable time, money and energies are to be invested in the process of research. It is an area in which **vision** plays an important part.

Formulation of a research design is a very crucial problem and a pivotal point for the success of a research programme. Research design is a plan of activities that, in the final analysis, would contribute to a research study. It is the logical and systematic planning and directing of a piece of research. Normally while formulating a research proposal, the whole scheme of things is categorised in a three-fold fashion as **plan, structure and strategy** of research study.

Plan includes the major framework of activities that the researcher wishes to accomplish under the banner of a particular research study. It includes every thing the investigator will do from writing the hypothesis and their operational implications to the final analysis of the data.

The second part of a research design, the **structure**, is the outline and the scheme of integrating the variables of investigation into a cohesive piece of analysis.

The third part of a research design, the **strategy**, refers to the specific methods which are to be used for obtaining the research data, analysing it, verifying its reliability and interpreting the results. As a matter of fact it refers to the methodology used for conducting the research investigation.

Research design in a way tells us what observations to make, how to make them and how to analyse the qualitative representations of the observations. There may not be such a thing as a single or correct research design. Research design represents a compromise dictated by the many practical considerations that go into social research. Research design is not a highly specific plan to be followed without deviations, but rather a series of guideposts to keep one headed in the right direction.

2.2 RESEARCH PROBLEM

Selection of a research problem is a very important job for a researcher. The range of potential topic for socio-economic research problem is as broad as the range of socio-economic behaviour itself. Socio-economic life in modern times is full of innumerable problems which provide a researcher ample scope for identification of a research problem. The problem begins to appear when the need for its expansion or solution is urgently required. The existence of unsolved problems is the first symptom for locating a research problem.

As is usually the case every problem may not be researchable and, therefore, if the researcher has not been able to select the problem properly, his efforts may not yield the desired results and in some cases it may lead to frustration and desperateness in the researcher. Selection of a suitable problem for research is in many ways the most difficult task. On the face of it, the selection of a problem may appear to be simple but one envisages the operational difficulties of putting a design into effective research, he may find that the selection of problem is a very complex phenomenon and warrants several considerations.

2.2.1 Research Objective *vis-a-vis* Research Problem

There is no foolproof rule that will guide the researcher in formulating significant questions about a given research area. There are many pitfalls to be avoided and there are many students who have failed to complete their research, not because they were lay or badly organised, but because their topic was not suitable for research. Probably the most common mistake is to choose a topic that is too large at the level at which the student is studying/working. A topic which may be suitable for a doctoral thesis may not be so for a dissertation for a post-graduate degree, which has to be completed in less than a year.

Obviously, extensive social research, involving large samples and presenting considerable problems in terms of analysis, cannot usually be undertaken by a single research scholar. Some topics are simply not researchable at any level. This may be because the sources do not exist or difficulty may arise because no means have yet been devised for investigating that particular field.

The selection of a research problem, firstly depends on whether research is being conducted as a requirement of a degree or for academic interest. If the research is only to fulfill some requirements for obtaining a degree, the problem may be more specific and limited in scope and may offer itself for completion within a specific time. If the researcher fails to select such a problem as would be completed in a reasonable amount of time, he would often run into difficulties and may go without the desired degree. On the other hand, if the research is of academic interest, time may not be a binding factor. Many a time the entire life-time of a researcher goes towards exploring something new through his research efforts. He does not want any research degree but only does the research with the sole objective of enriching his personal knowledge as well as that of the general academic world.

Selection of the problem as a requirement for degree further depends on the level of research or learning. Many a time, research problems are selected at Masters Degree level, M.Phil or Ph.D. levels, but it is obvious that the levels of learning at these stages are different and, therefore, demand different treatments to research problems. Usually the problems selected at initial stages such as Masters Degree level constitute an exploratory or a learning process. At the next stage such as M.Phil level it aims at the consolidation and improvement in the earlier research efforts. At Ph.D level, it may aim at masterly application of advanced research methods and seek to make significant contribution to advance the existing knowledge.

2.2.2 Precautions for the Selection of a Research Problem

1) A researcher should select a topic in the field in which he has familiarity and deep understanding. According to Tyrus Hillaway "If a scholar is familiar with his field and knows what research has already been completed in it, he will also know something of the many problems that remain".

2) A research problem should be time-bound and cost-bound and the selection of the problem should be made accordingly.

3) Selection of the problem should be as strong as the formulation of a building and it requires enormous reading, reference and discussion.

4) Selection of the problem should not be made in a hurry but it should be carefully considered and objectively selected. Spending considerable time in selecting a topic will not be a waste at all, on the contrary it will create confidence in a researcher. Best John has rightly remarked "It is a serious responsibility to commit oneself to a problem that will inevitably require much time and energy, and upon which so much academic significance is based".

5) While selecting a topic the researcher need not be over enthusiastic. He must be balanced and ready to face the unexpected twist or unreasonable delay, in completion of data.

6) The capacity of the research problem should be within the physical resources available to a researcher.

7) The research problem must be worth studying. It must have potentiality to stand as a research problem.

8) The study must have social relevance. It must be capable of drawing the attention of the policy makers, experts and academicians who are familiar with the subject.

9) There must be a felt need for research over the problem. As far as possible it should be untouched by other researchers or even if touched must be in need of further elaboration and exploration.

10) A research problem must come out with pragmatic solution to the issue.

11) As far as possible a social research problem must be uptodate or relevant to the current social happenings, it should not be sterile.

After demarcating the area in which one would like to do research, he must at least possess two characteristics, viz. (a) interest, and (b) competence. To do sustained work on a problem and to put one's intellectual effort fully into it is a difficult task and will turn out to be a drudgery unless one has a genuine interest in the problem itself.

Interest in a subject develops due to various reasons. For example, an agricultural background student may have interest in some area related to agriculture. On the contrary, he may avoid agriculture altogether. Academic interest in a particular branch of subject may be the result of reading a good book or because the teacher dealt with it extremely well. Interest and inclination must be matched by competence before one can decide on an area of research. Competence can certainly be cultivated and learning new techniques is part of any research programme.

A common tendency in the approach to research is to feel that it must result in some extraordinary '**findings**'. Ambition is good, but too much ambition is dangerous because it may lead to the rejection of manageable problem in search of a brilliant and path-breaking contribution.

Once the main area of research is decided upon, the next step is to narrow it down to a topic. The researcher cannot be a passive recipient of suggestions from others but should know the field himself. This can be acquired by reading literature. After the survey of literature, the area of research problem has to be narrowed down. It means select an area on which one can get the guidance of a few competent articles or books, but not of too many. In social science research this '**rule**' is utmost important. In brief, identifying a research problem is an active but subjective process, and possibly the most important aspect of research. The research is a process, process of extraction and expansion; and of arriving of at the end '**reforming**'.

2.2.3 Sources of Research Problem

The research problem may be selected from the following sources:

1) Theory of one's own interest
2) Daily problems
3) Technological changes
4) Unexplored areas
5) Library
6) Discussion with supervisor

1) A researcher may select a problem for investigation from a given theory in which he has considerable interest. In such cases a researcher must have thorough knowledge of that theory and should be sufficiently capable of exploring some unexplained aspects or assumptions of that theory.

2) A research problem can also be selected on the basis of daily experience of the researcher. Day-to-day problems constantly present something new and worth investigating. A researcher by his sharp intellect may knit his daily experiences into a research problem.

3) Technological changes in our fast changing modern society are constantly bringing new problems and opportunities for research. The impact of changed technology on the existing socio-economic set-up creates interest in the research and tempts him to undertake such studies. Newly established institutions also provide an opportunity for the researcher to evaluate their functioning and impact on the socio-economic condition of the society.

4) Research problem can be both abstract or of applied interest. A research problem may also be selected from the areas which have not been explored so far. Such unexplored areas may be theoretical or empirical in nature, eg. an economic analysis of hill farming system in a region where such an investigation has not been undertaken earlier is a useful place of research. In a similar fashion many geographical regions about which no knowledge exists can easily be undertaken for research purposes.

5) A researcher may choose a research problem with the help of books, journals, periodicals, research reports, research journals, professional journals, bibliographies and indexes which constitute the flesh and blood for social research. Not only will the scholar himself notice a number of unsolved problems and unexplored areas during the course of his general study in his field, but also he can lay his hands on summaries and reviews of research that point out their gaps. Periodicals published for the scholar/reader often contain long lists of topics upon which additional research is needed. Also from published articles and reports of research one can gain ideas and suggestions for subjects requiring further study.

6) Sometimes the researcher while discussing his interest with his proposed supervisor may come across a problem that can be researched by the investigator.

2.2.4 Preliminary Survey of Literature for Topic Selection

A preliminary survey of the literature on the topic should be carried out to find out the possibility of original contribution to the concerned area of knowledge. **A study of the current literature** in your chosen field will indicate the problems that are being investigated and will suggest further problems for investigation. The researcher should be familiar, too, with recent doctoral studies in his field of interest. Special assignments, term papers, dissertations, and theses usually conclude with suggestions for further research.

Generally it is very difficult to know about the current research work in a specific area. One may institute inquiries at all important places known for research on one's proposed or intended topic. The possibility of such project topics being under way may be indicated by reading professional journals and latest papers, discussions in workshops, conferences and seminars. *Encyclopaedia of Social Science, Encyclopedia of Library and Information Science, Dissertation Abstracts International* and similar publications are rich sources for problem seekers.

The research fellow, lecturer, reader or professor who is active in research is usually a **source for research problems**. Generally, as problems are investigated, more problems emerge for investigation. Specialists or experts who are recognised as research guides normally provide clues to the area of research work at different centres. The more one knows about a particular field, the more able one is to detect gaps in it and to recognise problem areas that require investigation. It is the ability to detect problems that the research scholar must develop, for every thesis or dissertation may potentially shed new light on the solution to a particular problem.

Once the field of interest is narrowed down and several problems/areas for possible investigation are identified then the following questions should be asked about the topic for its **feasibility.**

Does the topic really interest you?

Is the problem a significant one ?

Is there adequate supervision/guidance?

Can the topic be completed in the required time?

Are subjects available to experimental treatments and testing?

Are the necessary equipments and adequate library facilities available?

2.2.5 Review of Relevant Literature

Once a topic has been decided upon, it is essential to review all relevant material which has a bearing on the topic. In fact, review of literature begins with a search for a suitable topic and continues throughout the duration of the research project work. Since a research report, either dissertation or a thesis, is supposed to be a study in depth aiming at contribution to knowledge, a careful check should be made that the proposed study has not previously been carried out.

The main reason for a full review of research in the past is to know the outcome of those investigations in areas where similar concepts and methodologies had been used successfully. Further, an extensive or even exhaustive process of such review may offer vital links with the various trends and phases in researches in one's area of specialisation and familiarise him/her with the characteristic precepts, concepts and interpretations, with the special terminology, with the rationale for undertaking one's proposed investigation. Each research has a formal and systematic structure which generates appropriate norms for evaluating its outcome. In this connection, a review of previous related research projects will help the research to formulate a satisfactory structure for his project.

A teacher who lets students learn by doing the project work should also feel proud that his students are doing things on their own. After all, learning is worthwhile and meaningful only if the learner is able to use the knowledge to solve practical field problems and is able to operate on his own.

2.2.6 Topic Suggested by the Guide/Choice by oneself

There are basically two ways in which one arrives at a topic. Either the research guide suggests it or one arrives at the choice by oneself. In the former case, the chances are that a topic suggested by one's guide will be suitable for research and that its scope will be sufficiently limited. On the other hand, the main pleasure to be derived from research is to immerse oneself in the topic in which one is really interested; if the researcher finds the subject boring then the special study will become a millstone around his neck. For this reason, therefore, it is probably better to choose the topic one self. It is important for the student to understand that the ultimate responsibility for the choice of a topic is his and that although the research guide may offer advice, he will expect the student to make the final decision.

Thus, a student by choosing a topic of his liking can pursue his own interest to the farthest limits. That is, if one is interested in the topic, then there is the opportunity to spend a long time on something that is continuous source of pleasure. It is satisfaction that is a continuous source of pleasure. It is satisfying to be able to study a subject in depth instead of having to break off at the very point when one becomes interested. At the end of the research, one should have a fair claim to be a specialist in one's field and this tends to give one confidence.

2.2.7 The Technique of Skimming

All skillful readers use the well known reading techniques of skimming. Everyone develops his own particular order and techniques, but skimming is absolutely essential if the student is not to be swamped in a sea of print. The technique of skimming is explained below in seven steps.

1) Flip through the book from cover to cover and the reader will get some impression of the book;

2) Turn to the title page and look on the back of the title page also to find out its relevancy and recency respectively.

3) Turn to preface and read it. A preface usually explains the author's purpose in writing the book and the particular approach he has taken. The approach may promise to be useful to one's research.

4) Skim the table of contents to get some idea of the organisation of the book. If one finds a section which is right on his topic, read that section of the table carefully;

5) Flip through the book again, pausing to read the opening paragraphs of some chapters. By now the reader should be able to pick some which are more likely to be of use to him than others;

6) While going through the book, note carefully any foot note, questions etc, which are more likely to be helpful for indepth study of the research area;

7) Finally, if the book has an index (at the back), check in it for specific items which are relevant to the topic of research. (For more details see Unit 4).

2.2.8 Hypothesis

When a researcher observes known facts and takes up a problem for analysis, he first has to start somewhere and this point of starting is the hypothesis. In other words, one has to proceed to formulate tentative solutions as soon as the problem to be tackled is finalised. These proposed solutions or explanations constitute the hypothesis which the researcher proceeds to test on the basis of facts already known or that can be made known. The tentative explanations or solutions are suggested to us by some thing in the subject matter and by our previous experiences.

2.2.9 Role of the Guide in Research Project Work

1) Guide the students to select appropriate projects

2) Always be willing to cooperate and coordinate the work

3) Provide necessary resources and obtain help from various agencies and people both in the institution as well as outside as necessary.

4) Encourage the students to aim at completion. If students encounter difficulty or frustration persuade them to proceed by boosting their morale. Sense of fulfillment is an important gain, and an incomplete project can shatter a student's self-confidence and self-respect.

5) Encourage and help students by appreciating their efforts now and then.

6) Be careful about lazy and indifferent students. They can, if not properly dealt with, influence others too.

7) Have a fair evaluation plan and inform students.

2.3 RESEARCH PROBLEM - SCRUTINY AND EVALUATION

Before discussing about scrutiny and evaluation of research problem, let us understand what are the criteria for a good research problem.

2.3.1 Criteria for a good Research Problem

A good problem selected by the researcher should conform to the following criteria.

1) **Relation between variables Expressed:** The problem should express a relation between two or more variables.

2) **Clarity and unambiguity:** The problem should be stated clearly and unambiguously.

3) **Empirical Verification:** The problem should be such as to imply possibilities of empirical testing. This means not only that an actual relation as stated, but also that the variables of the relation can somehow be measured.

4) **Availability of guidance:** Every research activity needs the patronage of a guide and the approval and sanction of a competent authority.

5) **Level of Research:** It is another criterion to help in the selection of a problem. The nature and scope of a study will be determined in the light of levels like Master's Degree, M.Phil. Degree and Ph.D. It may simply be an action research or a research to produce a research paper or an experimental project.

6) **Experience and Creativity:** Good research problems stem from a clear understanding of the theoretical, empirical and practical aspects of the subject derived from personal experience and from a thorough review of the literature. Conversely, lack of familiarity with the subject is almost sure to result in a poor choice.

7) **Novelty:** It should be sufficiently original so that it does not involve objectionable duplication. Originality is the basic credit point of any research. Ignorance of prior studies may lead a researcher to spend time on a problem already investigated by some other worker. Moreover the study should employ the latest data. While originality is an important consideration, the fact that a problem has been investigated in the past does not mean that it is no longer fit for study. There is constant need for verification of the findings of previous investigations, using newer and better devices and procedures. There is also a need for the testing of former findings under changed cultural conditions.

8) **Interesting:** The problem should be interesting for the investigator himself. If he is not interested in it, he will not be able to face and overcome the obstacles which come at every step in research. He should have a strong inherent motivation for it. If it seems to him dull and boring, there is little hope that he would do justice to it. A major reason for encouraging a researcher to select his own research problem is that the results should prove better in terms of the growth of the investigator and the quality of his work. His interest should be purely intellectual and should not be there only for a reward, material benefit, advancement in position, increased authority and so forth.

9) **Importance:** It should be significant enough and involve an important principle or practice. If it is not worthwhile, if it neither adds to knowledge nor leads to any improvement in the current practices, it would be in vain.

2.3.2 Formulating and Stating the Problem

After the problem has been selected, it must be formulated and stated clearly. If it is to serve as a guide in planning the study and interpreting its results, it is essential that the problem is stated in precise terms. The type of statement to be employed depends on the preference of the worker and the nature of the problem. There are two different ways of stating a problem: (i) posing questions and (ii) making declarative statements.

One may choose any of these ways keeping in mind that the question form has an advantage in sharpening and focussing the issue, but the declarative form is perhaps more common, and both the ways may be combined easily in an initial statement.

2.3.3 Definition of the problem

To define a problem means to specify it in detail and with precision. Each question and subordinate question to be answered is to be specified. Sometimes it is necessary to formulate the point of view or educational theory on which the investigation is to be based. If certain assumptions are made, they are explicitly noted. It implies the separation of the problem from the complex of difficulties and needs. It means to put a fence around it, to separate it by careful distinctions from similar questions found in related situations of need.

It is important to define and elucidate the problem as a whole and further define all the technical and unusual terms employed in the statement. By this the research worker removes the chance of misinterpretation of any of these crucial terms. The definition helps to establish the frame of reference with which the researcher approaches the problem..pa

2.3.4 Delimitation of the Problem

It is a statement of the limits or scope of the investigation. It will determine the boundaries of the project in hand. The delimitation will mention the geographical limits of the study i.e. whether the study will be covering a single town, a district, a region, a state or a country. Again it will be important to mention as to how many subjects will constitute the sample of the study and how that will be distributed over the institutions, geographical areas or time intervals.

2.3.5 Justification of the Problem

The urgency and worthiness of the project have to be justified. It has to convince the readers about the significance of the investigation. This step would prevent wastage of research effort on unimportant, trivial, superficial or insignificant problems. The problem should be broad enough to provide an investigation of real significance.

2.3.6 Evaluating the Problem

Before the proposed research problem can be considered appropriate, several searching questions should be raised. Only when those questions are answered in the affirmative can the problem be considered a good one.

1) Can this type of problem be effectively solved through the process of research? Can relevant data be gathered to test the theory or find an answer to the problem under consideration?

2) Is the problem significant? Is an important principle involved? Would the solution make any difference as far as theory or practice is concerned? If not, there are undoubtedly more significant problems waiting to be investigated.

3) Is the problem a new one? Is the answer already available ? Ignorance of prior studies may lead a researcher to spend time needlessly on a problem already investigated by some other researcher.

4) Is research on the problem feasible? After a research project has been evaluated, there remains the problem of suitability for a particular researcher.

(a) Am I competent to plan and carry out a study of this type? Do I know enough about this field to understand its significant aspects and to interpret my findings?

(b) Are pertinent data accessible? Are valid and reliable data-gathering devices and procedures available?

(c) Will I have the necessary financial resources to carry on this study? What will be the expenses involved in data collection, printing, test materials, board and lodging, investigators help and computer.

(d) Will I have enough time to complete the project ? Will there be time to devise the procedures, select the data gathering devices? Gather and analyse the data, and complete the research report ?

(e) Will I have the courage and determination to pursue the study in spite of the difficulties and social hazards that may be involved? Will I be willing to work aggressively when data are difficult to gather and when others are reluctant to cooperate?

2.3.7 Errors in Selecting a Problem

While selecting a problem for research the researcher who is a beginner may commit certain errors which are as follows:

1) Selecting a broad field of a specific aspect. For example, the researcher may take up the study, "the Education in India" or "Agricultural Development". The study of these is difficult because they are vast areas.

2) Assuming that the investigation is impossible. For example, "The Effects of mothers on economic gains of their offspring".

3) Narrowing a problem. The problem may be narrowed to too small an area which may be insignificant from research point of view. For example, "The History of A Village School Library".

2.4 RESEARCH DESIGN

Now-a-days there is an increasing tendency on the part of students and teachers to do research. The selection of a research topic is the first and foremost problem that the prospective researchers face. Quite often, in the initial stages, one has very vague and hazy motions on what he would research upon. Often in the initial stages, the canvas is very wide. It is necessary to delimit the area of investigation in view of the limited resources of an individual research worker to handle adequately and intensively large canvas. This will facilitate intensive enquiry into the problem so that well founded conclusions can be drawn on the phenomenon under consideration. All this requires proper planning. But one of the most common defects noticed in research after the selection of the problem is lack of smooth flow of research work.

2.4.1 Formulating a Research Design

As soon as the problem is selected, one hastily leaps forward with the work of collection of data or writing the plan for the thesis or dissertation and even sometimes writing the initial introductory chapters in the thesis or dissertation. This shows that the researcher is not denoting considerable time in thinking about his future work. The smooth scaling in the field of research is possible only when the researcher thinks considerably about the problem under study and about the various aspects of the problem.

It is because of such importance, the research design or plan occupies a key position in the research work. An attempt has been made in this section to explain the steps that are normally necessary in formulating a research design. Those are not rigid but flexible and can be adopted to suit the problem under investigation.

2.4.2 Meaning of Research Design

Let us approach the meaning of **design** and understand what all is involved in **designing**.

An architect is said to **design** a building. In the process of **designing** he considers each decision that will have to be made in constructing the building. Bearing well in mind the purpose for which the building is to be used, the architect takes decisions such as how large the building will be, how many rooms it will have, how these rooms will be approached, what building materials will be used and so on. The designer does all this before the actual construction begins. On the basis of this paper picture, he can correct mistakes and make improvements before the construction actually starts.

In the same way a soldier prepares a strategy before launching an attack. An artist makes a design before he executes his ideas. Any prudent man makes a plan before he undertakes work. So also the researcher makes a plan of his study before he undertakes his research work. This will enable the researcher to save time and resources. Such a plan of study or blue-print for study is called a Research Design or Research Strategy.

A research design or model indicates a plan of action to be carried out in connection with a proposed research work. It is a tentative plan which undergoes many modifications as the study progresses. It presents a series of guide posts to enable the researcher to keep track of his actions and to know that he is moving in the right direction in order to achieve his goal.

A research design should be based more or less on methodology. It should be made once the topic and problem for research have been selected and formulated objectives have been properly framed.

2.4.3 Definitions of Research Design

Let us examine some of the definitions of **Research design**:

Claire Seltiz *et al* states, "Research Design is a catalogue of the various phases and facts relating to the formulation of a research effort. It is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure".

Paul Green and Donald Tull state, "A research design is the specification of methods and procedures for acquiring the information needed. It is the overall operational pattern or framework of the project that stipulates what information is to be collected from which sources by what procedures. If it is a good design, it will ensure that the information obtained is relevant to the research questions and that it was collected by objective and economical procedures".

According to Bernard Phillips, "The research design constitutes the blueprint for the collection, measurement and analysis of data. It aids the scientist in the allocation of his limited resources by posing crucial choices: Is the blue print to include experiments, interviews, observation, the analysis of records, stimulation or some combination of these? Are the methods of data collection and research situation to be highly structured? Is an intensive study of a small sample more effective than a less intensive study of a large sample? Should the analysis be primarily quantitative or qualitative".

To quote P.V.Young "It is the logical and systematic planning and directing of a piece of research. The design results from translating a general scientific model into varied research procedures. The design has to be geared to the available time, energy and money; to the availability of data, to the extent to which it is desirable or possible to impose upon persons and social organisations which might supply the data".

2.5 STEPS IN PREPARING A RESEARCH DESIGN

A good research design consists of the following components. They are:

2.5.1 Components of a Research Design

Title of the study:

- should be as brief as possible
- should be as precise as possible
- should project the scope of the problem in generalised terms

The Importance of the Study:

Analysis of the Problem with Implications for Policy and Programme Implementa

Here an analysis of the problem in the context of the academic, socio-cultural or politi situations should be made in order to indicate its relevance and the need for such a study contribute in whatever way to the existing thinking on the theme or solution to the problem o possibility to develop a theoretical framework.

Attempt can also be made here to indicate the implications of the study either for polic y formulation or programme implementation.

Review of Literature:

A brief survey of studies undertaken earlier has to be given. This is meant to fill the gaps in the contribution of knowledge. Such survey would entail a study of a relevant literature. The review has to be done in such a manner that when the proposed study is completed, it would add its own contribution to the quantum of knowledge in the subject concerned. Earlier studies are not only reviewed but also critically appraised. This type of review would pinpoint the stage from where further research is called for.

Objectives of the Study:

The main purpose of the research study is to lay down the objectives precisely. It may be mentioned in the form of questions or an explanation to a particular issue or phenomenon. In other words, the objectives should be worded in a lucid and precise form. For example, if the investigation is on rural leadership in a village organisation, the objectives may be:

- Assessing the extent of rural leadership
- Identifying the socio-economic characteristics of such rural leaders.
- Evaluating the effects of such rural leadership on thegrowth of institutions involved.

Formulation of Hypothesis:

Hypothesis is a tentative solution of the problem facing the scientist. It is usually framed depending on one or more of the following sources:

- The facts established by previous investigations in the related areas
- The history of science
- The analogies
- The findings of other studies
- The body of theory

In framing the hypothesis, the researcher has to resort to null hypothesis which means possible reasoning rejection of the variables under study. A hypothesis should be empirical, conceptually clear, specific, close to things observable and related to the body of theory.

Identification of Variables:
(Dependent, Independent, and Intervening)

The most important aspect of research is to identify the variables as this would help the researcher to operationalise the methods to be used for collection of data. The independent, dependent and if possible the intervening variables should be identified. Eg. Impact of IRDP on Generation of Employment: IRDP (scheme) becomes the independent variable (or input) and the extent of employment generated, either self or wage employment becomes the dependent variable (or output). The intervening variables (process) in this could be the thinking, feeling and behaviour of the beneficiary in terms of awareness, perception, motivation and attitudes. This includes the study of the way in which the independent variable is managed, the beneficiaries' reactions and the outcome in terms of employment generation.

Definition of Concepts:

The researcher needs to know clearly the meaning and contents of every term he uses. It clarifies the issue and explains to the researchers, respondents and other readers of the research report specifically and pointedly. The term concept must be defined both the abstract term and the general meaning it is intended to convey. This may be called the 'formal definition'. The operational term by which it will be represented in a particular study should be given. This type of concept is known as 'operational definition' (Sometimes, the working definition). The operational definition helps in collection of data and in carrying out all research.

To illustrate the operational definition of the concept, take for example, size of the landholding: the land owned by a cultivator minus the land leased out plus the land leased in.

Development of Bibliography:

As soon as consultation of available source is begun, the development of bibliography should be undertaken. Each reference should appear on a separate card with last name of the author first, followed by his initials or given name and then the title of the reference (with underline), place of publication, publisher's name and the year of publication.

Nature of the Study:

The next step in formulating research design is to ascertain the nature of study, whether it is statistical study, case study, comparative study, experimental study or a combination of these or of any other type should be decided.

2.5.2 Methodology

(a) Method of Investigation:

The decision pertaining to method of investigation is to be done in this section. The researcher has to decide whether the research project is going to be empirical, in the sense of being based on primary data collected from the field, or analytical in the sense of being based on secondary data collected from published reports.

(b) Methods of Collecting Data (including tools to be used for collecting primary data and sources for secondary data):

Depending upon the nature of the respondent and the type of data to be collected, the methods of collecting data need to be decided. Also various sources and nature of documents or records should be listed out.

(c) Types of Data to be Collected (Primary and Secondary):

Primary Data: The nature of the data, and the target group or sources from which it has to be collected should be identified.

Secondary data: The sources of secondary data which is relevant for study need to be identified and collected as a supplement to primary data.

(d) Sampling Design:

The researcher has to decide the method, either census method or the sampling method to be adopted in the study with justification and the reasons for the same.

Census method involves comprehensive or complete enumeration. If a survey covers all the universe under investigation, it is called census. The study of a few representative individuals or units is called a sample study.

Complete coverage of population on a scientific enquiry is never possible or available; even when it is possible, the population will be very small. What is to be done is to take a sample from the population which is numerically adequate and culturally representative.

(e) Area to be covered:

It is essential to determine the geographical area to be covered in connection with the research study. Therefore, the physical boundaries of the area are to be specified in the research design.

2.5.3 Data Analysis, Interpretation and Reporting of Results

Data Analysis: After collecting the data, classification and analysis are to be taken up. Depending on the nature of the data and the information required to fulfill the objectives or hypotheses, the researcher should analyse the data subject to the appropriate statistical analysis besides tabulation. Tabulation of results in a meaningful way is by itself a technique and an art. The data given in the tables must be in self-explanatory form. Tabulation may be done either manually or by computer method.

Interpretation of Results: After organising the data into a meaningful form the researcher should draw inferences. Drawing inferences or interpreting the data will call for statistical inference based on usual test for significance.

Reporting the Findings: No research is an end in itself. The investigator report must what he discovered or innovated based on the data collected to fulfill the need for which the study was taken up and to ensure proper directions to other investigators in carrying out of similar researches. Report of the findings must be clear, simple and directly relating to the objectives of the study.

2.5.4 Chapterisation (Chapter Scheme)

Preparation of a chapter outline is the last step in planning the thesis/dissertation and it is an useful first step in writing a rough draft. Some of the headings may need to be changed as the investigation progresses. The final form is determined by the nature of the study itself and by the conventions. The number of chapters that the study contains and the name of each chapter must be mentioned. The basis for chapter scheme is the objectives of the research problem.

Chapter I	-	Introduction
Chapter II	-	Review of Research
Chapter III	-	The Problem, Objectives and Hypotheses
Chapter IV	-	Analysis of data and discussion of the results
Chapter V	-	Summary and Conclusions
Chapter VI	-	Implications and Suggestions for future research

2.5.5 Time and Financial Budgeting

In carrying out the research, time factor is very important, otherwise the work carried by the researcher and the findings of the study become a mere waste. Since research is meant to solve the problems of the society the researcher has to respond quickly and in time.

Financial factor is also a governing one in carrying out the research successfully.

A) **Time Plan:**

i)	Preparation of research tools/schedules	1 month
ii)	Pretesting of the tools/schedules	1 month
iii)	Finalisation of the tools	1 month
iv)	Collection of data/field work	1 month
v)	Data processing and analysis	2 months
vi)	Preparation of the report	
	(a) First draft	3 months
	(b) Final Report	3 months

B) **Personnel Requirements and Budget Estimates:**

i)	No. of Respondents	300
ii)	No. of Temporary Field Investigators	4
iii)	No. of Schedules to be filled per day by an Investigator	3
iv)	Training of Investigators	3 days
v)	Internal Travel	4 days
vi)	Field Work	26 days
vii)	Coding and Transferring of data	10 days
		43 days

C) **Financial Estimates:**

i)	Salary for temporary Field Investigators @ Rs.50/- per day per Investigator	Rs. 8,600
ii)	Travel Expenses for the temporary Field Investigators	Rs. 1,000

iii)	Stationery and Cyclostyling of schedules	Rs. 2,000
iv)	Tabulation of data (by Computer)	Rs. 2,000
v)	Contingencies	Rs. 400
	Total	Rs.14,000
vi)	Overhead @10%	Rs. 1,400
vii)	TA and DA for permanent staff	Rs. 6,600
	Grand Total	Rs.22,000

2.6 LET US SUM UP

In a developing society complex problems which need to be researched are bound to arise. But all problems are not researchable. Some problems are socially useful and need to be researched while others cannot and should not be researched because: (a) these are not socially useful (b) the labour is not worth the results which might come out of the research and (c) the parties from whom data are to be collected would not at all be willing to provide information and as such results of the research would not be dependable.

In the formulation of a research proposal there are different stages and all are inter-linked with each other. The first and most essential thing in a research proposal is to have a clear idea and definition of the subject matter to be enquired into. It is really difficult to formulate a research proposal. Some researchers believe that it is more difficult to formulate a research proposal rather than to solve it.

Some considerations which weigh in the formulation of a good research proposal are:

(a) the area of the study should be manageable

(b) before selecting the problem the researcher should study available literature having a bearing on the general area.

(c) he should see that necessary and relevant information is available and also that he knows the sources from which the data are to be collected.

Before proceeding further it is essential that the problem is properly defined and the hypotheses are clearly formulated. On the face of it, the selection of a problem may appear to be simple but one envisages the operational difficulties of putting a design into effective research. A researcher should follow some precautions in the selection of a research problem. He should select a topic in such a field in which he has familiarity and deep understanding. The research problem must be worth studying and should have social relevance.

The research problem may be selected from a theory of one's own interest, daily problems, technological changes, unexplored area, library and discussion with supervisors.

The criteria of a good research problem are: having relation between two or more variables, clarity and unambiguity, possibility of empirical testing, creativity, novelty, interesting and important.

A research design or model indicates a plan of action to be carried out in connection with a proposed research work. It is a tentative plan which undergoes many modifications as the study progresses. It presents a series of guide posts to enable the researcher to keep track of his actions and to know that he is moving in the right direction in order to achieve his goal.

A research design should be based more or less on methodology. It should be made once the topic and problem for research have been selected and formulated, and objectives have been properly framed.

A good research design should possess the following four characteristic features, viz. objectivity, reliability, validity and generality.

In a narrow sense research design refers to the procedures for the collection of data and its analysis. In its broader sense a research process involves identification and selection of a research problem, choice of theoretical framework for research problem and its relationship with previous research.

2.7 REFERENCES

BEST, John W. *Research in Education*. New Delhi: Prentice-Hall, 1977.

GOODE, W. J. and Hatt. P. K. *Methods in Social Research*. New York: Mc Graw-Hill, 1952.

KERLINGER, Fred N.*Foundations of Behavioural Research*. Delhi: Surjeet Publications, 1978.

LUNBERG, G. A. *Social Research: A Study in Methods of Gathering Data*. New York: Longmans Green, 1951.

REDDY, C. R. *Research Methodology in Social Sciences*. Delhi: Daya Publishing House, 1987.

SADHU, A. N and Amarjit Singh. *Research Methodology in Social Sciences*. Bombay: Himalaya Publishing House, 1980.

SARAVANAVEL, P. *Research Methodology*. Allahabad: Kitab Mahal, 1987.

SINGH, R. A. P. *Methods in Social Research*. Jaipur: Printwell Publishers, 1990.

VERMA, R. K. and Gopal Verma. *Methodology and Techniques of Research*. New Delhi: Anmol Publications, 1988.

YOUNG, P. V. and Schmid, C. F. *Scientific Social Surveys and Research., An Introduction to the Background, Content, Methods, Principles and Analysis of Social Studies*, 3rd ed. Bombay: Asia Publishing House, 1961.

2.8 ASSIGNMENT

1) State the components of a research design and draw a brief design for your proposed research problem.

2) Give a brief critical review of literature in the field of your research work.

2.9 RECOMMENDED BOOKS

GOPAL, M. H. *An Introduction to Research Procedure in Social Sciences*. Bombay: Aria Publishing House, 1964.

HAKIN, Catharine. *Research Design*. London: Allen & Unwin, 1987.
HANS Raj. *Theory and Practice in Social Research*. Delhi: Surjeet Publications, 1979.

JASPAL Singh. *Methods of Social Research*. New Delhi: Sterling Publishers Private Limited, 1991.

NACHMIAS, David and Chava Nachmies. *Research Methods in the Social Sciences*. London: Edward Arnold Ltd., 1996.

PHILLIPS, Bernard S. *Social Research Strategy and Tactics*. New York: The Macmillan Company, 1966.

SELLTIZ, C., Jahoda, M., Deutsch, M., Cook Stuart. *Research Methods in Social Relations*. New York: Holt, Rinehart and Winston, 1964.

SJOBERG, Gideon and Roger Nett. *A Methodology for Social Research*. New York: Harper & Row, 1968.

SANDERS, William B. and Thomas K. Pinhey. *The Conduct of Social Research*. New York: Holt, Rinehart & Winston, 1983.

WILKINSON and Bhandarkar. *Methodology and Techniques of Social Research*. Bombay: Himalaya Publishing House, 1988.

2.10 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) Define a research problem. What criteria would you advocate in selecting a problem for research ?

2) A fresh Post-graduate student is problem-blind. He has eyes but he does not see. He has ears but he does not hear. Comment.

3) What precautions should be followed while selecting a research problem ?

4) What are the steps involved in evaluating the problem ?

5) What is meant by Research Design ?

6) Discuss the role of research design in social research ?

7) In devising a research strategy, what things would you have to take into account ?

8) Explain the relation between research problem formulation and research design ?

II. SHORT NOTES

a) Topic suggested by a guide

b) Sources of research problem

c) The technique of skimming

d) Delimitation of the problem

e) Role of Hypotheses in making a research design

UNIT - 3: SURVEY AND REVIEW OF LITERATURE

Structure

3.0 AIMS AND OBJECTIVES

Survey and review of literature is an important step in any research. The present unit focuses on how a learner should proceed on this aspect before he or she undertakes a research project.

After studying the unit, you should be able to

- explain the need for literature survey
- describe how to survey various documentary sources for information
- discuss the importance of reference books and archival data
- demonstrate how to gather information through database searching and
- explain how to do a literature review.

3.1 INTRODUCTION

The very first thing in a research project is to construct a theoretical problem. After stating the problem we have to review the literature on similar phenomena as our study. Usually, the researcher reviews the literature before the research has begun so as to get an idea of where to proceed and to see if there are any pitfalls to watch out for. However, researchers will often go back over the literature after the research is completed to see if they can account for some unexpected findings. Whatever may the case, it is important to describe other research in the area and to tell where our research fits. Not only does this heighten the relevance of our own work in a scholastic community, it may also show how our research blazes a new path that no one else has thought of.

As a practical question, students wonder exactly, how much literature review should be put in a paper. To some extent this will depend on the topic and the way it is being examined. Some areas have an immense amount of research while others have hardly any. It would take far too long to review every book and article on social class and belief systems, and so the literature review would narrow its focus to those items that are most similar to the research being presented. At the same time, it is possible to summarize several studies that found the same thing or the one that used the same theoretical and methodological approach. What is important in reviewing the literature is to provide an overall picture of what has been done and how our own work adds to knowledge in that area.

3.2 LITERATURE SURVEY

There are several reasons why an important step in dissertation writing is to have literature survey on the concerned subject. You might have done some preliminary survey while finalising your topic. The same will have to be extended to wider areas. Presumably, your topic has not been completely investigated by others in the past, otherwise you would not have settled on it. There is still enough scope for research on the topic with a view to reinterpreting the known facts, or discovery of new facts, or application of the theories valid elsewhere, to your specific topic.

3.2.1 Need for Literature Survey

No conscious investigation can begin without taking into account the historical background of the problems. It is only common sense to find out where others have floundered so that you may be cautious. You need not embark on your quest in the manner of a fledgling. The failures of the attempts by the investigators in the past need not deter you from trying again in a more organised manner, provided you can be certain enough about the causes of the failures so as to overcome the same. There is an additional glory in succeeding where others have failed, and you might make a special point to identify the failures of the investigation in the past, while surveying the previous literature. That will prompt you to know the methodologies adopted in the past to solve those specific problems. Mere knowledge of failures will be of little practical use.

The second reason for a full literature survey is to know the outcome of those investigations in areas where similar concepts and methodologies had been used successfully. That is no guarantee of success in your research but it should encourage you to try your own approach. This is particularly helpful in hypothesizing. Both the positive and the negative aspects of the fate of related research projects will provide you latitude to think and imagine.

Lastly, an extensive or even exhaustive process of such a survey should offer you vital links with the various trends and phases in the researches in your area of specialisation, familiarising yourself with the characteristic precepts, concepts and interpretations, with the special terminology, with the rationale for undertaking your proposed investigation. Each research has a formal and systematic structure which generates appropriate norms for evaluating its outcome. Only through a patient survey of the previous related researches will you be able to formulate a satisfactory structure for your project.

3.2.2 Locating the Sources of Information

Locating the sources of information becomes your next logical step in a survey of the previous research. If your work would be wholly derivative, like an assignment, it can be completed by reading in the libraries and analysing the data collected by you in a presentable form. But most of the topics of the dissertations and theses demand investigations beyond finding out facts and opinions of others in the documents available to you in libraries.

It will be essential to know exactly the reasons for your use of libraries. Libraries are the principal sources of information which give you an indirect, second-hand knowledge of the reality. It is a systematic storehouse of data of encyclopaedic nature. It contains a record of the results of human thinking. The written word represents a very effective medium of communication provided you are able to understand the meaning.

Basically the sources of information are of two kinds. The primary sources are the publications, like those of books, journals, reports, based on first-hand field work, explorations, experiments or direct observations which can provide original data to you. If their accounts are found authentic, you might rely on them for your investigation. On the other hand, the secondary material consists of those books, articles in journals and encyclopeadias, dissertations, theses which already exploited the direct first-hand data. They give you pre-digested, processed second-hand data. Both these types of data have their uses but should not be confused with one another.

Facts and opinions based on direct observations are more valuable for your investigation. The speculative opinions, interpretations by others might also help you in generating your ideas, provided you subject them to a critical review. The first-hand data will contribute to the inductive reasoning. The second-hand data, if demonstrated as reliable, will support your deductions. Your own investigation will need a judicious combination of both inductive and deductive inferences from the two kinds of data. In addition, you will use your own personal observations, fieldwork, experiments for further inductive reasoning. In your **library research** you will collect both types of data and keep them recorded separately.

3.3 DOCUMENTARY SOURCES OF DATA

In anthropological studies of primitive societies the field worker in the past often had no written sources of information which he could call upon at all. The primitive societies were preliterate; written information did not exist. While this obviously placed great limitations on the evidence available to the social anthropologist it also spared him the problems of trying to find out what help he could get from documentary sources. But most sociologists, and now many social anthropologists also, are dealing with complex literate societies in which the accumulation of documents of one sort or another has been going on for centuries. For the sociologist to derive the maximum amount of benefit from published data, he needs to bring some order into his understanding which calls for some classification of documentary sources. It is customary in research to distinguish between the sources of the documents by classifying them as primary and secondary.

3.3.1 Primary and Secondary Sources

Primary sources provide data gathered at first hand. They are original sets of data produced by the people who collected them. They are contrasted with secondary sources, which are the data obtained second-hand; that is, sets of data not collected at first-hand but called from other people's original data. The distinction between the two types is generally not too difficult to make if one knows enough about the sources of the data.

3.3.2 Official Records

These should, on the face of things, be the most reliable sources, so long as we can trust the writer. The verbatim Parliamentary record of what is said in the Rajya Sabha and Lok Sabha is possibly the most trusted document one could find. Any disagreements between speakers in the two houses and the official record are carefully looked into and such is the integrity of the record that any attempt to alter it, except for errors, would undoubtedly result in a national scandal. One would also tend to accept verbatim reports of judicial proceedings, evidence given before special commissions and special committees of enquiry, so long as one can genuinely believe that the shorthand writers were quite impartial, skilled in their work and that all they took down has been printed without editing.

3.3.3 Newspapers

Newspaper reports, where a reporter was present at the scene, might be thought to be valuable, but unfortunately it has been shown only too often how little reliance can be placed on them. It is important to recognise under what pressures newspaper correspondents work. Many of them do not have shorthand and use their own personal notation system. Their reporting of speeches is likely to be inaccurate in many ways, sometimes embarrassingly so. Also a reporter can only extract a tiny part from any total occurrence, and in many cases he extracts what will be eye-catching and provocative.

Newspapers love a good disaster, and a wedding where the bridegroom's trousers fell down at the marriage would stand a far better chance of being reported on than one that went without a hitch. It should also be remembered that newspapers work very fast-the very latest news is always the most desirable, as is instanced by the occasional **scoop** of the event which has not yet even happened. Reports are cut by sub-editors to fit into available space and tuned to the particular **angle** of the newspaper's political policy. In all, it is little short of a miracle that the subsequent report bears any relation to the actual occurrence at all - and sometimes it does not.

3.3.4 Official Statistics

In this country we tend to trust our official governmental statistics and then to distrust the use made of them by politicians. The census of population, the monthly and the annual statistical reports of the various ministries and other national bodies and local authorities produce a great deal of extremely valuable data for the sociologists to work on. We will not go into details of the dangers of using statistics but merely point to some precautions which must still be taken even when using statistics.

The research worker must be absolutely sure he knows what the statistics are about. In criminal statistics, for example, the definition of a **crime** is operationally made in a number of ways. Obviously **crimes known to the police** are greater than '**people prosecuted**' which in turn are greater than **persons found guilty**. To use criminal statistics wisely necessitates some prior work in getting to know the difference between indictable and non-indictable offenses and also some legal history.

The use of index numbers and special ratios should always be undertaken with care since they may contain hidden dangers if one does not know how they are compiled and, very often, how they are **corrected** for a variety of reasons.

Not all **official** statistics, of course, are governmental or even quasi-governmental. We would probably accept the newspaper circulation figures put forward by Bureau of Economic and Statistics and we would be happy to quote the figures of National Sample Survey (NSS). But how the first is calculated is probably unknown to most social scientists, and although the problems encountered in gathering the latter are described in detail at the beginning of each publication, probably few researchers bother to read through the technical details.

Educational statistics are also prone to mislead. If one uses statistics of children in over-crowded classes it is absolutely fundamental to bear in mind that to be over-crowded a primary school class must exceed forty pupils, yet a secondary school class need only top the thirty mark. This is a charming example of how to use different bases for calculations so as to divert attention from the needs of the youngest pupils.

3.3.5 Diaries, Memoirs and Autobiographies

The divisions between diaries, memoirs and autobiographies are not easy to make since one form of record can shade into the other or even contain parts of each other. For our purposes, the divisions could be that a diary is written at the time of the events, memoirs are a writer's recollections of a particular period (perhaps aided by diaries) in which he was not necessarily the central character, while an autobiography is an attempt to give a systematic and chronological record of the author's life with himself at the centre of the story.

With politicians and military leaders the keeping of a diary is not only the gathering of data for future book but can also be a form of self-defence preparation for any postmortems which might be held in the future. In a sense it is a personal record of **the minutes** of what occurred at conferences or in battles which will remind the writer of what happened at a precise time in the course of the events.

3.3.6 Biographies

Autobiographies and biographies are avidly read and often appear high in the lists of best sellers of the day. The biographer is likely to work on people of some fame, whatever may be their spheres of activity, and so the biography is more likely to be about the famous (or notorious) person, the outstanding success or the eye-catching unusual personality. Controversial figures in public life will probably attract biographers for years to come because it is felt that there are still new things to be said about them, new perspectives in which their work can be seen, and new interpretations of politics or the arts in which their contributions should be judged. But even if we tend to look for the scholarly, non-involved biography of a great man/woman we cannot be sure that the work will be beyond criticism.

While the academic biographer may wish to prove a hypothesis, the professional biographer may present even greater problems. It is sad but undoubtedly true that sensationalism can be used to promote sales in this present day. In other cases of course, the biographer may be concerned with the defence of a dead person, and this is particularly true of biographies written by loving sons or daughters. In yet another instance the biographer may have been commissioned by the family of a person who has died to write a biography with their authorisation and cooperation. The latter types of biography must be particularly carefully scrutinised since it is hardly likely to contain much criticism of its subject.

Another interesting point in biographical-historical work is the use made by later writers of their subjects' own diaries and records. In an interesting review of two books on the 1914-18 war, the historian and critic A.J.P.Taylor wrote, 'There is one piece of advice which can be given to all generals, successful and otherwise: "Do not keep a diary". Statesmen usually follow this advice; so do most ordinary mortals, unless they are intent to provide a figure of fun for posterity. Generals ignore it, to their subsequent discomfort.

One last example may be given here to demonstrate the dangers of biographers copying each other's errors. Too many biographers copy one another. When Peter Mann was writing an essay on Bernard Shaw, the author read in one biography that he had attended a Wesleyan school and had always been at the bottom of his class. Before printing the essay the author thought that it is safer to send it to Shaw. He returned it with this comment: "I have never been to a Wesleyan school nor been at the bottom of my form".

3.3.7 Correspondence - Private and Public

Although these types of data are always mentioned as source for sociological enquiry, their use for research is limited. For specifically sociological purposes, however, there are likely to be few types of correspondence which come easily to hand, and even if they did, of what value could they be ? It is obvious today that **collections** of correspondence will be of historical value; today we do not find the same sort of letter writing as in the past when people had no telephones and were not so mobile. It is probably only young lovers today who are separated for some reason or other who keep their sweet hearts' letters, and publications of these would be extremely embarrassing to everyone. Apart from this uncertain source it is hard to think of letters which could be obtained after the correspondence had taken place. The use of correspondence would, therefore, seem to be virtually nil unless the letters were historical.

But even if letters are available for research, they still have their limitations. Skill in letter writing by no means is evenly distributed among the population and letter writers cannot be said to be representative of a general population. Also in writing of events the writer is certain to have to abbreviate any descriptions very sharply. So letters condense events enormously, and they are also probably written from a particular angle with the recipient in view. One can imagine a university student writing two letters, one to his parents, and the other one to his best friend, describing what the rag dance last Saturday had been like. It is very unlikely that the details would be the same for the two recipients.

3.3.8 Historical Documents

In a sense it is almost impossible to separate out a section on documents which can be called **historical** since all documents are historical. But conventionally historical documents refer particularly to events of the past about which the main source of information is documentary, the participants now being dead. Even this definition is by no means satisfactory, since there is plenty of history of the twentieth century, and plenty of people are alive today who have memories running from its beginning.

The sociologist is by no means universally beloved of historians. The prejudice against sociology often derives from the historian's belief that sociologists are just bad modern historians.

On the other hand sociologists frequently find that historians fail to satisfy because of their lack of general theoretical framework within which their histories are placed; in other words they see the historians as chronologers without frames of reference.

Sociologists are rarely fully trained historians, even though their education may contain a certain amount of economic, social and political history. Of course, sociologists do not usually want to get too historically involved, since the depth of study in any particular period of time in which they can so easily become involved can lead them away from their own original sociological purposes. But in a way this is all the more reason for the sociologist to know something about sources of historical documents and the care that must be taken in their use.

Much sociological investigation begins with a historical background. The community study begins with the historical development of the settlement, from village to town; the study of the organisational structure of the trade union can only gain a good perspective by tracing the development through from the inauguration of the union - or probably before then. In all work such as this, the sociologist, just like the historian, must turn to documents to establish the sequence of events and try to understand the processes and interactions that took place in time now past. Perhaps the work centres on detailed analysis of the books of the census enumerators as they become available a hundred years after the census takes place; perhaps the researcher becomes immersed in tracing records of landholdings or he may, yet again, work on the minute books of committees of long past. In all probability he will use a combination of sources, some local, some national, some public, some private, some primary and some secondary.

The sociologist, like the historian is trying to re-assemble events in a way that the jigsaw puzzle player works.

One prime warning for the sociologist is that his historical work should be based on sound sources, and perhaps the best thing is to advise him to consult the historian before he begins. The sociologist who dives into historical research without seeking advice from the historian will only waste time and effort in seeking out the correct sources and is likely to use the sources badly if he knows little or nothing about the way in which the documents were compiled.

3.4 REFERENCE SOURCES

Any written or printed composition of some length is called a book. The word 'book' has a very wide connotation. It covers all kinds of reading materials in any form which serves the purpose of the readers whether for general or recreational reading, or for study and research, or for mere consultation as a source of information. There are various types of books which may be called as 'text-books', 'treatises', 'monographs', 'classics', 'reference books', etc. The librarian must have a fair idea of the nature of such books and their use.

A 'text book' is meant for providing basic knowledge of a particular subject to the student in an easy language. It is a presentation of the principles of a subject in a systematic way and is divided into various chapters giving paragraphs to give a continuous exposition of the subject. For example, Esdaile's *Students' Manual of Bibliography* is one of the text-books for the students of library science.

A 'treatise' is a comprehensive, authoritative and systematic compilation of known information on a subject with the detailed coverage of the subject. For example, Laski's *Grammar of Politics*.

A 'monograph' is a scholarly treatise on a specific area of knowledge and is studied as a part of a wider subject. For example, Indu Bhushan Bannerji's *Evolution of the Khalsa*.

A 'classic' is an outstanding work of scholarship and high literacy merit known to the world for all times. Homer's great epic poems '*Iliad*' and '*Odessey*' are among the world's classics.

Valmiki's *Ramayana* and Vyasa's *Mahabharata* represent the great epics from the Orient. With the development of reference service, the necessity arose to have a collection of such books as could be consulted as sources of information. These are called "reference books".

A reference book differs from an ordinary book in many respects. An ordinary book, that is, a book other than a reference book, has the following characteristics; It is made of continuous exposition. Sentences mount into paragraphs. Paragraphs mount into a chapter. Chapters get woven into a single swelling exposition, in the continuous pursuit of single idea simple or complex. On the contrary, a reference book may have disjointed entries, the sequence of which is determined by the scheme of arrangement chosen for the purpose. It might be alphabetical, classified, or some other type of arrangement; or, a reference book may consist mostly of formulae, statistics, diagrams, tables, maps, charts, or lists of documents with or without abstract or annotations.

An ordinary book is written primarily to be read consecutively for inspiration, enjoyment or information. But the reference book is designed to be consulted or referred to from time to time for a specific piece of information. Ordinarily they are not meant for continuous reading.

According to *ALA Glossary*, a reference book has been defined "as a book designed by its arrangement and treatment to be consulted for definite items of information rather than to be read consecutively". It may be stated that there is no strict line of demarcation between reference book and an ordinary book. Often, the ordinary books may have to be looked for specific information.

Ordinarily, the reference books are not allowed to be issued for outside use as these are required from time to time for consultation. Some one may require it at any time. Therefore, the use of reference book is restricted to library building. Due to the nature of their use, these are usually kept in a separate sequence in the library. In such books, facts and information are usually brought together from a vast number and a variety of sources, and are arranged conveniently for expeditious use. The main purpose is to supply information.

3.4.1 Characteristics of Reference Books

Reference books have some distinctive characteristics, which are as follows:

(1) They are intended primarily for occasional consultation

(2) They are not meant for comprehensive reading

(3) They are consulted for definite items of information

(4) The information included in them is collected from a vast number of sources.

(5) It is a miscellany of information or facts.

(6) The arrangement of information is such that it can be conveniently and quickly recalled.

(7) They usually include only the bird's eye-view of the topics and rarely deal with them in depth.

(8) They usually concentrate on facts.

(9) They are normally not issued on loan and are kept for use in the library only.

3.4.2 Kinds of Reference Books

There are different kinds of reference books, which are useful for gathering information for research. They are

1) Language Dictionaries
2) Encyclopaedias
3) Biographical sources
4) Yearbooks
5) Geographical sources: Gazetteers, guidebooks, maps, atlases.
6) Current Sources: Handbooks, manuals and sources of statistics
8) Bibliographies
8) Audio-Visual Sources

According to William Katz, the reference sources may be divided into two large categories; the control access-directional type and the source type. The control access-directional type sources include bibliographies, catalogues - indexes, abstracts, etc. They serve as control device - a kind of check-list. They themselves do not give definite answers, but serve to direct the users to the sources of answers.

Works of source type are usually self-contained. Unlike the access type of reference work, they are synoptic. These include encyclopaedias, year-books, almanacs, handbooks, manuals, directories, dictionaries, biographical sources and geographical sources. Apart from this, there are government documents which are official publications ordered and normally published by central, state and local governments.

3.4.3 Evaluation of Reference Books

In order to achieve a thorough understanding of reference books, it is essential to know how to evaluate a reference book. A reference librarian should know the criteria of evaluation of a reference book. A reference book must be able to provide specific and accurate answers to queries without much difficulty. The librarian must ask at least four basic questions about a reference work: What is its purpose? Its authority ? Its scope? Its proposed audience? Finally, the format of the book must be considered.

According to Louis Shores, the important points to serve as check-list for evaluation of reference books should consist of (i) Authority, (2) Scope, (3) Treatment, (4) Arrangement, (5) Format, and (6) Special Features.

The purpose of a reference work should be kept in mind. It should be evident from title, table of contents, preface or introduction, and the index. The other clues to purpose can be found from publishers catalogue, book reviews, jacket of the book, etc.

i) Authority: The authoritativeness of a reference work can usually be judged from the qualifications and experience of the authors, contributors, editors, etc. and the reputation of the sponsoring body, if any, and the publishers. The librarian has to rely on the authors' known scholarship and experience. Some publishers have excellent reputation for issuing reference works.

ii) Scope: The scope of reference book should be understood properly. To what extent the purpose of the book has been fulfilled? What is the coverage of the book? What is the range of the subject matter, and what are its limitations? Does the work provide recent information on the subject? To what extent do the bibliographies indicate scholarship and additional information?

iii) Treatment: A critical appraisal of the treatment of the matter recorded, involving accuracy, objectivity and style should also be made. How thorough, reliable and complete are the facts? Does it show any biased or controversial topics? Is the level of writing for layman or scholar, adult or child?

iv) Arrangement: It is essential that one should be able to locate information easily and quickly. The arrangement can be alphabetical, classified or chronological or some other. The classified or chronological arrangement must be supported by alphabetical index. The index should provide various approaches and cross references.

v) Format: It refers to the physical get up of the book. The quality of binding, the quality of paper, the type face used for printing, the illustrations, the layout of the book are also to be judged to see whether satisfy the minimum specifications. Is the physical get up attractive?

vi) Special features: What are the distinctive features which distinguish this reference book from other reference books of similar type?

3.5 ARCHIVAL DATA

The ancient archives contained the best of societies' accumulated knowledge in the form of clay tablets, palm leaves, copper plates etc on which royal orders, official letters, etc. were inscribed. These were preserved and stored at a place which was known as archives or library. Earlier civilisations made no distinction between the archives and the library. The distinction between library and archives material dates back to the 15th Century, when the invention of printing technology brought a tremendous change. It is now possible to produce multiple copies of the books previously available in manuscript form only. With this the distinction between archives and library became apparent.

3.5.1 Library and Archives - Today

A library is a collection of written, printed or other graphic or visual material (including films, photographs, tapes, discs, microforms and computer programmes) organised and maintained for reading, study and consultation, whereas Archives is a repository for an organised body of records, produced or received by a public, institutional or business entity in the transaction of its affairs and preserved by its successors.

Archives are public records and documents. According to Schellenberg "archives may be defined as those records of any public or private institution which are judged worthy of permanent preservation for reference and research purposes and which have been deposited or have been selected for deposit in an archival institution". The term record means all books, papers, maps, graphs or other documentary materials regardless of physical form or characteristics, made or received by any public or private institution in pursuance of its legal obligations, or in connection with the transaction of its proper business and preserved or appropriated for preservation by that institution or its legitimate successor as evidence of its functions, policies, decisions, procedures, operations or other activities or because of the informational value of the data contained there in.

3.5.2 Library and Archives - Analysis of Activities

The above conceptualisations (library and Archives) point out some of the similarities and differences of both the 'institutions. These also indicate some common issues shared by both the institutions. Now-a-days the archives and libraries mainly carry out two distinct functions.

1) They deal in information and specifically writings as a communication medium; and

2) Control information by selecting/collecting from a mass of data, portions which may be useful to the society.

Despite their common interests, both are different too. Each deals with different formats of information packages, controls information by different means. On the contrary, both have some common activities like collection, building, public relation work, techniques used, services offered, bibliographical control of information, etc. An analytical look into the activities and services of both will help to reveal their relations.

i) Collection

It is easy to state, that the difference between the material housed by the two is visible. A cursory look leads one to say, that archives contain records and libraries contain books. This statement, however, would be misleading and inaccurate. It is because, now in addition to records, archives also procure reference sources, maps, gazetteers, reports and books etc, needed by research scholars and creating agencies. Libraries, on the other hand house documents in all physical forms. Thus, even the distinction on the basis of printed format cannot be made, because archives also deal with printed matter like reports, biographies, printed books, etc.

It may be deduced that, except for the relative quantity of books or records, there is little difference in the collection of large libraries and large archives. Only one major difference revealed is, that libraries usually do not have records.

The ultimate aim of both the institutions is to have a reasonably comprehensive collection, to organise it in an orderly manner, in order to achieve an efficient retrieval of documents when called for.

ii) Acquisition Policy

Libraries acquire documents mainly through purchase. In addition some documents are acquired through exchange arrangements and donations.

Archives on the other hand, are rightful inheritors of all records created by corporate bodies. Therefore, archives get the public records automatically transferred to it, when these are no longer of current administrative value to the creating agencies. In addition to records transferred free of cost to it, the archives acquire private records through purchase.

Archives like libraries, though at a much smaller scale, acquire reading material (books etc.) through purchase and donation. Exchange of copies (duplicates or reprographs) of records, though possible, is not often resorted to by the archives.

iii) Selection of Documents

Libraries use bibliographical aids. To this end, often the documents themselves are physically available before these are selected for purchase.

On the contrary, in archives a process similar to document selection in libraries is termed as "records-appraisal". The appraisal of records is conducted at the record centres of the creating agencies i.e. before the transfer of records to the archives. Archivist mainly judges the historical or research value of such of the records as are not to be retained permanently according to retention schedules because these are not required by the creating agencies for future administrative reference. Except these, all other records are to be kept permanently for future reference and thus transferred to the archives.

In addition, archives have a library unit. For this unit, invariably the director of archives selects different documents which includes reference sources, maps, books, etc. on history and allied subjects. There is no regular budget fixed for this purpose. A large number of documents, published by the Government departments and other corporate bodies, are received free. The purchase committee constituted in the archives mainly selects and purchases a number of private records.

iv) Processing of Documents

In both the institutions, documents are accessioned after receipt. The columns of the register used in archives and library are different. This is because records lack physical description like author, title, imprint, etc.

The librarian classifies the documents according to a predetermined accepted classification scheme. He analyses the subject matter of the document and assigns the class number. The classification helps to arrange the documents on the shelves. An archivist on the other hand, has no such scheme for classifying the records. Each creating agency, has its own functional arrangement given to records.

Archivist only assigns location number in the form of stack number, etc. to locate the records when required. Other documents are classified with the help of a classification scheme, like in libraries.

After this step, documents in the library are catalogued, but in the archives the subject content of the document is described to create various retrieval tools or 'finding aids'. The library catalogue offers various approaches to readers, like approach by author, title series, collaborator, subject, publisher, year of publication etc. But as records lack physical description, archivist prepares indexes, guides to the holding, etc. to provide different approaches to identify records.

3.5.4 Services to Users

The essential difference between libraries and archives is the access to the stack rooms. Invariably we have open access in libraries and closed access in archives. Open library stacks afford browsing opportunity to users, which is denied in the archives. The uniqueness of the records, the fact that papers are loose in the boxes, and the impossibility of arranging shelved items in any observably logical order, all vouch against the researcher's free access to the stacks.

a) **Reference Service:**

Mainly two types of activities are performed by the reference staff i.e. they give personal assistance to users in locating or selecting the material; and provide answers to their inquiries.

In archives, due to closed access, users are always familiarised with different finding aids viz. indexes, summary inventories, guides, descriptive lists, etc. with the help of which users can select the records. Archives with the help of handbook/guide book, etc., acquaint the readers about their holdings. Archives also answer user's enquiries received by post, telephone or in person regarding information contained in the archives.

b) **Loan Service:**

Users in archives can only use the material in the search room (reading room). However, records are loaned out to the creating agencies, courts, other government departments (with the permission of the creating agencies) for a specific period. Even documents other than records are not usually loaned out for house reading.

3.5.5 Care and Preservation of Documents

In libraries, the predominant format in the collection is usually the bound volume whereas in archives this place is held by the single sheet format. Thus, preservation problems and techniques used for care and restoration are basically the same, with varying emphasis on different techniques.

Both use techniques like binding, laminating chiffon repair, mounting etc. for restoration of documents. Fumigants, naphthalene balls, etc. are also used for preservation of documents. In archives, as records are not weeded out, the dependence on these techniques is much higher, as compared to the libraries. Also, as quite a large number of records are handwritten, archives have to take special care of these, so that readability may be maintained.

3.6 DATA BASE SEARCHING

A database is an organised collection of information. There are a number of databases which include telephone directory, dictionaries, encyclopedias and card catalogues in the library. The term **database** as used here means a computer programme that is designed to store and retrieve information of all kinds. Data can be entered in any order and then sorted out using any factor stored. For instance, if the information in the telephone directory for the city of New Delhi is fed into a computer, an answer to the question - "How many people living in Karol Bagh have a surname ending with **ta**?" - could be obtained in minutes, while a manual search for the same query could take months.

The Anthropological Survey of India (**ASI**) has compiled a database entitled 'People of India'. This is an ethnographic profile of the various communities in India. Besides bio-anthropological data, it includes information on the language spoken, food habits, social divisions, marriage customs, family structure, status of women, life cycle rituals, economic activities, mechanisms of social control and religious attitudes. This database is of immense value to any one preparing to do research work within the community.

With the rapid growth of knowledge, it has become very difficult to keep abreast of the latest developments in a field without spending hours in the library. This problem has partially been overcome by the abstracting/indexing services. These services compile abstracts/indexes of the published material in a given discipline. **The problem, however, is that from the printed version of the abstract/index it is often very difficult to retrieve all the references pertaining to a topic, especially if it is interdisciplinary in nature.**

There are a number of databases relevant to the social and behavioural sciences. A good way to discover which data base is likely to yield the most fruitful result is to consult the abstract/index it is based on. For example, the American Psychological Association's **Psyc INFO** is a data base built along the lines of the printed version of the Psychological Abstracts. It covers 1300 journals from approximately 50 countries published in over 25 languages, and contains over 6,50,000 references (most with abstracts) dating back to 1967. It is updated monthly and provides citations and abstracts for the international literature in psychology and related disciplines.

The databases mentioned in **Table-1** are accessible to researchers residing anywhere in India.

The National Social Science Documentation Centre (NASSDOC) at the ICSSR, New Delhi, conducts manual bibliographic searches for a nominal fee. Ordinarily searches can take over a month. A database on CD ROM, POPLINE (Population Information On Line), is also available. The database gives information on sociological and biomedical aspects of the population.

Table-1: *Some Databases Relevant to the Social and Behavioural Science*

Data Base	Printed Version	Subject Area	Also Includes:
BIOSIS	Bilogocal Abstract	Bio-Science	Behavioural Science, Nutrition, Public Health
Dissertation Abstract	Dissertation Abstracts International American Doctoral Disertation Comprehensive Disser-tation Index	All	Science & Social Science
ERIC	Current Index to Journals in Education, Resources in Education	Education & Education Material	Counselling Early Childhood Exceptional Child Rural Educa-tion Tests, Measurements & Evaluation
MEDLINE	Index Medicus Inter-national Nursing Index	Medicine & Related Areas	Anthropology, Education,Health Care, Psychiatry, Psychology, Sociology
Psyc INFO	Psychological Abstract	Psy and Related Areas	Anthropology, Behavioural Sci-ences, Education, Health, Neu-rology, Psychiatry,Sociology
SCISEARCH	Science Citation Index	Science & Technology	Anthropology, Behavioural Sci-ence, Psychiatry, Psychology
Social SCISEARCH	Social Science Citation Index	Social & Behavioural Science	Anthropology, Criminology, Sociology
SOCIOFILE	Sociological Abstracts	Sociology	Anthropology, Education

The Indian National Scientific Documentation Centre (INSDOC), New Delhi, has on-line access to over 300 international databases which cover all the Sciences including the social and behavioural sciences. The charge for an on-line search in 1992 was Rs. 100 for every 10 references or part without an abstract and Rs. 150 with an abstract. If the information sought is available on a CD-ROM than the charge is Rs. 50 for every 10 references.

The National Informatics Centre (NIC), New Delhi, provides on line access to MEDLINE using MEDLARS (Medical Literature Analysis Retrieval System). This provides access to one of the world's largest medical databases compiled by the National Library of Medicine at Bethesda, Maryland, USA. It covers approximately 3000 medical journals in 69 languages from all over the world. In addition, NIC also has a number of databases on CD-ROM (see Table - 2).

The University Grants Commission (UGC) has recently sponsored a number of information centers throughout the country designed to help researchers in universities and colleges with bibliographic searches. These include the following.

Table - 2: Some Centres in the Country where Computer Aided Searches are carried out

Name of Centres	Data Base		Facility open to	Charge
	On-line	CD-ROM		
NASSDOC	POPLINE		Anyone	No/Nominal
INSDOC	All major databases	SCISEARCH	Anyone	Yes
NIC	MEDLINE	ERIC Dissertation Abstracts Psyc LIT SOCIOFILE Social SCISEARCH	Anyone	No/Nominal
NCSI		BIOSIS	Researcher in Indian University	No/Nominal
INFORMATICS	All major databases	ERIC Psyc LIT	Anyone	Yes
British Libraries	ERIC		Anyone	No/Nominal

If your area of interest lies within the domains of the life sciences than the National Centre for Science Information (NCSI), Bangolore, will conduct a computer search for you from the database BIOSIS (Bio Science Information Service, USA).

The SNDT Women's University Library Information Centre provides bibliographics support in the fields of sociology, womens studies, home science and special education. A locally developed database 'SUCHAK' built in scanning approximately 400 Indian and foreign journels in English, Hindi, Marathi, and Gujarathi is used to conduct the search.

Another centre for conducting computerised literature search using a locally developed database for education and psychology is situated at the University of Boroda. This facility has become operational in 1992.

Write directly to the centre that has access to a database of interest requesting an application form. The centre will issue certain guidelines for filling the form. Fill the form carefully. Any misguiding words could result in a number of useless references. After the receipt of the completed application, delivering of bibliographies could take from a week (NIC) to a month (NASSDOC).

3.7 LITERATURE REVIEW

Science is an accumulative enterprise; each researcher builds on what has been done before. Finding out about previous work is called **reviewing the literature.** There is an excitement to following the thread of scientific inquiry - like reading a detective story in which the reader is asked, on the last page of the book, to join in the search for the critical clue that will solve the mystery.

3.7.1 Networking

There is not one source of information on a topic, but many. The two places to look first are the table of contents in the front and the index in the back. Check any relevant sections and note down the names of pertinent references. If this is a one time project, you can write these titles in a notebook. The note book is better than cards which are likely to be lost. If you intend to pursue the topic further, you can list relevant references on 3 x 5 cards and file them according to the author's last name.

A basic strategy of information gathering is called **networking**. A single source can supply certain bits of information and point lead you forward to other relevant sources. Each source becomes a springboard to other sources. When all the names and titles that you encounter begin to look familiar, then you have come close to a good overview of the area. Your information on a topic will never be complete. Although it is desirable to learn as much as you can before the study, there will always be new bits of information coming in while the project is underway.

3.7.2 Using the Library

Set aside a specific time for visiting the Library. Here are some of the steps in a beginning literature review.

1) Check the card catalogue under subject headings. Most of the titles will not be pertinent to your interest, so skip over these quickly to find the few that are. Copy down the call numbers, authors and titles.

2) Check the library shelves around the call numbers that were mentioned most often in the card catalogue. Libraries tend to place books on a topic close together. You are likely to find titles that were missed in the catalogue search.

3) Visit the Reference Room to find out what is available. There may be specific reference on your topic. Look at copies of **Library and Information Science Abstracts, Psychological Abstracts** and **Sociological Abstracts** on your topic.

If you have identified one or more key individuals who have done work on the topic of your interest, look under their names in one of the relevant Abstract volumes and in the Social Science Citation Index. The latter index covers most social science journals and is available in bound volumes. The reference librarian is a trained professional whose job is to help people gain access to information. Do not just ask and run; explain your needs in detail.

Source books available are often listed as 'Guides to Sources of Information' in education, psychology, sociology, political science etc. These source books are probably indexed in the card catalogue under the subject matter heading or may be found in the reference area of the library.

4) Next step is the Periodical Room. It houses current issues of journals, magazines and newspapers. There are scientific journals that cover virtually every field of knowledge imaginable. You will have noticed some of them cited in the reference books you examined. Current issues are likely to be found on open shelves in the Periodical Room. Earlier (non-current) issues may be housed separately by subject area in another part of the library. Start looking through the tables of contents of the journals that were mentioned most often. Remember that you are not looking for dozens of relevant articles but a few.

5) Government documents are probably kept in a special document section. Documents can also be borrowed from the other libraries or directly from a government agency.

3.7.3 Direct Consultation

Find out the names of local individuals and agencies knowledgeable about the topic. They can supply information about your topic.

It is gratifying to find how helpful outside people can be if they are approached in the right way. A further advantage of local consultation is the possibility of making valuable contacts. During conversations with local officials, students may learn about new programmes and job opportunities.

District officials and authorities are more difficult to use as information sources. If you write to the author of a book on a topic, make your requests specific. Some famous persons are extraordinarily helpful with student inquiries. If you have a specific question that you think some distant authority can answer, don't hesitate to write a personal letter. Enclosing a stamped self-addressed envelope will increase the likelihood of a reply.

3.8 RECORDING OF SURVEYED INFORMATION

The degree to which notes reflect exactly what was said by the respondents depend on how the interview will be used. When the information is the final source of data, a precise transcript is essential. The most common means of recording is to write down the respondents' replies exactly as spoken. This generally requires rapid notation system in order to avoid the respondents losing interest. On the other hand, it is flattering to have someone write down one's statement in detail. Thus, there is a positive aspect even though some delays and interruptions are required.

The type of interview and scoring system will help determine how much the interviewer should record. The task of classifying answers into fixed categories is called **coding**. This task becomes more time-consuming when the coder has to wade through a lot of irrelevant detail. **Verbatim responses** are those written down exactly as spoken and would include ungrammatical statements, requests for more information, and so on. Paraphrased responses are the interviewer's impressions of what the respondent meant. Things that are irrelevant and unnecessary are omitted.

It is probably better for the beginner to write down more, rather than less details. What is irrelevant can be ignored at the time of coding. When dealing with groups of people who have a special way of expressing things, such as adolescents or professional engineers, it is important to record statements exactly as they are said. Instead of talking about the members' attitude towards authorities, the researcher can express their attitude in the words of the group members.

Tape recorders provide an advantage of precision and also will record hesitation phenomena, the various 'ahems' and 'ahs' and 'ers' made by the respondent. However, transcription is a long process, taking four or five times the amount of time spent in the interview itself. Many respondents are made uncomfortable by the presence of a tape recorder and confidentiality may be compromised. Thus, tape recorders are not recommended for most interviews.

In many instances an exact word- forword transcription is not necessary nor is it necessary to code the material for statistical analysis. In conducting in-depth interviews as a preliminary step to questionnaire construction, discovering categories of response is more important than recording every single word or anticipating statistical analysis. Jotting down notes from significant responses will be sufficient.

3.9 LET US SUM UP

There are several reasons why an important step in dissertation writing is to have literature survey on the concerned subject. Usually, the researcher reviews the literature before the research has begun so as to get an idea of where to proceed and to see if there are any pitfalls to watch out for. How much literature should be put in a paper depend on the topic and the way it is being examined.

Locating the sources of information becomes your next logical step in a survey of the previous research. Basically the sources of information are of two kinds viz. the primary and the secondary. A basic strategy of information gathering is called networking. When you visit the library for literature review you have to check the card catalogue, library shelves, reference section, periodical section and special document section. A single source can supply certain bits of information and point you toward other relevant sources.

Social scientists are dealing with complex literate societies in which the accumulation of documents of one sort or another has been going on for centuries. The important documentary sources of data are official records; newspapers; official statistics, diaries, memoirs and autobiographies; biographies; correspondence - private and public and historical documents.

With the development of reference service, it becomes necessity to consult a set of useful books, called reference books. These need to be evaluated before use.

A library is a collection of written, printed or other graphic or visual material organised and maintained for reading, study and consultation whereas archives may be defined as those records of any public or private institution which are judged worthy of permanent preservation for reference and research purposes and which have been deposited in an archival institution.

The ultimate aim of both the institutions is to have a reasonably, comprehensive collection, and to organise it in an orderly manner in order to achieve an efficient retrieval of documents when called for.

Despite these common interests, both are different too. Each deals with different formats of information packages and each controls information by different means. Invariably we have open access in libraries and closed access in archives.

A database is an organised collection of information. There are a number of databases relevant to the social and behavioural sciences. A good way to discover which database is likely to yield the most fruitful result is to consult the abstract/index it is based on.

3.10 REFERENCES

AGARWAL, S.P., *Social Information and Documentation*, New Delhi: Concept Publishing Company, 1987.

BEDEKAR, V.H., *How to write Assignments, Research Papers, Dissertations and Theses*, New Delhi: Kanak Publications, 1982.

GUPTA, P.K., and Usha Pawan, *Library and Information Science*, Jaipur: RBSA Publishers, 1986.

JAGDISH Saran Sharma, and D.R.Grover, *Reference Service and Sources of Information*, New Delhi: Ess Ess Publications, 1987.

MANN, P.H., *Methods of Sociological Enquiry*, Oxford: Basil Blackwell, 1968.

SANDERS, W.W., and Pinhey, T.K., *The Conduct of Social Research*, New York: Holt, Rinehart and Winston, 1983.

SOMMER, R., and Barbara B. Sommer, *A Practical Guide to Behavioural Research, Tools and Techniques*, New York: Oxford University Press, 1986.

VANIT, Nalwa., *The ABC of Research for Behavioural and Social Sciences*, New Delhi: Wiley Eastern Limited, 1992.

3.11 ASSIGNMENT

Prepare notes for group discussion on:

(a) Documentary sources of data

(b) Evaluation of reference books

(c) Library and Archives

(d) Steps involved in Literature review

3.12 RECOMMENDED BOOKS

BROADUS, R.N. *Selecting Material for Libraries*, New York: The H.W.Wilson Company, 1981.

FOSKETT, D.S. *Science, Humanism and Library, London:* Crosby Lockwood & Son Ltd, 1964.

KHANNA, J.K.*Advances in Librarianship*, New Delhi: Ess Ess Publications, 1985.

MANN, Thomas. *A Guide to Library Research Methods*, New York: Oxford University Press, 1987.

RANGANATHAN, S.R. *Documentation and its Facets*, Bombay: Asia Publishing House, 1963.

YOUNG, P.V., *Scientific Social Surveys and Research*, Bombay: Asia Publishing House, 1960.

3.13 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) What are the reasons for full literature survey before writing the dissertation?
2) Describe documentary sources of data and its utility for the social scientist.
3) How do you evaluate the reference books?
4) Explain the similarities and differences of activities between libraries and archives.
5) What is database? In what way it is useful to the social scientist?
6) What are the various steps involved in reviewing the literature?
7) Explain in brief the procedure involved in recording surveyed information.

II. SHORT NOTES

a) Characteristics of reference books
b) Difference between ordinary books and reference books
c) Locating sources of information

UNIT - 4 : FORMULATION OF HYPOTHESIS

Structure

4.0 AIMS AND OBJECTIVES

Formulation of hypothesis is a basic and important step in research methodology. The unit aims at introducing you to the need, importance and the process in the formulation of hypothesis.

After studying this unit you should be in able to

- describe what hypothesis is ?
- explain the need for hypothesis
- distinguish different types of hypothesis
- list out the qualities of good hypothesis
- describe the process from hypothesis to theory, law and axiom.

4.1 INTRODUCTION

In this age of science and logic we do not admit anything as valid until a satisfactory test of its validity is carried out. We may often start with an assertion or hypothesis and use our research data to prove or disprove it. Any hypothesis is tested with known statistical procedures and unless such tests are carried out, a research study may not be treated as complete. If the observed data are found to be inconsistent with the hypothesis, the hypothesis is rejected and if there is agreement it is accepted. Formulation of hypothesis and its verification are fundamental aspects of research. A hypothesis is a speculation about the phenomena. It is based on common sense, theory or empirical facts. The hypothesis is a necessary link between the theory and the investigation which leads to the discovery of additions to knowledge.

4.2 FORMULATION OF HYPOTHESIS

Formulation of a hypothesis gives a definite direction to the inquiry, aids in establishing direction in which to proceed, and helps to delimit the field of enquiry by singling out the pertinent facts on which to concentrate and by determining which facts should be set aside, at least for the time being. The use of hypothesis, thus, prevents a blind research and indiscriminate gathering of data which may later prove irrelevant to the problem under study. In this connection Lundberg observes:

"The only difference between gathering data without a hypothesis and gathering them with one is that in the latter case, we deliberately recognise limitations of our senses and attempt to reduce their fallibility by limiting our field of investigation so as to permit a greater concentration of attention on the particular aspects which past experience leads us to believe are significant for our purpose".

4.2.1 How to formulate your Hypothesis?

The formulation of hypothesis is very crucial and the success or the failure of a research study depends upon how best it has been formulated by the researcher. We may conclude by saying that it is hard to conceive modern science with all its rigorous and disciplined fertility without the guiding power of hypothesis.

Hypothesis is the statement of relationship between two or more concepts and/or social patterns.

A hypothesis is a piece of theory; it is a theoretical statement. It tells us where to look for if we want to find some particular sort of experience. A hypothesis is "a human device for anticipating the events that are about to happen to us". A hypothesis has the general form: "If **this** happens, then one will (with some probability) find himself observing **that**. An equivalent form is: "Suppose I were to do **this**: what would ensue? Would **that** perhaps occur?" The important thing about this mode of expression is that it is tentative and exploratory. The chief idea is not to "prove" that an idea is "true", but to open oneself to new possibilities and contingencies.

4.2.2 Working Hypothesis *Versus* Research Hypothesis

Very often a hypothesis being researched is referred to as a working hypothesis or research hypothesis.

Research hypothesis usually specify which variable is independent, which variable is dependent and what the anticipated form of the relationship is. A hypothesis seldom provides predictions about the extent and precision of a relationship in the social sciences. It is left to the research analyst to conclude if the form of the relationship is as predicted by the research hypothesis. The analyst should also provide information about the extent and precision of the relationship which will expand understanding of the theoretical and practical usefulness of the relationship found in the data.

4.2.3 When to make a Hypothesis?

Hypotheses are not given to us readymade. This is so especially in social sciences where there has not yet evolved a highly developed theoretical system in many areas of its subject matter affording fruitful bases for hypothesis formulation. As a result, in social sciences at least, a considerable portion of research endeavour is directed understandably toward 'making' hypotheses rather than testing them. Hence, it is to be remembered that research can begin with a well-formulated hypothesis (in the planning state of the project) or it can come out with hypothesis as its end product (at a later stage).

4.2.4 Need for Hypothesis

While hypothesis is useful it is not always indispensable. In physical sciences (including fundamental research) the hypothesis is most often necessary but in social sciences useful facts may be discovered, organised and presented purposefully even without a hypothesis other than the one which might incidentally be suggested in the course of the investigation. However, it is desirable to have hypothesis/hypotheses.

4.3 HYPOTHESIS - DEFINITIONS AND TYPES

4.3.1 Definitions

Goode and Hatt say, "hypothesis is a shrewd guess.... that is formulated and provisionally adopted to explain observed facts, or conditions, and to guide in further investigations".

In the words of George Lundberg, "A hypothesis is a tentative generalisation, the validity of which remains to be tested. In its most elementary stage the hypothesis may be any hunch, guess, imaginative idea, which becomes the basis for action or investigation".

According to Webster, "A hypothesis is a proposition, condition or principle which is assumed , perhaps without belief in order to draw out its logical consequences and by this method to test its accord with facts which are known or may be determined". Goode and Hatt have defined it as "a proposition which can be put to test to determine validity".

In very simple terms Kerlinger states "A hypothesis is a conjectural statement of the relation between two or more variables. Hypotheses are always in declarative sentence form, and they relate either generally or specifically variables to variables".

Rummel and Balline say "A hypothesis is a statement capable of being tested and thereby verified or rejected".

William Emory says "when propositions are formulated or empirically tested they are called hypotheses". Propositions are combinations of concepts designated by statements that may be judged true or false if they refer to observable phenomena. Concepts are abstractions from reality and are designated by words or phrases.

According to M.H. Gopal "It has been defined as a tentative solution posed on a cursory observation of known and available data and adopted provisionally to explain certain events and to guide in the investigation of others. It is in fact, a possible solution to the problem".

Palmer Johnson has beautifully described hypothesis as "islands in the uncharted seas of thought to be used as bases for consolidation and recuperation as we advance into the unknown".

It is clear now that a hypothesis is a provisional formulation or possible solution or tentative explanation or suggested answer to the problem facing the scientist. A hypothesis, therefore, is essentially tentative, likely to be modified during the investigation if the facts discovered in the course of the enquiry demand it. It is an explanation that need to be established before it can be accepted.

If the hypothesis is proved the problem of which it was a tentative solution is answered. If it is not proved, alternative hypotheses or situations would need to be formulated and tested. A hypothesis thus stands somewhere at the **mid-point of research;** and also look forward to data.

4.3.2 Types of Hypothesis

There are different approaches to the classification of hypothesis. Young refers to mainly to two types:

1) Working hypothesis, and

2) Explanatory hypothesis

However, these two terms have similar connotation. They are used on a very tentative basis in order to find out the fact. The word working emphasises instrumentality of the proposition whereas explanatory refers to the scope of going into the depth and width with various possibilities so far invisible. Explanatory hypothesis and relational propositions which strongly imply or state the existence of, or a change in, one variable causes or leads to an effect on another variable. The first variable is typically called the independent variable and the latter the dependent variable.

Hypotheses may further be classified into descriptive or rational. **Descriptive hypotheses** are propositions that typically state the existence, size, form or distribution of some variable. Eg. "The current unemployment rate in Tamil Nadu exceeds 20 percent of the work force" is a proposition about the size of the unemployment group. A proposition about the form of unemployment might be "The unemployment in Tamil Nadu is concentrated on male adults under the age of 25 who have less than 12 years of education". If these propositions are advanced as tentative statements about conditions in Tamil Nadu, subject to verification by research, they are hypotheses. Few more examples of descriptive hypotheses are:

"Husbands and wives agree in their perception of their respective roles of purchase decisions"

"Private brand purchasers constitute an identifiable market segment".

When a statement describes the relationship between two variables, it is called a **rational** proposition. eg.," Families with higher incomes spend more for recreation". In this proposition the relationship suggested is a positive correlation between variables. While the correlation implies some degree of interaction, the linkage strength is unclear and may not imply a cause-effect relation. On the other hand, when relationships between variables can be directly and more accurately observed, the linkage may state that A causes B or that B is a function of A. Eg.Loyalty to a particular store increases the probability of purchasing private brands sponsored by that store.

"High unemployment among male adults in Tamil Nadu is a function of current industrial recession.

4) **Hypothesis should be Related to a Body of Theory:** Science cannot develop if each study is an isolated survey. Science, however, can be cumulative only by building on the existing body of fact and theory. If the hypothesis is related to some theory, research will help to qualify, support, correct or refute the theory.

5) **A Hypothesis should be Empirically Testable:** It should be framed so that it makes possible to deduce logically certain inferences which in turn can be tested by observation in the field.

6) **Simple:** The hypothesis should be simple and to the point. Insight governs the essentials of simplicity. The more insight the researcher has into the problem the simpler will be his hypothesis about it. William Occam, an English Philosopher, described hypothesis as 'razor'. This means the hypothesis should be as sharp as a razor's blade.

7) **Brevity:** The hypothesis must be brief so that its observation may be possible. It would help for a better understanding of the underlying concept and meaning.

8) **Hypothesis should be Closest to Things Observable:** Failing this, it would not be possible to test their accord with empirical facts.

4.4.3 Difficulties in Formulating the Hypothesis

The formulation of useful hypothesis is one of the most difficult steps in scientific method. The chief difficulties in the path to the formulation of useful hypothesis are three.

i) **Lack of Theoretical Background:** First of all, what the researcher should have at hand is a scientific framework. The chances are that he may get one already in existence, the other is that he is having no such theoretical outline suitable for his proposition.

ii) **Lack of Logical Background:** Hypothesis not only lacks a definite and clear cut background, but also lacks the logical use of the theoretical background. In other words there is no logical background.

iii) **Lack of Knowledge of Scientific Methods:** It is not always possible to have complete information of and acquaintance with the scientific methods for the formulation of a hypothesis. This lack of scientific knowledge presents difficulty in the formulation of a hypothesis.

4.4.4 Functions of Hypothesis

Science would not have progressed so much had the aim of hypothesis been mere collection of data. A scientific inquiry is objective, but it does not mean aimless collection of data. In fact, no experiment is possible, no scientific work can be carried on without a purpose or proposition. The hypothesis has got unlimited importance for the scientific inquiry. The following functions are important:

(1) The hypothesis makes researcher sensitive to his research work.

(2) It performs the function of an intelligent guide.

(3) It establishes relationship among data.

(4) It helps in the formulation of the relevant concepts.

(5) It helps in explanation of the data.

(6) It suggests methodology to be followed in a particular inquiry.

(7) It also functions as a stimulator for the researcher.

(8) It gives a definite point to the inquiry.

(9) It aids in establishing the direction in which to proceed.

(10) It delimits the scope and the field of an inquiry.

The function of a hypothesis is to direct our search for the order among facts..... It is of considerable advantage if a systematic inquiry is begun with a suggested explanation or solution of the difficulty which originated it. Such tentative explanations are suggested to us by something in the subject matter and our previous knowledge.

Goode and Hatt have emphasized deep relationship between theory and hypothesis. One scientist in this connection has stated: "In practice a theory is an elaborate hypothesis which deals with more types of facts than does the simple hypothesis"... The distinction is not clearly defined". They have stated further that a hypothesis looks forward. It is a proposition which can be put to a test to determine its validity. It may seem contrary to or in accord with common sense. It may prove to be correct or incorrect. In any event, however, it leads to an empirical test. Whatever the outcome, the hypothesis is a question put in such a way that an answer of some kind can be forthcoming. It is an example of the organised skepticism of science, the empirical verification. Every worthwhile theory, then permits the formulation of additional hypotheses. These when tested are either proved or disproved and in turn constitute further tests of the original theory. In either case they may be of use to existing theory and may make possible the formulation of still other hypotheses.

There is little doubt that hypotheses are important and indispensable tools of the scientific research. They are the working instruments of the theory. Hypotheses can be tested and shown to be probably true or probably false. Hypotheses are powerful tools for the advancement of knowledge because they enable man to get outside of himself. Though constructed by man, hypothesis exists, can be tested and can be shown to be probably correct or incorrect apart from man's values and opinions. This is so important that we venture to say that there would be no science in any complete sense without hypothesis. In terms of functions, working hypotheses serve to guide and organise the investigation, providing us something to go on with.

Thus, hypothesis is the integral part of scientific research and without the former the latter shall neither be valuable nor possible. To sum up, hypothesis is the foundation of scientific social research.

4.4.5 Process : Hypothesis -> Theory -> Law -> Axiom

Discovering facts, establishing relationships between them and explaining situations and events so as finally to lead to rational generalisations and wherever possible to help to predict, constitute the process and purpose of research. The process itself is broadly in four successive stages, viz., the formation of a **hypothesis**, of a **theory**, and of a **law** culminating, if possible, in an **axiom**.

If the hypothesis has been tested and established at least to some extent and if such a conclusion on further testing is established over a number of investigations, perhaps over years of work, it becomes a theory.

4.4.6 Hypothesis and Theory Distinguished

A hypothesis is not the same as theory although the two are closely related. According to William George, theory is elaborate hypothesis. The hypothesis actually emerges from the theory. It is a generalisation drawn from the theory itself and when it has been tested and found correct it becomes a part of the theory itself. Thus, theory itself in its early form is only hypothesis and the two are interdependent upon each other. There is only difference in degree and form. Hypothesis is a stage that precedes formulation of a theory. According to Pauline Young, "provisional central idea which becomes the basis for fruitful investigation is known as working theory".

A theory verified on a large scale, more firmly establishing a regular relationship which may be believed as true and adopted as the basis of further inferences becomes a law. Thus, for instance, Malthus started with the hypothesis that population increases faster than food supply, and on the basis of the study of contemporary conditions formulated the famous theory of population. During the century that succeeded Malthus this theory was further tested and established the Malthus or Malthusian Law of Population. It is true that the law, has, in the 20th century, questioned as to its universal validity, but with regard to underdeveloped countries the Malthus Law still holds good.

If the law is so basic that it becomes the foundation of the belief that other ideas in the particular science can be organised around it and make other ideas deducible from it, the law becomes an axiom. Thus, the doctrine of the earth's gravitation, when it first occurred to Newton's mind was a hypothesis; the idea of universal gravitation, when confirmed by evidence from the moon's orbit, it became a theory; when the latter became verified by a study of the motions of many kinds of bodies, it stood as a law; and when the law of motions became central to astronomy and mechanics, it became an axiom.

4.5 LET US SUM UP

Hypothesis is the statement of relationship between two or more concepts and/or social patterns. A hypothesis is a statement capable of being tested and thereby verified or rejected. The formulation of hypothesis is very crucial and the success or failure of research study depends upon how best it has been formulated by the researcher. Formulation of hypothesis gives definite point to the enquiry, aids in establishing direction in which to proceed, and helps to delimit the field of enquiry by singling out the pertinent facts on which to concentrate and by determining which facts should be set aside, at least for the time being.

It is clear now that a hypothesis is a provisional formulation or possible or tentative explanation or suggested answer to the problem facing the scientist. A hypothesis, therefore, is essentially tentative, likely to be modified during the investigation if the facts discovered in the course of the enquiry demand it. It is an explanation that needs to be established before it can be accepted.

Hypothesis may be classified into working, explanatory, descriptive and relational hypotheses. When a hypothesis is stated negatively, it is called a null hypothesis. The object of this hypothesis is to avoid the personal bias of the investigator in the matter of collection of data. A null hypothesis is used to collect additional support for the known hypothesis.

There are several sources of hypothesis. It may be formulated from (1) problem or topic (2) cultural environment (3) cross cultural phenomena (4) personal experience (5) scientific theory (6) previous studies (7) analogies (8) current popular beliefs and practices.

In view of the importance of hypothesis a workable hypothesis must possess some characteristics viz., it should be (1) conceptually clear (2) specific (3) related to available techniques (4) related to a body of theory (5) empirically testable (6) simple (7) brief (8) closest to things observable..pa

Formulation of hypothesis is one of the most difficult steps in scientific methods. The main difficulties are (1) lack of theoretical background (2) lack of logical background and (3) lack of knowledge of scientific method.

A hypothesis is not the same as theory although the two are closely related. A hypothesis actually emerges from the theory itself and when it has been tested and found correct it becomes a part of the theory itself. Thus, the theory itself in its early form is only a hypothesis and the two are interdependent.

4.6 REFERENCES

GOODE, W. F. and Hatt, P. K. *Methods in Social Research.* New York: McGraw Hill Book Company, 1952.

KERLINGER, Fred N. *Foundations of Behavioural Research.* Japan: CBS Publishing, 1986.

REDDY, C. R. *Research Methodology in Social Sciences.* Delhi: Daya Publishing House, 1987.

SADHU, A. N. and Amarjit Singh. *Research Methodology in Social Sciences.* Bombay: Himalaya Publishing House, 1983.

SARAVANAVEL, P. *Research Methodology.* Jaipur: Printwell Publishers, 1990.

VERMA, R. K. and Gopal Verma. *Research Methodology.* New Delhi: Commonwealth Publishers, 1989.

4.7 RECOMMENDED BOOKS

KAPLAN, A. *The conduct of Inquiry.* Jaipur: Sachin Publishers, 1980.

LUNDBERG, G. A. *Social Research.* Longmans Green & Co., 1982.

RUNKEL, Philip J. and Joseph E. McGrath. *Research on Human Behaviour.* New York: Holt, Rinehart and Winston Inc., 1972.

SANDERS, William B. and Thomas K. Pinhay. *The Conduct of Social Research.* New York: Hold Rinehart and Winston, 1983.

WALIZER, Michael H. and Paul L. Wienir. *Research Methods and Analysis - Searching for Relationships.* New York: Harper & Row Publishers, 1978.

YOUNG, P. V. *Scientific Social Surveys and Research.* New Delhi: Prentice Hall, 1977.

4.8 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) What is a hypothesis ? Explain how to formulate a hypothesis ?

2) What are the main sources of a hypothesis ?

3) What are the qualities of a good hypothesis ?

4) Explain the functions of a hypothesis in social science research.

5) Trace the interrelationship between hypothesis, theory, law and axiom

II. SHORT NOTES

(a) When to make a hypothesis ?

(b) Working hypothesis

(c) Null hypothesis

(d) Difficulties in formulating hypothesis

BLOCK - II : RESEARCH METHODS

Scientific research is a systematic, empirical and critical investigation of hypothetical propositions about the presumed relations among natural phenomena. Therefore, research is an activity of solving problems which leads to new knowledge, using methods which are currently accepted as adequate by the scholars in the field. In order to understand the varieties of factors affecting the development of the people, there is a need to have methods of enquiry which will yield facts as they are. This requires collection of data in terms of figures, facts, numbers, etc. which can be processed, analysed and interpreted.

In social science research, the following terms are often used with reference to the data namely, primary data and secondary data.

(a) *Primary data*: Information which is collected by the researcher directly from the respondent by way of interview, observations, questionnaires or such other techniques. Primary data can be qualitative or quantitative.

(b) *Secondary data*: Provides information which is already available from various sources such as census, national sample survey, data from state bureau of economics and statistics, and all the sectoral departments dealing with different programmes or monitoring and evaluation cells of special programmes. The secondary data are generally quantitative.

An appropriate research method for collection of data should take into account the nature of the research problem, the purpose for which data is required, nature of the respondent, the locality and socio-cultural context in which the research has to be carried out and the time, effort and money required for managing the data collection.

A description of the important research methods in social sciences is presented in this block.

Unit-5 deals with Historical and Experimental Research Methods. Historical Research helps in understanding why, how and when past events occurred. There are a number of studies in LIS field using historical method. Though not many studies in LIS using experimental method, however, some significant studies like Cranefield Research Projects help us understand the importance of this method.

Unit-6 provides an overview of Survey Research and its application to library and information science field. Social surveys are generally confined to social problems of an area or situation to find out facts and to suggest remedial actions. It employs observation, questionnaire and interview methods in data collection. Libraries have long been conducting surveys in an attempt to understand about the use of their services.

Unit-7 deals with Case Study Method and Evaluation Research. Case study is a detailed and intensive study of a social unit. Evaluation research deals with testing of the application of knowledge within a specific programme of project. Case study and Evaluation research methods are used in LIS extensively.

Unit-8 provides an overview of Content Analysis, Documentary Research and other methods of research such as Delphi, Participatory research, Comparative studies, Sociometric techniques, Constructive typology, and Projective techniques.

UNIT-5: HISTORICAL AND EXPERIMENTAL RESEARCH METHODS

Structure

5.0 AIMS AND OBJECTIVES

The unit aims to introduce you to two important research methods used in social sciences - historical and experimental research.

After studying this unit, you should be able to

- explain the historical method of research and its use in LIS research
- describe how experimental method is used in LIS research with examples.

5.1 INTRODUCTION

The survey research attempts to examine the present situation and attempts to interpret it. In the present unit we will discuss about the historical and experimental research methods.

The historical research method investigates a past event and attempts to analyse it in the present situation. In other words, by discovering the past event the present situations can be studied. Thus, the historical research has great significance in understanding the evolution of a thing. In library and information science, historical research method can be applied to study the growth and development of different libraries, reading materials and their organisation, procedures and techniques, etc.

Though experimental method has a limited application in social sciences, library and information studies have used this method successfully in understanding the performance of indexing languages. The basic factor in this method is the control over the subject of study and manipulation of the independent variable to study its effect upon the dependent variables. Thus, an experimental study involves two groups of subjects - experimental group and controlled group. By subjecting a change in one variable, its impact can be observed. The famous Cranefield Research Projects conducted forty years back by Cyril Cleverdon and his colleagues at College of Aeronautics were the examples of experimental design in library and information science.

5.2 HISTORICAL RESEARCH

An inquiry into the past in a systematic manner is called 'historical research'. Historical method helps us to understand why, how and when past events occurred. By this study of past events, we can predict the future developments in a subject field.

The past has deep-rooted influence on our present social systems. P.V. Young said, "though today is different from yesterday, it was shaped by yesterday". Today and yesterday influence tomorrow. The past events stored in documents contribute so much in understanding the social system centuries later. Study of past also help librarians how they grew and how the procedures/ techniques of organisation of materials and services modified over the period.

In one sense all research is historical as it provides an accurate record of what happened in the past. Historical research depends upon observations which cannot be repeated in the same sense that a laboratory experiment or a descriptive survey. It always involves intensive library usage and scholarly patience. Historical research may be concerned with a historical problem in a special field as distinguished from the historical approach to the current problem in the field.

In other words, the primary aim before the researcher is not to make a deep-study of the history but to confine himself to that part of the history which is relevant to his/her social research. The researcher should always keep in mind that he/she is out for social research and not a historical research.

5.2.1 Elements of Historical Research

There are seven elements in the historical research. They are

(1) Choice of a problem

(2) Collection of material

(3) Classification of material

(4) Formulation of tentative conclusions

(5) External criticism of sources of data

(6) Internal criticism to determine the meaning of data

(7) Establishment and grouping of facts for interpretation.

Let us study these elements in detail.

(1) *Choice of a Problem*: The problem chosen for investigation by the researcher under this method should rigorously follow all the procedures. The study is, generally, carried out by only one individual and may not always carry an hypothesis.

(2) *Collection of materials*: The collection of materials is done as evidence for a solution to the problem. The procedure involves collection of data from all identified sources including oral research, critically examining the data in a sequence of occurrence of events/happenings, drawing inferences, tentative conclusions along with possible interpretations and presentation of the information in a narrative form.

(3) *Classification of materials*: The materials collected as evidence to solve a problem may be of different nature. They are generally classified as documentary and non-documentary type. Information for historical research is mostly obtained from available documentation viz., library, chronicles, biographies, diaries, genealogies, certain classes of inscription, records maintained by the government departments on different subjects, newspaper reports of events at a given point of time, etc. One has to scan through the subject indices of all national centres of documentation to identify the relevant sources. Now that most of these centres have computerised the sources of information, it would not be that difficult as it was in the past to get at the sources required. The only difficulty would be to conclude whether the sources required are available but not traceable or not available and hence not accessible. This would be a very difficult judgment for the historical researcher to make.

Other sources of information are, archival records maintained by the National or State Archives, visiting sites or places where events had taken place to identify historical landmarks (like in Ayodhya) and conducting oral research among those who were witnesses to an event or victims of an event. For example, the erstwhile members of the Constituent Assembly could be talked to find out the nature and extent of discussion on any provision made in the Constitution of India.

In the event of oral research being not possible, observation of places where events had taken place based on documentation could augment the information available or authenticate it. Certain physical factors (archaeological) could be delineated where relevant. Apart from this, artistic productions like historical paintings, portraits, scenic or portrait sculpture, coin types and figures on ancient coins could also serve as sources for certain types of historical research. Hence historical research purely depends on documentary evidence available and accessible though not verifiable.

(4) *Formulating tentative conclusions or Hypotheses*: As already noted that the historical research may not carry any hypotheses. Many social scientists feel that there is little scope for formulating tentative conclusions or hypotheses. This is because the reasoning in historical research is inductive. The style of reporting the results is mostly narrative. When evidence is not found to support a hypothesis, it may be because of a likelihood that the appropriate material has been lost rather than never having existed. Therefore, rejection of the hypothesis is not possible.

However, the historical research may be carried on certain assumptions. For example in LIS, one who studies the growth and development of libraries may formulate his/her assumptions as follows: "Academic library development in India in Post-Independence Period is largely due to support from UGC"; or The library movement in Andhra Pradesh is entirely people's movement".

(5) *External criticism of sources of data*: The external criticism deals with the value of material collected as sources of data. It evaluates the material used for analysis to detect anything spurious.

(6) *Internal criticism to determine the meaning of data*: Internal criticism aims at finding the actual meaning of the evidence collected. This relates to interpretation of the evidence gathered. The researchers have to be careful while relying upon the documentary sources as evidence.

Certain criteria may guide the evaluation of sources of information, viz., the reputation of the scholar whose documents are examined, time gap between the event and the writing of the document, the method used to prepare the document (consultation with others, from memory or in reference to other documents), the background of the author, his prejudices, commitments, etc. repetition of certain information of events by some scholars over a period of time, comprehensiveness and logicality of the information provided, whether document is the result of intensive research or observational in scope or only documentation on the subject available regardless. In view of these uncertainties, as already pointed out, emphatic conclusions, due to lack of verifiability of the documents and generalisation beyond reasonable limits cannot be done in historical research.

(7) *Establishing and grouping of facts for interpretation*: If the conclusions to be valid, the researchers need to be absolutely unbiased in their approach to interpretation. This suggests that the facts collected from the documentary evidence to be examined thoroughly and if possible oral testimonies should supplement them. The method of reflective thinking is applied to establish the interdependence among the facts through cause and effect relationship.

5.2.2 Essential Requirements of Historical Method

Historical research method appears to be very easy but it is not so. Besides the difficulty in collecting proper and reliable evidence to the problem on hand, several factors may influence the researcher in his/her interpretation. Many a time it is very difficult to adopt dispassionate attitude in adopting historical method. The researcher needs a high level of competency to gather data from various sources as well as to establish its authenticity. For this purpose, a researcher must have following qualities:

(1) *Social Insight*: The researcher should have adequate insight to understand the facts and interpret the social events to establish cause and effect relationship.

(2) *Dispassionate Attitude*: The researcher should never be misguided by preconceived notions, biases, emotions, guesswork and dogmas. He should be dispassionate while studying the facts and events.

(3) *Imaginative Capacity*: It is very difficult to collect minute details of the facts. Many times the events are not available in coherent form. Available fragments are to be assembled to get a complete picture. The researcher should have imaginative power to understand the events and interpret them properly.

(4) *Discriminating capability*: The researcher should possess discriminating capacity in selecting or rejecting a source of evidence according to its relevance and use.

(5) *Knowledge of the Study Field*: The researcher should possess complete knowledge of subject field. He should acquaint himself with the trends in his study area.

5.2.3 Advantages and Disadvantages

Some of the **advantages** of the approach are

i) There are some problems which cannot be solved in any other way.

ii) It sometimes is not possible or even desirable to attempt to duplicate a past situation in a laboratory or in a small field situation by examining the effect of a set of variables.

iii) It might be used as a practical device for alleviating an emotionally charged situation as well as providing some useful information towards the solution of a problem.

The **disadvantages** are:

i) Lack of vigorous control in matching past situations with present ones. Only the gross effects can be detected and seldom can the cause of these effects be directly attributed in a specific way to particular variables.

ii) There is a great tendency to generalise the findings far beyond the justified limits.

iii) The extent of data to be gathered before a conclusion is reached is a subjective decision.

iv) Many a time, there is heavy reliance on secondary sources.

5.2.4 Historical Method in LIS Research

An application of historical method to LIS field is usually called "Library History". Busha and Harter (1980) defined Library History as "the systematic recounting of past events pertaining to establishment, maintenance, and utilization of systematically arranged collections of recorded information or knowledge". Carefully conducted library history relates causes and results of events; it also often recognizes the social, economic, political, intellectual, and cultural environment in which these events occurred. The term *library history* is often applied to an account of events that affected any library or group of libraries, as well as to the social and economic impacts of libraries on their communities. An understanding of the origin of libraries and library techniques help us practise librarianship better. Nevertheless, a number of non-librarians have also conducted significant studies on the library history.

Historical research method can be applied to LIS field to study origin and growth of various libraries over a period of time - public libraries (Bodlean Library, Boston Public Library, Library of Congress, British Museum and Library, National Library of India), academic libraries (College and University Libraries) and Special Libraries), biography of a person (C.A.Cutter, John Cotton Dana, Melvil Dewey, S.R.Ranganathan, Poole and others), library associations (LA, ALA, SLA, ILA, etc), library movement (in India, in A.P. or in other states), or development of a library techniques (catalogue codes, classification schemes, indexing techniques), etc.

5.3 EXPERIMENTAL RESEARCH

It is considered to be the prototype of the scientific method of problem solving. It involves procedures for gaining knowledge by collecting new or fresh observations under controlled conditions. It involves the use of independent and dependent variables. The former could either be controlled or manipulated (experimental) to study its effect on dependent variables or criterion variables.

According to Festinger, the essence of an experiment may be described as observing the effect on a dependent variable of the manipulation of an independent variable". It describes the results of what will happen when a condition is changed in a carefully controlled condition.

A scientific experiment is a special kind of observation made under conditions that have been pre-arranged and carefully controlled. The procedure is designed to yield quantitative data and contributes to identify the causative factors for any phenomena. In other words, experimental approach can provide an answer as to why certain phenomenon occur or what could happen if a particular factor is changed. This helps us to establish the cause and effect relationship.

5.3.1 Basic Concepts in an Experimental Design

An experimental design involves two groups of subjects: a controlled group and an experimental group. The aim is to observe the change and developments in the variables dealt within it when a factor is manipulated in experimental group. In this approach we often come across terms like variables (independent and dependent), manipulation, law of single variable, causal hypothesis, cause and effect relationship, etc. Let us understand these concepts before studying about the experimental design.

Subject: A subject in an experiment is the basic unit on which an experiment is performed. Eg: library users (students or patrons) and librarians are the subjects in library-based experiments.

Experimental group: a group of subjects on which experiment is performed to observe effect of treatment (ie manipulation of an independent variable associated with it).

Controlled group: A group of subjects on which experiment is not performed. The group is selected for the purpose of comparison.

Variable: A variable is a changeable element in a phenomenon. There are two kinds of variables - independent and dependent.

Independent Variable: An independent variable is actually the experimental variable, which is manipulated in an experiment. It is also known as cause or causal variable.

Dependent Variable: A dependent variable is a subject variable associated with control group. It is also known as the effect or criterion variable.

Manipulation (Treatment): Manipulation is controlling the independent variables or handling it with skill. The independent variable in an experiment is introduced, changed or withheld to study its effect on dependent variable.

Cause-effect relationship: The results of the experiment showing the influence of manipulating the independent variable (cause) on the dependent variable (effect).

Causal hypothesis: The assumptions about the cause and effect relationship between the two variables.

5.3.2 Classification of Experiments

Experimental designs may be classified on different bases. The usual bases generally observed for classification of experiments are -

(a) Type of control:
 scientific or practical
 formal or informal
 actual or inferred

(b) Means of approach:
 analytical, comparative
 qualitative or quantitative

(c) Place/Setting where experiments conducted:
laboratory or field (classroom, library)

(d) Treatment of subjects:
individually or in groups

(e) Function or purpose:
to study direct effects or indirect effects (or causes)

(f) Time involved:
short or long duration

(g) Investigation agency (Sponsorship):
independent or sponsored
independently or cooperatively
by an individual or an institution

5.3.3 Approaches to Experimentation

Repeating the experiment on several groups under similar conditions is a possibility. In social sciences, the experimental approach can be undertaken under laboratory conditions or field situation. Three possible approaches are generally followed in experimentation:

One, in which there will be control and experimental groups, identical in all major characteristics contributing to the variance except for the treatment to be given or intervention to be made in the experimental group. This is sometimes referred to as matched groups' approach.

The *second* approach involves only one group, whose condi- tions as obtaining or to be intervened are studied before the experimentation and again after exposure to the experimental variable. The difference if found any is attributed to the intervention.

In the *third* approach, a group is identified and exposed to the experimental variable and outcome if any is attributed to the intervention made.

The first two approaches are relatively better than the last one where the pre-condition need to be assumed to be different than the resulting condition after the experimental exposure. Thus we have the control and experimental groups or matched groups approach, pre-post experiment approach and post approach.

5.3.4 Experimental Formulations

There are three categories of experimental formulations:

i) After-only or Post-test only experimental design,

ii) Before - After or Pre-test - Post-test experimental design, and

iii) Ex-post facto experimental design.

i) ***After-only or Post-test only Experimental Design***: This type of experimental study is carried out under social conditions which are not at all under control of the physical or natural conditions. Two groups of subjects, which are similar in all conditions, are chosen. One is called experimental groups and the other is controlled group. Experiment is carried out on the experimental group as per the pre-determined method. After the prescribed period both groups are observed and the results are measured. The results are compared and the changes that are observed in the experimental group are recognised as a result of manipulating the variables in the experiment.

ii) ***Before - After or Pre-test - Post-test Experimental Design*:** It is often referred to as a simple classical experimental design and is more reliable as it represents the so-called *four-cell design* (see Figure-1). In the pre-test (ie before experimentation) all the groups are selected, observed, and measured. There is one independent variable—the treatment—and one dependent variable. Subjects are assigned randomly to the control group as well as to the experimental group. The dependent variable is measured for both groups. After the pre-test, the treatment is introduced to the subjects in the experimental group only. The dependent variable is then measured again for both groups and compared. This is the *post-test*. However, the design has one limitation - it does not ensure to be free from the influence of external factors.

	Value of Dependent Variable before Treatment	Value of Dependent Variable after Treatment
Control Group		
Experimental Group		

Figure-1: The Four-cell experimental design

iii) ***Ex-post Facto Experimental Design*:** The Latin term 'ex-post facto' means after the fact, or retrospectively. This technique is generally used to study the varying influence of two identical factors. In this method, the investigator selects two or more groups of subjects. In group an event has taken place, while in the other the event has not taken place. The investigator proceeds from the present to the past and seeks to analyse what has already happened in an effort to isolate the cause of the events. The limitations of this method are i) Many factors influence a phenomenon and hence it is difficult to match subjects on a number of variables, ii) the investigator has no control over the events that have already occurred. Since there is no control, many researchers do not consider this method as an experimental research design.

5.3.5 Laboratory Experiments

In laboratory experiment, the variance of all or nearly all of the possible influential independent variables not pertaining to immediate problem of investigation is kept at a minimum. This is done by isolating the research in a physical situation away from the routine or ordinary living and manipulating one or more independent variables under rigorously specified, operationalised and controlled conditions. The laboratory experiment can and often does, isolate the research situation from the life around the laboratory by eliminating the many extraneous influences that may affect the dependent variable.

The greatest weakness of the laboratory experiment is probably the lack of strength of independent variables. Since laboratory situations are created for special purposes - the effects of experimental manipulations are usually weak. Another weakness is the artificiality of the experimental research situation. The temptation to interpret the results of laboratory experiments incorrectly is great. The laboratory experiments though have internal validity, lack external validity. However, laboratory experiments have three related purposes:

i) to discover relations under 'pure' and uncontaminated conditions

ii) to test predictions derived from theory primarily and other conditions secondarily.

iii) to refine theories and hypotheses, to formulate hypotheses related to other experimentally or non- experimentally tested hypotheses and perhaps most important to help build theoretical propositions.

5.3.6 Field Experiments

A field experiment is a research study in a realistic situation in which one or more independent variables are manipulated by the experimenter under as carefully controlled conditions as the situation will permit. Most field studies must operate with less control which is a severe handicap to the experimenter.

The *weaknesses* of the field experiment are mostly of practical nature.

i) The control of the experimental field situation is rarely as tight as in a laboratory and is contaminated by uncontrolled environmental variables.

ii) Another practical obstacle is the problem of randomisation.

iii) Attitude of the researcher guided by negativism.

iv) Lack of precision as the measures are not sensitive enough to pick up the messages of independent variable.

The field experiment has certain *advantages*:

i) The variables in field experiment usually have a stronger effect as the research situation is more realistic.

ii) Field experiments are well suited to the testing of the theory and to the solution of practical problem.

5.3.7 Planning an Experiment

We need to plan out things much in advance and in their minutest detail to get the best results out of an experiment. The experimenter needs to involve himself with the experiment. The following are the steps involved in planning an experiment in any field of study.

(1) Selecting the Problem

The researchers opting for experimental approach need to investigate the requirements in the subject field and decide upon a problem. A review of the literature related to the similar problems may give them a broad perspective. All kinds of problems cannot be studied through experimental method. One of the pre-requisites for the problem to be selected for experimental study should is that it should enable us to manipulate the independent variable to study its effect on the dependent variable. For example, study of the performance of various indexing techniques, impact of IT on library use, effect of innovative methods of teaching on students' performance in LIS schools, etc can be studied through experimental approach.

After selecting the problem, the hypothesis must be stated in conceptual and clear terms. The researcher has to decide the place, time, duration and materials of the experiment at this stage. The subjects and groups on which the experiment is conducted are to be selected. The variables that effect the phenomenon must be examined. The independent variables should also be decided and the plan for its gradual manipulation.

(2) Selection of Setting

The setting for a laboratory experiment is created automatically, while for a field experiment a natural setting has to be selected. The main features of a good setting are -

i) The situation must contain the phenomenon to be studied. For example, a study on the impact of IT on library use should have a library setting where the library should be computerised.

ii) The phenomenon must occur in the setting so that experimental research is feasible.

iii) The experimenter must have power to manipulate the independent variables. The experimenter must also control other factors which might confound the experiment. The library authorities should permit the researcher to conduct the experiment and the library users (subjects) should cooperate with the experimenter.

iv) The researcher must have freedom of access to the data as well as publish it.

(3) Pilot Study

Many a time a pilot study is necessary to study the realities, which may pose problems later. Pilot study enables the researcher to understand precisely the various causative factors involved with the experiment. This will also help the experimenter to appraise the authorities of the setting and the subjects about the aims and objectives of experiment and seek their cooperation.

(4) Research Design

Research design lays down the procedure for conducting an experiment. It lists the independent and dependent variables and shows the manner in which the experimenter to manipulate the situation to study it. The method of selecting of control groups and experimental groups, randomisation of samples to eliminate the problem of choice and personal bias, measurement of the effect of manipulation of independent variables, replication of experiment to confirm the findings, etc are decided in the research design. Therefore, it is an essential step for any experimental method.

(5) Conducting the Experiment

The major steps in conducting an experiment include

(a) Controlling the variable(s) or non-experimental factors

(b) Keeping a careful record of steps in the procedure

(c) Applying the experimental factor or factors, ie., manipulating the independent variables

(d) Measuring the experimental results

(e) Classifying, analysing and interpreting experimental findings

(f) Verifying the findings with already established facts

(g) Drawing conclusions from the findings

(h) Reporting the results

(6) Evaluation of an Experiment

Once an experiment is conducted, the results (the experimental data) must be evaluated. Evaluation involves a number of concepts, including hypothesis testing, experimental error, sensitivity, internal validity, and external validity. Busha and Harter (1980) have discussed these concepts in their book entitled, *Research methods in librarianship*.

In experimental research, generally, a null hypothesis is formulated. A null hypothesis in experimental design is a statement indicating that no significant difference exists between the control and experimental groups. A test of null hypothesis intends to show that the experimental variable has not effected any change in the value of dependent variable. For this purpose, *t-test*, *chi-square test* and *analysis of variance* have to be used appropriately.

The *sensitivity* of an experiment is its ability to detect relatively small effects. To increase sensitivity of an experiment the investigator increases the number of subjects.

The *experimental error* or *random error* is decreased by exerting additional control over the experiment. This is done by *matching* the subjects on the basis of as many variables as possible.

The *internal validity* of an experiment is assessed through the following questions:

Are variables sufficiently controlled?

Has randomization been employed throughout?

The *external validity* of an experiment is tested through the following questions:

Can the results of the experiment be generalised? If so, what is the population to which it can be generalised? When a sample is taken, to what extent the sample is representative of the target population?

Experimental findings are also evaluated in terms of the reliability of the data, the scientific importance of the results, and the extent to which the data can be generalized.

5.3.8 Experimental Method in LIS Research

Experimental method has limited use in LIS research. It was used at the library of College of Aeronautics at Cranefield, England by Cyril Cleverdon and his colleagues in his famous Cranefield Research Projects I and II. Project I was aimed at assessing the relative performance of various indexing methods and systems. The results of the experiment showed that all indexing languages performed equally well, the difference in performance between them being marginal.

Project II was a refinement over Project I in experimental design. The problem chosen for assessment was whether recall and precision devices improved in indexing language performance. A null hypothesis was formulated. To test the hypothesis, the documents (1400 research papers) and the questions (221 real research questions) were used as the controlled variables. These controlled variables were used consistently for each of the tested indexing language devices. The dependent variable was performance. Performance was measured by *precision* - the proportion of retrieved documents that are relevant and by *recall* - the proportion of relevant documents that are retrieved. A single performance measure, *normalized recall*, was also used. The operative rule of the single circumstance at a time was followed, ie. one device at a time being tested. The investigators reported the results of the experiment in the form of a ranked list of 33 indexing languages.

Experimental method can also be used to assess the performance of various indexing languages (such as PRECIS and POPSI) and classification schemes (eg:DDC and CC). The method can also be applied to study - impact of IT on library use, impact of bibliographic instruction programmes on library use, relative performance of library application software, impact of innovative methods of instruction on students' performance in LIS schools, etc.

5.4 LET US SUM UP

Let us recapitulate what has been discussed so far in this unit.

* The historical research method investigates a past event and attempts to analyze it in the present situation.

* The historical method is used to study the library history. There are a number of studies on growth and development of libraries, reading material, organisation, procedures and techniques, etc.

* The data for historical research is collected from documentary as well as non-documentary sources (including oral testimonies). Evaluation of these sources of data to determine their authenticity is a critical part of the research.

* Experiment is essentially a study of the effect on a dependent variable of the manipulation of an independent variable.

* There are four categories of experimental formulations: After-only (Post-test only), Before- After (Pretest - Posttest), Ex-post facto and Panel study.

* The pre-test - post-test experimental design represents the Four-cell design. It is often referred to as simple classical experimental design.

* The famous Cranefield Research Projects conducted by Cyril Cleverdon and his colleagues at College of Aeronautics, England was a significant study in LIS as it used experimental design systematically.

5.5 REFERENCES AND RECOMMENDED BOOKS

BUSHA, Charles H and Stephen P Harter. *Research methods in librarianship: techniques and interpretation*. New York: Academic Press, 1980.

INDIRA Gandhi National Open University. *Research methodology* (MLIS-E2). New Delhi: IGNOU, 1994.

KRISHAN Kumar. *Research methods in library and information science*. New Delhi: Har-Anand Publications, 1992.

SUKHIA, S.P., P.V. Mehrotra and R.N. Mehrotra. *Elements of educational research*. 3rd rev ed. Bombay: Allied Publishers, 1974.

5.6 ASSIGNMENT

Go to the Thesis Section of a University Library and critically examine the research methodology used in the M.Phil. and Ph.D. dissertations. Evaluate the historical and experimental formulations used in the theses.

5.7 MODEL EXAMINATION QUESTIONS

I ESSAY QUESTIONS

1) What is Historical research ? Describe the elements of historical research.

2) Enumerate the characteristics expected from a researcher working on a historical problem.

3) Describe the steps involved in designing experiments, with special reference to LIS research.

4) Discuss the formulations used in experimental research.

II SHORT NOTES

a) Library history

b) Field experiment

c) Panel study

d) Cranefield Projects

UNIT - 6 : SURVEY RESEARCH

Structure

6.0 AIMS AND OBJECTIVES

The aim of this unit is to introduce you to the Survey Research Method and its use in library and information science.

After studying this unit, you should be in a position to

- explain survey research
- describe the various methods involved in survey research
- discuss its application in LIS research.

6.1 INTRODUCTION

The word *survey* has been derived from the words 'sur' or 'sor' and 'veeir' or 'veoir', which mean 'over' and 'see' respectively. Therefore, the literal meaning of survey is 'to look over something' from a high place, but it has come to be used as as a term with specific connotation of its own. The word 'Survey' is used for a technique of investigation to collect information on a phenomenon through observation, interviews, questionnaires, etc.

According to Mark Abraham, a social survey is "a process by which quantitative facts are collected about the social aspect of a community's composition and activities".

Herman Morse defines the survey as "a method of analysis on scientific and orderly form for defined purpose of given social situation or problem or population".

Social surveys are generally confined to social problems of an area or situation to find facts and to suggest remedial actions. Libraries have long been conducting surveys in an attempt to gather information about their services and use.

The four main characteristics of the survey method are:

(1) It is directly concerned with social life as it exists here and now; what is observed, described, collected as a body of facts about current situation and problems.

(2) It focuses upon a given locality or geographical area.

(3) It leads to measures of reform on the data assembled both for improvement and the means to obtain them.

(4) It must be conducted in an objective spirit.

6.2 SURVEY RESEARCH

Survey implies taking a general view of all cases or situation - a broad general consideration of a whole situation or field of circumstances. A social survey is usually an enquiry into the composition, activities and living conditions of a group of people.

Social surveys may be divided into two broad categories: 'topical' and 'general'. *Topical* surveys treat certain specific aspects, while *general* surveys follow a broad approach and deal with a number of facets of a given social situation.

Surveys are commissioned for a variety of purposes: the general aims are *description* and *explanation*.

Most of the surveys have description as their main aim where an entire sample can be examined or different subgroups could be compared. Surveys undertaken for descriptive purposes can take the form of factual reporting of a population based on different characteristics (eg.

Census, Children's atlas) or description of urban slums in terms of living conditions (housing, water supply, electricity, roads, facilities and services) or village surveys or monographs giving a detailed descriptive account of the village - geographical, ecological, social, economic, political, occupational, educational, climatic aspects, etc.

A *descriptive survey* tries to answer questions like 'what' and 'how' it is. One may attempt a descriptive survey of population based on only age and sex, how many are children, how many are old people, what is the percentage that constitutes the workforce, what is the proportion of females to males, how it has changed over time, etc. A descriptive survey provides the base for an explanatory survey and hence is useful to understand a given situation, problem area, or groups of people.

A schedule for descriptive survey takes the form of a check-list, usually precoded and exhaustive in scope covering all physical, social and human dimensions.

Surveys with an *explanatory purpose* attempt to establish the connections among variables. An explanatory objective requires a multivariate analysis, examining the distribution of respondents *on at least two* and usually more than two variables simultaneously.

6.2.1 Approaches to Surveys

Based on the approaches, surveys have been further classified into i) cross sectional survey, and ii) panel and trend studies.

(1) Cross Sectional Survey:

Cross Sectional Survey is the most frequently used survey design, sometimes called the single shot survey. In this design, people are asked questions at one point of time. We might use this data either to describe our sample on each of the variables studied or to demonstrate the association between the variables.

The cross sectional survey has important limitations:

i) it is difficult to establish the time order to variables;

ii) it is difficult to analyse change over a period of time.

(2) Panel studies and Trend studies:

These two types of survey designs help to compensate the limitations of the cross-sectional survey.

In a *panel study*, the same people are interviewed two or more times. For example, some library users are interviewed at different times to seek their views on library product. In a trend study two or more different samples of people are drawn at different times from the same population. Trend studies provide information on net changes whereas panel studies provide information on both the net and gross changes. Trend studies can describe how the distribution of a variable is changing in the population studied. Panel studies can describe how individual members of a population are changing.

Trend studies are particularly valuable for describing long term changes in a population. Researchers may also use control procedures to determine which part of a population is changing. They may also examine the trends in the relationship between variables. While trend studies can be based on a comparison of survey data originally collected for other purposes, panel studies require specially collected data. Panel studies can be used to measure variables based on information collected at two or more points at a time. Panel studies are valuable for unraveling the causal order of variables, a problem that confronts the one-shot survey.

A popular type of trend study which in some ways approximates a panel design is 'cohort analysis'. A cohort is a group of individuals (other units of analysis) where all experienced a significant event during the same period of time. For example, there can be a cohort of people who entered the college at one time, a birth cohort can include all people who were born in a specific year. Marriage cohort might include people who married between specific years, etc. Usually a cohort analysis involves comparison of experiences of a number of different cohorts over time. The term generation cohort is sometimes applied to birth cohort. People in one generation may behave differently or hold different attitudes than people in another generation for various reasons. Cohort analysis traces the experience in each generation over time. Thus it allows us to determine whether a particular generation is responding differently to current conditions than other generations.

6.2.2 Data Collection in Survey

Surveyors gather data by contacting and questioning individuals, the unit of analysis is most frequently the individual. The contextual analysis is concerned with the effects of the group characteristic on the individual respondent where the group characteristics becomes the independent variable and individual's behavioural dimension,the dependent variable.

Survey research is usually based upon probability sampling from a large population. Therefore, survey research refers to a method of collecting data where individuals are questioned, and their answers are compared or aggregated to understand the group characteristics.

The method of collecting data involves using a schedule coupled with personal interview. There can be open ended or closed questions or with fixed responses. Open ended questions are primarily used in exploratory surveys. Our goal should be to write clear, simple and understandable questions that are both accurate reflections of the respondent's true feelings and behaviour and appropriate measures of important variables.

Some of the ***guidelines*** for writing questions are

i) The questions on a particular survey must be understood in the same way by all respondents in the sample who differ in education, experience and belief.

ii) Each question should contain a single idea.

iii) Questions must provide categories for the respondents' answer which are exhaustive and which include all the answers that most respondents would wish to give.

iv) The response categories must be mutually exclusive.

v) Question must be specific.

vi) Questions must be unbiased and neutral in tone.

vii) The question writer should be familiar with the lives and languages of those being surveyed.

All the other precautions for questionnaire or schedule preparation as discussed under questionnaires and schedules should be scrupulously followed.

Personal interview method is followed in conducting surveys and the process of conducting the interview as discussed under interview method should be followed.

The survey data needs to be coded for statistical analysis. The process of transforming verbal responses into a set of numbers is called coding, the researcher divides the responses given to a particular question into several categories and then assigns numbers to each of these categories. The coded numbers are transferred to IBM sheets for computer analysis. Prior to coding, the coders should be given enough training with regard to the method of coding and also the need for high coding reliability.

6.2.3 Advantages and Disadvantages of Survey

Some of the **advantages** of this method are:

i) Survey is a powerful technique because, within known limits of error, the responses of a relatively small sample of people can be generalised to a very large population. The researcher should be familiar with the lives and frames of reference of prospective respondents before conducting a survey.

ii) The questionnaires developed need to be pretested to keep the error to the minimum. The interviewers need to be trained through videotape practice interviews and improve the interviewing techniques as it otherwise contributes to errors.

Some of the **disadvantages** are:

i) Survey does not penetrate deeply below the surface level of information.

ii) It is demanding of time and money.

iii) It is subject to sampling error.

iv) The survey interview can temporarily isolate the respondents from their social context which may make the results of the survey invalid.

v) The disadvantages indicated under questionnaire and schedules and interview method also occur in measurement in survey research too.

To conduct survey, therefore, one requires a good deal of research knowledge, a sound grounding in sampling procedures, questionnaire and schedule construction, techniques of interviewing and analysing data.

6.3 QUESTIONNAIRES AND SCHEDULES

Questionnaire refers to all data collection forms. The content is aimed at finding facts or opinions or both. The questions are directed to persons about themselves or their families or their occupations, etc. The questionnaires could be used to collect data from knowledgeable groups or sent by mail or administered in a personal interview.

There is a slight difference between a questionnaire and a schedule. Schedule is a form that is filled up during the personal interview. Questionnaire is a form used for self-administration, usually mailed to the respondents. In both the kinds of instruments, there will be a set of questions or items to which the individuals respond either through closed or fixed alternative responses or without any restriction on the nature of response.

6.3.1 Content, Factors and Wording of Questionnaires

The appropriate form a questionnaire will take depends upon (1) the mode of administration, (2) the nature of the subject matter, (3) the sample of people to be reached, and (4) the kind of analysis and interpretation intended.

The questionnaire *content* can be decided based on the following guidelines. Decide on

1) what information is to be sought
2) what the questions should cover, and
3) what type of questions to be used.

The content of the questionnaire depends upon several behavioural factors. Therefore, the content should mainly aim at finding out (1) facts, (2) beliefs, (3) the feelings, (4) discovering standards of action (norms), (5) present or past behaviour, and (6) conscious reasons for beliefs, feelings, policies or behaviour. In other words, questions to discover the 'why' of behaviour.

In deciding the question content and preparing questions for any kind of investigation, the following *factors* should be considered:

a) the nature of the question,

b) the number of questions - one or several questions, or

c) whether additional material is required or not,

d) whether intensity of feeling or conviction is being sought,

e) assessment of respondents having the information,

f) check whether the question calls for information which the respondent cannot give or give reliably,

g) the type of information sought, and

h) identify the respondent who may have the required information

i) whether the question needs to become concrete, specific and closely related to personal experience,

j) whether the question is biased or loaded in one direction or

h) whether the subject matter is unfair in anyway,

i) whether the respondent will give the information that is asked for,

j) whether the questions will encounter emotional influence (desire to falsify) and so on.

Some amount of thought also needs to be given in regard to the *wording of questions*. Here the decision should be in regard to the kind of words to be used to make the question clear and meaningful to the respondent by avoiding difficult sentences or difficult or unclear words, providing for adequate expression of alternatives, avoiding misleading questions, biased wording, emotionally loaded expressions, using stereotypes, etc. The wording to the extent possible should not be objectionable to the respondent.

It is also necessary to decide whether the question can be asked directly or indirectly or in projective form.

The decision regarding the form of response to the question should be in terms of whether one should have a checklist type, free answer, or with fixed alternatives on the scale. The decision in regard to using closed or open-ended question should depend upon the nature of the subject matter, the type of the respondent and the extensity or the intensity of the information that is sought.

6.3.2 Sequence of Questions

One has to take a decision in regard to the place of a question in the questionnaire in terms of the sequence to protect against 'mental set' or 'response set' either inclined favourably or unfavourably. It is believed that the content of the preceding question may influence the answer to the following question. The question placed first should be the easiest to answer. Interest catching questions may be used in the beginning. Placing a question early in the questionnaire that can affect the answers to later question should be prevented where possible. A time sequence should be observed in the arrangement of questions. The sequence based on subject matter may be retained. It is wise to have questions arranged from more familiar to unfamiliar and from more important to less important.

It is also necessary to safeguard the questionnaire from the influence of socially desirable responses as people tend to give responses that are socially acceptable; responses that indicate or imply approval of actions or things that are generally considered to be good. This may also be borne in mind while asking the questions for which there is always social approval or sanction.

A questionnaire also takes the form of a scale. A scale is a set of verbal items to each of which an individual responds by expressing degrees of agreement or disagreement or some prescribed mode of response.

In the preparation of the schedules, the same procedure can be followed. However, a schedule has certain amount of latitude in that there is a possibility of explaining the content to the respondent as it is coupled with personal interview. It also provides for probe in certain questions which implies calling for associative material as suggested by the responses to the primary question. There can be schedules of different types depending upon the purpose for which they are used like interview schedules, observation schedules, rating schedules, etc.

6.3.3 Advantages and Disadvantages of Questionnaires

Some of the **advantages** of the questionnaire method are:

(a) It can cover a wider geographical area and reach a much larger population with given funds than personal interview.

(b) It eliminates the expensive and time consuming task of training of field investigators.

(c) The respondent may answer many questions more frankly by mail since anonymity is assured.

(d) In answering mailed questionnaire, any personal antagonism to the investigator is avoided.

(e) The questions are standardised and thus eliminate the implied or suggested answers given by the investigators during personal interview.

(f) The questionnaire can be answered at the convenience of the respondent.

The *limitations* of this method are:

(a) The returns from mailed questionnaire sent out to the respondents are normally low, ranging from 10 to 20 percent. The respondent may misinterpret questions, omit essential items or send in material which cannot be tabulated.

(b) It is difficult to check the honesty and reliability of returns. It is practically impossible to return unsatisfactory/incomplete schedules to the informant for correction

(c) Preparing an upto date address of potential respondents is a difficult task.

6.4 INTERVIEW METHOD

Unlike questionnaire, interview places heavy reliance on verbal report. It gives greater opportunity for flexibility in eliciting information and also provides an opportunity to observe both respondent and the total situation.

6.4.1 Types of Interviews

There are different types of interviews classified depending upon the purpose or method. Some of these are:

(1) *Focussed Interview*, where the emphasis will be upon a given experience and its effects;

(2) *Clinical Interview or Depth Interview*, concerns itself with the broad underlying feelings or motivation in the course of one's life experiences;

(3) *Non-directive Interview*, where interviewer serves as a catalyst, used for intensive study of perceptions, attitudes, motivations; and

(4) *Directed Interview*, where the interviewer plays an important role in eliciting information from the interviewee on a variety of issues using a schedule.

6.4.2 Goals of the Interviewer

Interview is a very good tool for research in collecting data which is not otherwise accessible by any other method. The interviewer actually participates in data gathering and can modify the situation as required in terms of clearing misunderstanding, noting special happenings and picking up clues that are valuable for interpretation. However, the respondent's self-esteem may have to be ensured by not asking questions which will affect him. The goals of the interviewer are:

i) to secure the respondents' consent for interview;

ii) to motivate the respondents to give full and honest answers;

iii) to read each question exactly as written and in the order that it appears on the schedule;

iv) to ensure that the respondents' answer is relevant to the question;

v) to record the answer accurately.

6.4.3 Method of Conducting Interview

The quality of the interview depends upon the study design. The interviewer has to create a situation which can be characterised as permissive in order to facilitate the respondent to voice his opinion without fear.

The interviewer should encourage the respondent without showing surprise or making value judgment on his opinions.

The interviewer's manners should be friendly, courteous, conversational and unbiased. The whole idea is to put the respondent at ease and establish rapport.

The interviewer should also have a thorough mastery of the schedule in order to avoid awkward pauses without showing any surprise or disapproval of the answers given by the respondent. In case the respondent asks for the interviewer's views on any issue, he should avoid answering it directly by a pleasant remark and he should try to keep the direction of the interview by discouraging irrelevant conversation. The question should be asked as far as possible as it is worded in the schedule so that uniformity could be maintained across the respondents. Impromptu explanation of the question should be generally avoided as it may change the frame of reference or bias the response. The interviewer may repeat the question slowly with proper emphasis offering only the explanation as may be specifically authorised. However, the method of conducting interview can remain very rarely uniform as the personalities of the interviewers differ.

Obtaining a complete response for any question may be difficult. Whenever any answers like 'do not know' is given, care should be taken to find out whether the 'do not know' is due to lack of opinion, fear to speak, reluctance to focus on the issue, vague opinions or lack of comprehension.

6.4.4 Method of Recording the Responses

In regard to the reporting of the response, it is better to record the response verbatim and the interviewer should have some idea of the coding process. Some of the ways of recording could be as follows:

i) Write the response as the respondent talks.

ii) Can use common abbreviations.

iii) Should not erase any recordings but may cross when required.

iv) May depart from the ideal verbatim recording to the extent of using a telegraphic style.

6.4. 5 Advantages and Limitations

Some of the *advantages* of the interview method are:

i) It can be used with illiterate sample.

ii) Many people may cooperate in the enquiry if all they have to do is to talk.

iii) It provides for greater flexibility.

iv) It provides a better opportunity to appraise validity of reports.

v) The information received is likely to be more correct than that secured by other techniques, since the interviewer can clear up seemingly inaccurate answers by explaining the question to the respondent.

vi) Return visits to complete items on the schedule or to correct mistakes can usually be made without annoying the respondent.

vii) Questions about which the respondent is likely to be sensitive can be carefully sandwiched by the interviewer.

viii) The language of the interview can be adapted to the ability or educational level of the respondent.

Some of the *limitations* of interview method are due to interviewer's inaccuracies and biases.

We should not assume that all interviewers behave ideally. Even if they are completely conscientious and honest, they are capable of unintentionally distorting interviews. The errors made by the interviewers can be classified into four types:

(1) *Asking Errors*: Where the interviewer may ask questions in different order, change the wording of the question, neglect to mention one of the response alternatives, etc.

(2) *Probing Errors*: Where the interviewer failed to probe in cases where the answer had not matched the objective of the question, used the probes that changed the meaning of the question, used biased probes that emphasised certain of the response alternatives, prevented the respondent from saying all that he or she wishes to say, used probes that are irrelevant to the respondents' answer, etc.

(3) *Recording Errors*: Where the interviewer recorded something that was not said, failed to record something germane to the question that was said, put down 'No' when the respondent said 'Yes', etc.

(4) *Flagrant cheating*: Where interviewers did not ask a question but nevertheless recorded a response even when the respondents had refused to reply the question, skipped interview, etc.

Some of the *biasing factors* introduced by the interviewer are

(1) The differences between one interviewer and another is not only asking questions, eliciting responses but also in recording.

(2) One's own stand on any issue or opinion or strong views on the subject matter need not necessarily constitute interviewer bias but unguardedly may influence the interview.

(3) Human as they are, the interviewers' reports need not be complete, comparable and valid.

(4) Influence from the type of mutual perception between the interviewer and the respondent, each influenced by one's own insights, wishes, expectations and personality.

(5) Interviewers' perception of the study in terms of its expected results - positive or negative or support his committed theories or concepts in which case he is likely to introduce bias.

In view of these bias factors, the need for proper selection and training of interviewers is emphasised before they are inducted into collecting information. The selection may be used on their capacity to be objective, non-committal and have a positive attitude to people. The training should be in terms of developing rapport with the respondents, communication skills, method of asking questions, probing and recording the answers.

6.4.6 Rules of Interview

Different authors lay down different sets of rules governing the conduct of an interview. Some of them are:

(1) Make a gradual approach to the interviewer by developing rapport and identification and by establishing pleasant association.

(2) Help the interviewer to feel at ease and ready to talk

(3) Let the interviewee finish his story, then help him to supplement it by checking with him dates and events

(4) Take pains to phrase questions which are easily understood; avoid asking embarrassing questions which encroach upon personal privacy

(5) Be straightforward and frank rather than shrewd or clever

(6) Listen to the respondent in a patient and friendly way intelligently and in a critical manner

(7) Avoid the role of the teacher; make no display of authority

(8) Give no advice or more admonition

(9) Do not argue with the respondent.

(10) Listen to not only what the respondent wants to say but be sensitive to what he does not want to say or cannot say without help.

(11) Avoid discussing the implications of the answers to the questions.

(12) Never hurry the interviewee; give him ample opportunity to finish his views and qualify his answers.

6.5 OBSERVATION METHOD

Observation is not only the most pervasive activity of daily life but also it is a primary tool of scientific enquiry. Observation becomes a scientific technique to the extent it,

i) serves a formulated research purpose;

ii) is planned systematically;

iii) is recorded systematically and related to more general proposition rather than being presented as a set of interesting curiosa;

iv) is subject to checks and controls on validity and reliability.

6.5.1 Definition and Purpose of Observation

P.V. Young defines observation as "a systematic viewing, coupled with consideration of the seen phenomenon in which main consideration must be given to the larger unit of activity by which the specific observed phenomenon occurred".

Observation is probably the oldest method used by man in scientific investigation. Though it was crude in the beginning, with use of better instruments over the time, observations have become scientific and more perfect. In social sciences too the earliest method of investigation was probably the observation. Attempts were made to devise ways and means to make observation studies more clear, precise, unbiased, pointed and infallible. Now, observational study in social sciences has become a more popular and dependable method of social research.

Observation may have a variety of research *purposes*:

i) used in an exploratory fashion to gain insights and to cover supplementary data;

ii) used as a primary method of data collection;

iii) can take place in real life situation;

iv) can be flexible completely, having general ideas about possible aspects of the importance of the problem.

The *five broad questions* in observation method are

i) what should be observed ?

ii) how observation should be recorded ?

iii) what procedures should be used to try to ensure the accuracy of observation ?

iv) what relationship should exist between the observer and the observed ?

v) how can such a relationship be established?

There are *three* types of observations:

(1) Simple (uncontrolled);

(2) Systematic observation (controlled) and

(3) Mass observation (which is a combination 1 and 2)

6.5.2 Simple Observation

Most people rely on uncontrolled (simple) observation for knowledge about social situation. It consists of the simple forms of looking and listening, which not only contribute to the basic, varied stock of knowledge about social relations but which also serves as a principal technique for gathering social research data.

There are two ways to conduct simple uncontrolled observation: non-participant observation and participant observation.

Non-participant observation consists of looking at specific social situations but not taking part in them. The presence of the non-participant observer should be made as inconspicuous as possible. This method sometimes gives a better insight than all the formal methods combined on how people live and why they act as they do.

The *participant* observer is an outsider who temporarily becomes an insider. He thereby secures a better insight into the situation he is investigating as he is not personally involved and can remain detached while at the same time taking part in the groups' activities and sharing their feelings and prejudices.

The observer may participate actively in the group he is observing. He may be defined as a member of the group but keeps his participation to the minimum. The observer may not remain part of the group or his presence may be unknown to some or all the people he has observed. In general, the degree of structure and the degree of participation tend to vary with the purpose of the study. In an explanatory study, it may remain relatively unstructured and the observer is more likely to participate in the group activity than he is in a study focused on accurate description of a situation.

The participant observer study requires a special kind of devotion to research. To gather data, the investigator must give up temporarily at least his own comfortable way of life and adopt that of the others. At the same time, he should be especially skillful in maintaining his objectivity. Though he may not be able to seek intelligent information without considerable familiarity, becoming too friendly may lead him to become personally involved or to treat as common place the very behaviour he should be carefully describing as relatively unique and interesting. The success of such a study is also dependent upon personal characteristics of the observer in order to make him acceptable to the group rather than his skill as a researcher. These studies are generally unstructured in nature as a researcher may not know which variables are more relevant and, therefore, these kind of studies are more efficient in explorative work.

6.5.3 Systematic Observation

Systematic observation is controlled. It differs from simple observation in that its observational techniques are standardised and both the observer and the observed are under control. Controlled observation consists of -

i) Careful definition of the units to be observed and the information to be recorded.

ii) The selection of pertinent data for observation.

iii) Standardisation of conditions of observation - time, place, persons, etc.

iv) Wherever possible, the use of mechanical devices, tests and other aids to accuracy.

6.5.4 Mass Observation

Mass observation is a combination of controlled and uncontrolled observations. It depends on observing and recording of information by a number of people and the pooling and treatment of their contributions by a central person.

6.5.5 Prerequisites for an Observational Study

Some of the *directions* for observation study are,

i) **knowledge** of participants in terms of who the participants are and how they are related to one another and what kind of characteristics dominate the participants and what structures or groupings exist among participants,

ii) **setting**: social situation in different settings, kinds of behaviour encouraged, discouraged, permitted and preferred and so on,

iii) **purpose**: purpose of the meeting-official, social or chance,

iv) **social behaviour**: what actually occurs in terms of what they do, how, with whom, and with what they do it.

In regard to *behaviour*, the following may be of interest:

i) what was the stimulus or event that initiated it,

ii) what appears to be its objective,

iii) towards whom or what is the behaviour directed;

iv) what is the form of activity entailed in the behaviour;

v) what are the qualities of behaviour;

vi) what are its effects, etc.;

vii) Frequency and duration:
when and how long a behaviour occurred?
Is it recurring or unique occasion giving rise to it and
how typical are such situations as the one being observed.

These five aspects will give the advantage in planning the content of observation of activities.

6.5.6 Aids to Observational Study

Five devices are commonly used in field observations. They are

(1) *Diaries*: The keeping of private diaries where recording of events as they happen are useful because observations in the field are recorded while the impressions are fresh and vivid.

(2) *Note Books*: Notes serve as an aid to the memory and are particularly useful in stamping impressions upon the minds of investigators.

(3) *Schedules*: It draws up a list of points to be covered and helps the investigator to put down answers in the field.

(4) *Photographs*: These are more realistic and revealing than words and they remind field workers of certain details that the pen failed to record.

(5) *Maps*: These are useful in showing different types of information and facts by preparing a basic map of the community.

6.5.7 Advantages and Limitations of Observation

Some of the *advantages* of this method are,

i) It makes it possible to record behaviour as it occurs.

ii) Ir deals with respondents who are not able to give verbal reports of either of their behaviour or feelings.

iii) It is also independent of a respondent's willingness to report apart from ability.

Some of the *limitations* of this method are:

i) Impossible to predict the spontaneous occurrences of an event to enable us to observe it.

ii) The practical possibility of applying observational techniques is limited by the duration of events.

iii) Some occurrences which people may be able and willing to report are very rarely, if ever, accessible to direct observation.

6.6 SURVEY METHOD IN *LIS* RESEARCH

Survey method is the most extensively used research method in LIS than any other method. There are two types of surveys studied in librarianship: *Library surveys* and *Community surveys*.

6.6.1 Library Surveys

According to Busha and Harter (1980), "library surveys are systematic, in-depth examination of libraries, library systems, or networks of libraries". Library surveys are conducted in an attempt to gather information about every or many aspects of libraries.

Library surveys may be categorised as internal surveys and external surveys. *Internal surveys* are self-studies conducted by the collective library staff or by a single designated member or committee from the staff. *External surveys* are accomplished by a person or persons from outside agencies, ie., library associations, library schools, consulting firms, national/regional library authorities, etc. Due to increased professionalisation and academic excellence among the internal staff, librarians have reduced their dependence on outside consultants for surveys.

These surveys are not restricted to libraries of any one type (public, academic or special) or one location (national, regional or local; urban or rural).

In library surveys, three broad categories of information are gathered and reported: i) existing library conditions, ii) comparisons between present conditions and desired standards or goals, and iii) suggestions for the improvement of existing conditions.

The data collected in library surveys may be placed into the following categories:

(a) the library environment, including factors such as facilities, buildings, organisational structure, location, space, resources, and so on;

(b) the characteristics of library personnel, including features such as educational attainment, experience, sex, age, group memberships, and so on;

(c) the nature of users and non-users, including their age, socio-economic backgrounds, sex, ethnic origins, communication patterns, recreation activities, reading habits, and so on;

(d) the nature of certain library services (or the absence of services), including reference and interlibrary loan, audio-visual programmes, bibliographic and directional assistance to clientele, programmes for children and young people, book-mobile services, adult services, photocopying facilities, extension services, reserve book services, and so on;

(e) nature and extent of library resources, including books, periodicals, pamphlets, documents, films, recordings, micro forms, reproductions of paintings, and other collections of materials; and

(f) degree to which innovations such as data processing and computerized information retrieval technologies have been applied to library operations.

The data relating to the above categories of information from a library or libraries is obtained from different sources: records (such as annual reports, government documents); visits to those libraries and using survey methods - observation (direct and indirect), checklists, questionnaires and schedules, interviews, etc.

Some examples of prominent research studies using survey method are summarised below:

(1) A study on Information Needs of Research Economists, by Marilyn White. The hypothesis was 'academic economists change their communication behaviour during the conduct of research projects'. The survey was conducted through a self-administered questionnaire to 294 economists at ten universities. The communication behaviour of the economists in academic settings was studied in three stages of their research process - problem, methodology, and presentation.

(2) Gail Schlachter's study on "the attitude of selected librarians towards unionism and professionalism" was based upon a mail survey among a stratified sample of 884 academic librarians. The study challenged the traditional assumption that librarians are self-sacrificing, unaggressive, and indisposed to engage in labor union activities".

(3) "Citizen's Attitude toward a Public Library" was a survey conducted by Charles Evans. It was a questionnaire-based mail (postal) survey to a sample of 367 library users and 153 non-users from the registered voters in Oceanside.

(4) Lucille Wert used a survey research method and other techniques to find out "the relationship between school librarian' formal education and the quality of high-school library service" in eight Illinois school libraries. Survey revealed that librarians with higher education attainment levels had developed more extensive reader services than by less-educated librarians.

(5) A study on "user failure at the library card catalog" was conducted by Carol Seymour and J.L. Schofield, choosing four libraries - Cambridge University Library, Leicester University Library, University Institute of Education Library and Bradford University Social Science Library. The study combined use of slips which were aimed at obtaining the nature of search and it was followed by interviews.

(6) In India survey method is extensively used in research studies leading to Masters and Doctoral degrees. An analysis of topics of doctoral degrees awarded by universities in India during the years 1980 to 1991, shows that out of 58 theses more than 75 per cent used survey method in their research. Non-doctoral studies also use survey method to obtain data regarding

status of libraries and their programmes. For example, the Curriculum Development Centre (CDC) of the University Grants Commission (UGC) used survey method (by sending questionnaires to the LIS departments) to study the status of LIS education in the country and to compile a Report. Another example is the one used by the INFLIBNET Centre to assess and report the status of computerization and networking of university libraries in India.

In the literature of librarianship, we find a number of works dealing with the use of survey method and provide numerous examples of library surveys. The following works help those interested in using library surveys:

McDiarmid, E.W. *The Library Survey*. Chicago: ALA, 1940.
Line, Maurice B. *Library Surveys: An introduction to their use, planning, procedure and presentation*. New York: Archon, 1967.

Tauber, Maurice F. "Survey method in approaching library problems". *Library Trends*. Vol.13; 1964.

6.6.2 Community Surveys

According to Busha and Harter (1980), "the community survey is conducted in an attempt to gather recorded and unrecorded data about the various social, political, and economic facets of the library's community so that more intelligent decisions can be made concerning the planning, development, and conduct of services".

Community surveys are closely related to library surveys. These are designed to make a careful and in-depth study of the characteristics of communities and relate these features to goals and objectives of the library, which they serve. The community surveys attempt to analyse all the factors such as local traditions and values, geographical conditions, the nature of the population, the political climate, economic development, human resources, and the problems of the community. The community surveys can be conducted in academic, professional, industrial, business, and other communities served by libraries and information centres. Such organisations can be grouped as

(a) educational institutions (schools, colleges and universities: general, vocational, technical, professional; informal or formal)

(b) recreational services and cultural clubs: music clubs, book clubs, theatre groups, health clubs, etc.

(c) health facilities and services: health centres, medical clinics, hospitals, etc.

(d) government departments and administrative units:

(e) citizen's groups: neighbourhood councils, women forums, etc.

(f) religious institutions (churches, temples, mosques):

(g) industrial, trade, business bureaus: unions and associations

(h) individual persons: community leaders, elected representatives, etc.

(i) communication - information agencies: newspapers, radio, cinema theatres, television, bookstores, information and public relations units, information centres, etc.

There are two types of community surveys: descriptive community survey and exploratory community survey. According to Busha and Harter, a *descriptive community survey* is designed to characterise properties and conditions of a group of people living or working together in a district or within an institution. An *exploratory survey* devotes attention to interrelationships between these characteristics and the use of various library resources, programmes, and services.

An analysis of the library's community helps us to identify the actual or potential users and their information needs, evaluate existing programmes and decide on their continuation, design more meaningful and effective programmes, develop appropriate library personnel, buildings, resources and facilities.

6.7 LET US SUM UP

Let us recapitulate what has been discussed so far in this unit.

* Social survey is a process by which quantitative facts are collected about a social phenomenon, problem or situation.
* There are two major approaches to surveys: cross-sectional and panel & trend studies.
* In survey research, data is collected through questionnaires and schedules, interview and observation methods.
* A schedule is a form that is filled up during the personal interview.
* Interview provides an opportunity to observe both respondents and the total situation.
* Observation methods are of three types: simple (uncontrolled), systematic (controlled) and mass observation.
* Survey method in LIS research involves library surveys and community surveys. Library surveys involve a systematic and in-depth examination of libraries, library systems or networks of libraries.
* Community surveys are closely related to library surveys. They are conducted in academic, professional, industrial, business and other communities in libraries and information centres.

6.8 REFERENCES AND RECOMMENDED BOOKS

BUSHA, Charles H and Stephen P Harter. *Research methods in librarianship: Techniques and interpretation*. New York: Academic Press, 1980. (Chapters 3 and 6)

GOODE, William J and Paul K Hatt. *Methods in social research*. New York: McGraw-Hill, 1982.

KRISHAN Kumar. *Research methods in library and information science*. New Delhi: Har-Anand, 1992.

YOUNG, Pauline V. *Scientific social surveys and research*. 4th ed. New Delhi: Prentice-Hall of India, 1984.

6.9 ASSIGNMENT

(1) Go to any university library and survey the dissertations submitted in social sciences in the past five years. List out the dissertations using survey method.

6.10 MODEL EXAMINATION QUESTIONS

I ESSAY QUESTIONS

1) Define 'Social Survey'. Discuss its characteristics.

2) List out different approaches to survey research.

3) Trace the importance of questionnaires and schedules in social surveys.

4) Describe the techniques of Interview method.

5) Enumerate the various types of Observation and discuss the merits and limitations of each one.

6) Explain the role of survey method in LIS research.

II SHORT NOTES

a) Cross-sectional Survey

b) Rules of Interview

c) Aids to Observation

d) Descriptive Survey

UNIT - 7 : RESEARCH METHODS - CASE STUDY AND EVALUATION RESEARCH

Structure

7.0 AIMS AND OBJECTIVES

This unit aims at introducing you to two more types of research methods, namely, case study, and evaluation research.

After studying this unit you should be able to

- explain case study as a research method and its application to library and information science
- describe the method of evaluation research in LIS field

7.1 INTRODUCTION

The case study method in research in social sciences is used for intensive study of a single unit. It might have the focus on one person or an institution or a village or a community or any such single unit. It is also used when two or three representative instances have been selected and subjected to an unusually detailed description. Case study method is a popular method in medicine, psychology, sociology and management science. As libraries are organisations, case study method is applied mostly in library management studies.

As libraries and information centres are established for the purpose of users, they have the responsibility to show their worth to the society. They have accountability for the finances they spend. Libraries launch and carry out several programmes to promote their use and also to improve their operations and services. Librarians used to justify their programmes through numerous statistical data. The earlier studies were based primarily upon value judgements of library authorities. The current trend is toward more systematic accountability of library activities. Evaluation research allows to measure the success or failure of a programme. According to Busha and Harter, "studies conducted to obtain objective and systematic evidence of the success or failure of library projects and programs are often categorized as evaluation research'

In this unit, you learn about two important research methods, namely, Case study and Evaluation research.

7.2 CASE STUDY METHOD

The term 'case study' has been defined by a number of researchers in different ways. P.V.Young defined case study as "a method of exploring and analyzing the life of a social unit be it a person, family, an institution, cultural group or even an entire community".

A case study is an empirical enquiry that investigates contemporary phenomenon within its real life context when the boundaries between phenomenon and context are not clearly evident and in which multiple sources of evidence are used. Case study research includes both single and multiple case studies. Certain journalistic efforts can qualify as case studies. Case study can include and even be limited to qualitative evidence.

Case study serves a dual purpose. While intensive investigation of a single case is often made for the sole purpose of increasing general knowledge of a given area, perhaps more often than not, it is carried out with the ultimate purpose of making a practical improvement in the specific instance examined and only incidentally it contributes to general knowledge. As a research method, it is examined with respect to the way in which it can help to solve scientific problems. Therefore, the dual purpose of case study is to solve some of the problems and secondly to help in bringing about improvements.

Case study does not lead to well-established conclusions but rather to what might be better described as empirically developed hypotheses. Goode and Hatt opined that case study is not a specific technique, it is a way of organising social data so as to preserve the unitary characteristics of social subjects being studied.

Case study method provides for latitude and freedom to the investigator with respect to the type and amount of data gathered, sources of information and procedure to gather information. Sometimes even a good case study seems not to be a systematic study of a scientific problem at all, but a general hodge-podge of information gathered by the investigator following capricious whims and inclinations. Case study is used more and more in a preventive way rather than as a therapy after the situation has broken down.

7.2.1 Characteristics of Case Study

The following are the characteristics of the case study method:

(a) It is a detailed and intensive study of a social unit

(b) It is a method of qualitative analysis

(c) All the traits and variables are studied and linked with one an other

(d) It adopts a holistic approach and preserves the wholeness of the units

(e) It is a continuous study and covers a sufficiently wide cycle of time

(f) It does not limit to single source of data collection. The sources of data include life histories, personal records, personal documents, personal letters, biographies, observation, interviewing individuals, etc.

(g) The data gathered from the study is scientifically synthesized and used as a diagnostic study to find out what is happening and what can be done about it.

7.2.2 Steps in Case Study

The specific steps in the case study method are:

i) *Determining the present status* - To gather descriptive information which will determine as precisely as possible, the present status of the unit under investigation.

ii) *Gathering background information* - To collect information about and examine the circumstances leading to the current status. An attempt is made to list out the possible causes/problems in the present situation.

iii) *Testing suggested hypotheses* - To gather specific evidence about each of the hypotheses suggested from the background information just gathered. Since behaviour is multidetermined, the investigator is not likely to come up with a single cause for the breakdown of the situation.

iv) *Instituting remedial action* - Subsequent to verification of the contribution of one or two factors to the existing situation, some remedial or corrective or improvement programme should follow to check what effect the change has brought about.

7.2.3 Case Study Protocol Format

The case study protocol should have the following format:

i) *Overview of the case study project*:
Project objectives and auspices,
Case study issues,
Relevant readings about the topic being investigated.

ii) *Field procedures*:
Credentials and access to the study sites,
General sources of information and
Procedural reminders.

iii) *Case study questions*:

It should include specific questions in collecting data, and the potential sources of information for answering each question.

There are six sources of evidence which can be the focus of data collection for case study: documentation, archival records, direct observation, participant observation, physical artifacts or cultural artefacts, and tool or instruments.

iv) *Guide for the Case study report:*

The outcome format - narrative/specific,

Specifications for any bibliographical information and

other documentations.

7.2.4 Collecting Data in Case Studies

The major tasks in collecting data are:

i) Gaining access to key organisation/area/interviewees.

ii) Having sufficient resources while in the field.

iii) Developing a procedure for calling for assistance and guidance, if needed from other case study investigators/ colleagues.

iv) Making a clear schedule of the data collection activities that are expected to be completed within a specified period of time.

v) Providing for unanticipated events including changes in the availability of interviewees as well as changes in the mood and motivation of the case study investigator.

7.2.5 Advantages and Disadvantages of Case Study Method

The *advantages* of the case study method are:

i) It is an essential technique in exploring new fields.

ii) Tremendous producer of ideas and suggestions.

iii) Helps to formulate sound hypotheses about behaviour.

The *disadvantages* are:

i) It is inefficient in situations which are already well structured and where the important variables are already identified.

ii) It often assumes that all the past experiences of the individual or past happenings in the situation have contributed to the final result.

iii) It depends upon the recall of others as to what had happened, thereby giving room for fallibility of memory.

iv) Usually case studies often involve a problem case resulting in undesirable traits being over-emphasised and desirable characteristics tending to be under-emphasised thus limiting the extent to which results can be generalised to more typical situation.

v) It is also likely that the case worker becomes frustrated with incompleteness of data he obtains and if more than one case is involved, with the variation of information from one situation to another.

7.2.6 Skills Needed for a Case Study Researcher

The following are the prescribed skills required for a case study researcher:

i) Should be able to ask good questions and to interpret the answers.

ii) Should be a good listener and not be trapped by his or her own ideologies or preoccupations and be able to observe and sense more generally.

iii) Should be adaptive and flexible so that new encountered situations can be seen as opportunities, not as threats.

iv) Should have a firm grasp of the ideas being studied whether it is a theoretical or policy orientation. Such a grasp reduces the relevant events and information to be sought to manageable proportions.

v) Should be unbiased by preconceived notions including those derived from theory. Thus a person should be sensitive and responsive to contradictory evidence.

Multiple case studies can be undertaken on the same subject at different sites or in the same site covering different people or same institution located at different sites. This would facilitate to identify the commonalities and differences among a cross-section of the target group or institution to draw broad conclusions on parameters covered in the study. This also provides for comparison across the cases. For example, if we prepare success or failure cases either of institutions or individuals exposed to some intervention, then the factors contributing to the success or failure may be found to be different or same across these cases.

Sometimes the study of a case can lead us to study more number of cases. In case different factors emerge from a case, then more cases are needed to be studied to identify certain crucial factors irrespective of success. These crucial factors could guide programme modification or change in strategies. However, it is difficult to indicate how many cases one has to study to identify the common factors or to emphasise the variations across the cases. Such multiple cases could provide rich information about the various possibilities contributing to/ facilitating/ debilitating a programme flow or impact. Case studies, therefore, can be a useful tool for comparative studies of specific events on specific persons or institutions.

7.2.7 Case Study Method in LIS Research

According to Busha and Harter (1980), the case study is particularly appropriate for gathering extensive data about a single research object, and this allows a concentrated focus on a single phenomenon and the utilization of a wide array of data gathering methods". Data gathering methods used in case studies are based primarily upon direct observation: both participant and non-participant observation. Sometimes, questionnaires and interview techniques are used to supplement the above techniques.

In librarianship, case study method may be employed to study the following aspects: (a) organisations such as libraries, media centres, information centres or library schools; (b) librarians or other library personnel; (c) groups of library users (students, children, women, professional workers (such as doctors, lawyers, engineers, etc.); (d) techniques or processes in a library; (e) library services; (f) information systems and networks; etc.

Case studies closely resemble library surveys, yet these are not surveys. A case study may be distinguished from its characteristics, such as (a) these are undertaken with reference to a single subject or single situation, (b) the group selected as subject is homogeneous, and (c) it is an intensive study and extensive data is collected for a comprehensive examination over a wide period of time. For example, case study method may be used to study about a university library, a single institution. An intensive study may be conducted to examine its growth and development, contribution to the academic community, use, etc., using a wide range of data gathering methods.

Generally, case study method has been viewed as an *exploratory study* about a single subject or a single situation. However, a collection of case studies of similar nature may be used to derive generalizations about an observed phenomenon in an area/ subject, including librarianship.

7.3 EVALUATION RESEARCH

7.3.1 Definition

Ronald Powell defines Evaluation research as "a specific type of applied research whose primary goal is not the discovery of knowledge but rather a testing of the application of knowledge within a specific program or project".

Evaluation research is a process of applying scientific procedures to accumulate reliable and valid evidence in the manner and extent to which specific activities produce particular effects or outcomes. A programme can be evaluated if it has clearly specified goals and in which the manner of programme implementation is clear.

Evaluation implies making reasonable judgments about programme outcome based on objective evidence gathered through accepted methods of research. Like many other kinds of research formulations, evaluation research is conducted to describe a phenomena or situation, to reveal relationships between variables, and to identify cause-and-effect relationships. For this purpose, many research techniques are used to evaluate a programme or a project. Therefore, evaluation studies are, usually, not distinguished by their method, but by their purpose.

7.3.2 Elements of Evaluation

Evaluation is the determination of the results attained by some activity designed to accomplish some valued goal or objective. The objectives are the desired results of an activity. There can be hierarchy of objectives for any endeavour.

A programme is an organised set of activities, projects, processes or services oriented towards achievement of specific objectives.

A project is a planned undertaking which is a set of interrelated and coordinated activities designed to achieve certain specific objectives within a given budget and period of time. The purpose of a programme or project is to convert a set of resources into desired results (objectives) through a set of activities or processes. Evaluation is the process of determining systematically and objectively the relevance, efficiency, effectiveness and impact of activities in the light of the objectives. The resources are called *inputs*. The results are divided into three broad categories: output, effects and impact.

Inputs are goods, funds, services, manpower, technology and other resources provided for an activity with the expectation of producing outputs and achieving the objectives of the programme.

Outputs are the specific products or services which an activity is expected to produce from its inputs in order to achieve its objectives, physical outcomes like irrigated area, number of institutions established, kilometers of roads constructed, health facilities provided, schools built and so on. Services provided in terms of number of people trained/served or benefits given and so on.

Effects are the outcome of the use of project outputs like improvement of agricultural yields, improved credit services, increased use of health facilities, higher attendance in schools and so on.

Impact is the outcome of project effects. It is an expression of results actually produced and estimated, the changes in the living conditions like increased income, improved nutritional status, increased literacy rates, wider participation, self-sustained development of beneficiary both at the individual or household level.

In the light of this, it is required of a researcher to establish the criteria of evaluation in terms of evaluation of **effort** (input), **performance** (output), **adequacy** (output vis-a-vis need), **efficiency** (output/input ratio) through cost-benefit or cost effectiveness analysis and **process** (outcome). The six categories of evaluation are defined both by the purpose of the evaluation effort and by the kinds of activities that tend to be stressed.

7.3.3 Basic Categories of Evaluation Activity

(1) Front-end Analysis (Preinstallation, context feasibility analysis): These types of evaluation take place prior to installation of a programme to provide guidance in planning and implementing the programme as well as deciding if the programme should be implemented.

(2) Evaluability Assessment: This type of evaluation work includes activities at assessing the feasibility of various evaluation approaches and methods. The scope of the evaluation, technical matters, design limitations and cost parameters are established through evaluability assessment prior to undertaking a more formal evaluation especially a causal evaluation of programme outcomes.

(3) Formative Evaluation (Concurrent, Developmental, Process): These evaluations are aimed at providing information for programme management, modification and management.

(4) Impact Evaluation (Summative, Terminal, Outcome Effectiveness): These evaluations are aimed at determining programme results, and effects, especially for the purpose of making major decisions about programme continuation, expansion, reduction and funding.

(5) Programme Monitoring: This is the least acknowledged but the most practiced category of evaluation. The kind of activities involved in these evaluations vary widely from periodic checks and compliance with policy to relatively straight forward tracking of services delivered and counting of beneficiaries.

(6) Evaluation of Evaluation (Secondary evaluation, Meta-evaluation, Evaluation audit): This category includes professional critiques of evaluation reports, reanalysis of data, and external reviews of internal evaluations.

These types of evaluations are by no means mutually exclusive. Over time a particular programme might be involved in activities from all six general categories. Some activities may go on simultaneously, eg. programme monitoring, formative evaluation and impact analysis. Moreover, within each category there are numerous types of specific evaluations based on evaluation criteria.

7.3.4 Types of Evaluation

There are numerous types of evaluation. In each of these types we have enumerated a question or set of questions to pose them while evaluating a programme or a project.

1) *Accreditation Evaluation*:

Does the programme meet minimum standards of accreditation or licensing?

2) *Appropriateness-evaluation*:

What services should the clients be receiving? To what extent are the current services appropriate to client needs?

3) *Awareness focus*:

Who knows about the programme? What do they know?

4) *Cost/benefit analysis*:

What is the relationship between programme costs and programme outcomes (benefits) expressed in money terms?

5) *Cost effectiveness evaluation*:

What is the relationship between programme costs and outcomes (where outcomes are not measured in money terms)?

6) *Criterion referenced evaluation*:

To what extent has a specific objective been attained at the desired level of attainment (the criterion)?

7) *Decision-focused evaluation*:

What information is needed to make a specific decision at a precise point of time?

8) *Descriptive evaluation*:

What happens in the programme (No "Why" questions or cause/effect analysis)?

9) *Effectiveness evaluation*:

To what extent is the programme effective in attaining its goals?

10) *Efficiency evaluation*:

Can inputs be reduced and still obtain the same level of output or can greater output be obtained with no increase in inputs?

11) *Effort evaluation*

What are the inputs into the programme in terms of number of personnel, staff/ client ratios, and other description of levels of activity and effort in the programme?

12) *Extensiveness evaluation*:

To what extent is this programme able to deal with the total problem? How does the present level of services compare to the needed level of services?

13) *External evaluation*:

The evaluation is conducted by people outside the programme in an effort to increase objectivity.

14) *Formative evaluation*:

How can the programme be improved?

15) *Goal attainment scaling evaluation*:

To what extent do individual beneficiaries attain individual goals on a standardized measurement scale of 1 (low attainment) to 5 (high attainment).

16) *Goals-based evaluation*:

To what extent have programme goals been attained ?

17) *Goal-free evaluation*:

What are the actual effects of the programme on beneficiaries (without regard to what staff say they want to accomplish) ?

18) *Impact evaluation*:

What are the direct and indirect programme effects on the larger community of which it is a part ?

19) *Internal evaluation*:

Programme staff conduct the evaluation.

20) *Longitudinal evaluation*:

What happens to the programme and to the participants over time ?

21) *Meta evaluation*:

Was the evaluation well done? Is it worth using ?

22) *Need assessment*:

What do beneficiaries need and how can those needs be met ?

23) *Norm referenced evaluation*:

How does this programme population compare to some specific norm or reference group on selected variables ?

24) *Outcomes evaluation*:

To what extent are the desired client outcomes being attained?
What are the effects of the programme on beneficiaries ?

25) *Performance evaluation*:

What actually are the participants able to do as a result of participation in the programme ?

26) *Personnel evaluation*:

How effective are staff in carrying out their assigned tasks and in accomplishing their individual goals?

27) *Process evaluation*:

What are the strengths and weaknesses of day-to-day operations?
How are programme processes perceived by staff, beneficiaries and others?
What are the basic programme processes? How can these processes be improved ?

28) *Product evaluation*:

What are the characteristics of specific and concrete products produced by or used in a programme ?

What are the costs, benefits, and effects of those products?

29) *Quality assurance*:

Are minimum and accepted standards of care being routinely and systematically provided to the beneficiaries?

How can the quality of care be monitored and demonstrated ?

30) *Social indicators*:

What routine social and economic data should be monitored to assess the impacts of the programme (eg. health, statistics, crime statistics, housing statistics, employment statistics)

31) *Summative evaluation*:

Should the programme be continued ? If so, at what level ?

32) *System analysis*:

What are the available alternatives and given those alternatives, what is the optimum way to do this programme ?

33) *Utilization focused evaluation*:

What information is needed and wanted by decision makers, information users, and others concerned that will actually be used for programme improvement and to make decision about the programme?

7.3.5 Techniques of an Evaluation

Evaluation research has no independent method of study of its own but could use survey techniques, personal interview or field experimentation with its attendant limitations. In undertaking evaluation, one has to set the criteria as indicated in terms of programme effort, output or outcome or any one of the specific sub-processes and accordingly undertake the evaluation. While evaluating, one aims at getting the actual changes as a result of programme intervention in terms of economic benefits or improvement in productivity or material gains or it may be in terms of changes in the attitudes and feelings towards the programme measured in terms of reactions.

Generally, evaluation at times takes the form of social audit of a programme resulting in 'fault-finding' exercise, identifying all the limitations in programme conception and management. There should be healthy collaboration between programme implementors and the evaluators as each one needs to understand and appreciate the other's role in this transaction. Therefore, it is advocated that evaluation should involve the programmer and should be proactive instead of reactive in its scope. Ultimately, our objective is to bring about improvements in programme planning and management.

7.3.6 Steps in Evaluation Research Process

Busha and Harter (1980) have identified the following six steps in the process of conducting evaluation research.

(1) Determine what is to be evaluated and the reasons for the evaluation.

(2) Establish the desired level of performance of the object of the evaluation with particular reference to existing standards, goals, and objectives.

(3) Select appropriate investigative techniques, and, if necessary, prepare a suitable scale or instrument to measure levels of performance.

(4) Measure or test subjects (objects) to determine actual level(s) of performance.

(5) Compare the programme's goals with its objectives (i.e., determine the extent to which goals were achieved).

(6) Evaluate the project on the basis of analyzed data.

7.3.7 Skills expected from an Evaluator

The evaluator should be thorough about his understanding of the programme, the context and evolve appropriate measures to assess the variables based on the evaluation criteria. The coverage in terms of area and sample should be representative. Only a cautious and careful approach on the part of the evaluator would succeed in getting the relevant data and prepare a report which is objective enough.

7.3.8 Evaluation Research in LIS

Evaluation research can be applied to any aspect of library - library resources (documentary and non-documentary), personnel, infrastructure, procedures and techniques, services, extension activities, etc. You have studied about the famous Aslib-Cranefield Research Project carried out by Cyril Cleverdon and his colleagues to assess the relative efficiency of several indexing languages can also be said as an example of evaluation research. The Library Association Cataloguing and Indexing Group's 'Catalogue Use Study' conducted at 39 academic and public libraries to assess the user friendliness of the catalogue is another example of an evaluative research.

Evaluation research has been used in several doctoral research studies. Busha and Harter reported the uses of evaluation research in - "an evaluation of information retrieval systems in terms of cost"; " a study of the effectiveness of public library's information services"; "an investigation of information services to adults of low educational attainment"; "an evaluation of the reactions of ethnic minority readers to indigenous ghetto literature"; "a study to determine the benefits and level of performance of the Faculty of Document Delivery Service"; "etc.

There are several publications dealing with evaluation research method, which can help librarians to plan and conduct evaluation studies in their libraries. The following are important among them:

LANCASTER, F.W. *TheMeasurement and Evaluation of LibraryServices*. Washington, D.C.: Information Resources Press, 1977.

LANCASTER, F.W. *Evaluation and Scientific Management of Libraries and Information Centres*. Bristol: NATO Advance Study Institute, 1975.

ALTMAN, Ellen et al. *Data Gathering and Instructional Manual for Performance Measures in Libraries*. Chicago: Celadon Press, 1976.

AMERICAN Library Association. *Performance Measures for Public Libraries*. Chicago: ALA, 1973.

In India, there are several examples of evaluation studies conducted on academic, public and special libraries. An evaluative study conducted by Carl White on the University of Delhi Library is a classical example. It was a comprehensive evaluation, using multiple techniques. White used documentary sources (ie university records), survey method, interview techniques, comparative analysis, etc in his research. He used internal records to assess the pattern of library's growth and compared it with that of other libraries in the country. He compiled a comprehensive bibliography of over 6500 desirable titles from eighteen subject fields and evaluated the strengths and weaknesses of library collection. Questionnaires were sent out and students and teachers were interviewed to measure user satisfaction of library services. The evaluation study helped White to provide a comprehensive list of several recommendations for the improvement of library services and to chalk out a perspective plan to develop the central library and the libraries of the constituent colleges to form a library system.

7.4 LET US SUM UP

Let us recapitulate what has been discussed so far in this unit.

* A case study is an empirical enquiry that investigates contemporary phenomenon within its real life context.

* Case study is a detailed and intensive study of a social unit, situation or a phenomenon and covers a sufficiently wide cycle of time.

* In LIS research case study method is extensively used to study libraries, techniques and processes, services and information systems and networks.

* Evaluation research is an applied research method used to obtain objectives and systematic evidence of the success or failure of library programmes and projects.

* The major elements of evaluation research are input(effort), output(performance), effect, impact and outcome(process).

* There are various evaluation techniques. Evaluation research has no independent method of its own, but could use survey techniques, personal interview, and field experimentation to study a programme.

* Evaluation research in LIS is used to study different types of libraries, library resources, library techniques and processes, information retrieval systems, etc.

7.5 REFERENCES AND RECOMMENDED BOOKS

BUSHA, Charles H and Stephen P Harter. *Research methods in librarianship: techniques and interpretation*. New York: Academic Press, 1980.

INDIRA Gandhi National Open University. *Research methodology* (MLIS-E2). New Delhi: IGNOU, 1994.

KRISHAN Kumar. *Research methods in library and information science*. New Delhi: Har-Anand Publications, 1992.

LANCASTER, F.W. *The measurement and evaluation of Library services*. Arlington: Information Resource Press, 1977.

POWELL, Ronald R. *Basic research methods for librarians*. New Jersey: Ablex Publishing, 1985.

UNIVERSITY of Madras. *Research methodology* (MLISc course package). Institute of Correspondence Education, 1989.

YOUNG, P.V. *Scientific social surveys and research*. New Delhi: Prentice-Hall, 1966.

7.8 ASSIGNMENT

1) Go to a university library and prepare a list of doctoral dissertations in social sciences using case study and evaluation research methods.

7.9 MODEL EXAMINATION QUESTIONS

I ESSAY QUESTIONS

1) Define case study. Discuss its characteristics and merits and limitations.

2) Enumerate the various steps in case study method and discuss how these steps are applied in LIS research.

3) What is Evaluation Research? Discuss the techniques of evaluation research.

4) Trace the various types of evaluation techniques and describe them briefly with examples.

5) Discuss the application of evaluation research in LIS with examples.

II SHORT NOTES

a) Case study vs. Survey method

b) Impact evaluation

c) Elements of evaluation

d) Skills expected from an Case Study Researcher

UNIT - 8 : CONTENT ANALYSIS, DOCUMENTARY RESEARCH AND OTHER METHODS

Structure

8.0 AIMS AND OBJECTIVES

This unit aims at introducing you to research methods, namely, content analysis, documentary research and other research methods

After studying this unit you should be able to

- explain what content analysis is ?
- discuss documentary method of research
- describe the Delphi method, participatory research, etc.

8.1 INTRODUCTION

The unit deals with some of the research methods used in library and information science. Content analysis, Documentary research, Delphi method, Comparative study, Participatory research, Sociometric techniques, Constructive typology, Projective techniques, etc have been introduced.

Content Analysis is a procedure designed to facilitate the objective analysis of the appearance of words, phrases, concepts, themes, characters, or even sentences and paragraphs in printed or audio-visual materials.

Documentary Research is a generic term used to refer to inquire into the printed tools of librarianship - books, journals and indexes. Bibliometrics and citation analysis use statistical methods, while automatic indexing and abstracting, and literature searching and information retrieval the methods which mostly depend on the computer processing and analysis.

Bibliometrics is concerned with the techniques of measurement and evaluation of communication media. Bradford's Law of Scattering and Zipf's Law deal with core journals and number of words in texts respectively.

Delphi method uses a group of modified survey procedure to refine the judgmental data collected from a group of anonymous respondents.

You will also learn about Participatory research, Comparative study, Sociometric technique, Constructive typology and Projective techniques are other techniques in this unit.

8.2 CONTENT ANALYSIS

Content Analysis is a technique for making inferences by systematically and objectively identifying specified characteristics of messages in written and oral communication. It aims at identifying the actual motive of the communicator. Literary and drama critics often evaluate media such as books, plays, and films in order to identify the motive of the author(s).

8.2.1 Definition and Purpose

According to Bernard Berelson, content analysis is "a research technique for the objective, systematic and quantitative description of the manifest content of communication". The categories of analysis used to classify the content are clearly and explicitly defined so that others can apply them to the same content to verify the conclusions.

According to Busha and Harter, content analysis is "a procedure designed to facilitate the objective analysis of the appearance of words, phrases, concepts, themes, characters, or even sentences and paragraphs contained in printed or audio-visual materials".

LIS studies generally focus on various print and non-print media and their use in libraries; however, they are less concerned with the actual content of media. Content analysis is a useful method adopted from the field of Communication Science by librarians to investigate certain internal (content) features of media. Analysis of the contents of materials can be used to examine documents so as to make meanings or messages more explicit.

Content analysis is a method of data analysis as well as a method of observation. Instead of observing people's behaviour directly or asking them about it, the researcher obtains copies of the communication that people have produced and asks questions about these records.

Content analysis studies help to

* identify the attitudes and motives of the communicators;
* measure the readability of communication materials;
* understand communication trends;
* trace the development of ideas, empirical facts or theories;
* compare content *vis-a-vis* objectives of communication;
* construct and apply standards in communication.

8.2.2 Steps in Content Analysis

The major steps in content analysis are as follows:

a) Formulation of the research problem

b) Definition and establishment of the hypothesis to be tested

c) Identification and selection of the materials to be analysed (the target documents, messages, or media) and, if appropriate, selection of a sample of these; categories for analysis are to be determined;

d) Analysis of relevant documents and the measurement of contents according to the predetermined and well-defined categories (documents are read and coded, the relevant content is condensed into special data sheets; items placed in each category may be sealed);

e) Quantification and ordering of the categorised units;

f) Analysis and comparison of attained data with respect to the independent variables(s); and

g) Interpretation of the results.

8.2.3 Sources and Techniques of Content Analysis

(1) Selection of Sources

Investigators can systematically analyse the data obtained from print as well as from non-print media. These may be archival records, documents, and the mass media. The sources of data may be letters, diaries, newspaper articles, minutes of meetings, live reportage, films, television and radio shows. The materials are also scrutinised to determine the relative importance or occurrence of textual or graphical elements.

Content analysis can be used to describe the contents of communication. The content may be a message or theme focused or implicit. Various kinds of materials need different descriptive tasks to analyse their content or messages. For example,

Source material	Descriptive tasks for analysis	Message/ Content
Novels	the themes used	love, morality, idealism,power outcast, career, etc.
Newspapers	kinds of news	local, national, or foreign and editorials, sports, social items, art, entertainment, business, etc.
Reference tools	treatment of professional groups	medical doctors, lawyers, professors,psychologists, etc.

Researcher may come across large quantities of materials and this may create some problems. In order to overcome such problems the following measures may be used:

1) Sampling procedures can sometimes be used to help solve the problem of having too many materials.

2) Another useful method for reducing materials to a more manageable quantity is to limit the analysis to works published during a specific time period - for example, 6 months, 1 year, or 3 years, depending upon the latitude of an inquiry.

3) Limitations can also be imposed by the selection of a particular media format for analysis (ie., periodical articles about a topic rather than books and other publications relating to the same subject area).

(2) Technique

In analysis of communication's content, observed data is transformed into sets of symbols such as words, themes, motives, concepts, and so on. These are marked and indicated by codes: 'yes' (if present) or 'no' (if absent). This is known as classifying or categorising data. The categories into which researchers code content vary with the nature of data and the research purpose. However, one should be thorough about the coding procedure, developing thematic ideas and establishing inter-coder reliability.

Content analysis involves the interaction of two processes: *Specification* of the characteristics of the content that researchers are to measure, and *application of the rules* researchers must use for identifying and recording the characteristics appearing in the texts to be analysed.

i) Units of Content Analysis

Units are the "things" placed into categories representing the contents of a document. They are the individual observations or measurements which ordinarily comprise the dependent variable of a content analysis. There are two kinds of units - Recording units and Context units.

A *recording unit* is the smallest body of content in which the appearance of a reference is noted (A reference is a single occurrence of the content element). A *context Unit* is the largest body of content that may be examined when characterising a recording unit. For example, the recording unit may be a single term, but in order to decide whether the term is treated favourably, the researcher has to consider the entire sentence in which the term appears (the context unit). Thus, the whole sentence is taken into account when the researcher records (and subsequently codes) the term.

There are five major recording units, generally, used in content analysis: words or terms, themes, characters, paragraphs, and items.

Word: The word is the smallest unit generally applied in research. When recording unit is a word, the analysis yields a list of frequencies of these words or terms.

Themes:. A theme is subject and a predicate expressed in a simple sentence. Themes are also found in clauses, paragraphs, and as illustrations in most texts. Investigators use theme as a recording unit, particularly in the study of propaganda, attitudes, images and values.

Character: The researchers examine the personality traits of individuals appearing in various texts and count the number of persons appearing in the text rather than the number of words or themes.

Paragraphs: The paragraph is infrequently used as a recording unit because of its complexity. Coders have difficulty in classifying and coding the numerous and varied elements covered in a single paragraph.

Items: The item is the whole unit the producer of a message employs. The item may be an entire book, an article, a speech, or the like. Analysis by the entire item is appropriate whenever the variations within the item are small and insignificant. For example, news stories can often be classified by subject matter such as crime, labour, or sports.

ii) Categories of Content Analysis

A *Category* has been defined as "either the name given to any class of things, actions, or relationships which recur with sufficient (relative) uniformity and frequency as to render the class a useful subject of a prediction of the class itself" (Gould and Kolb).

Category is a partition or sub-partition set up according to some rule and established in regard to the purpose of a unique research problem.

Categories must be exhaustive and mutually exclusive. *Exhaustiveness* ensures that every recording unit relevant to the study can be classified.*Mutual exclusivity* means that no recording unit can be included in more than one given category within the system.

The researcher has to explicitly specify the criteria that determine which recording units fall into each category. This enables replication, an essential requirement of objective and systematic content analysis.

iii) Quantification

Most content analysis research is quantitative in one form or another. Some quantitative procedure is used in order to provide a measure of the importance and emphasis in the material of the various ideas found and to permit comparison with other samples of material.

In order to perform quantification, researchers employ one of the following four systems of enumeration:

1) A *time-space system* based on measures of space (for example, column inches) or units of time (for example, minutes devoted to a news item on the radio to describe the relative emphases of different categories in the analysed material.

2) *An appearance system* in which coders search the material for appearances of certain attributes. The size of the context unit determines the frequency with which repeated recording units occurring in close proximity to each other are counted separately

3) A *frequency system* in which every occurrence of a given attribute is recorded.

4) An *intensity system*, generally employed in studies dealing with attitudes and values.

Methods of quantifying for *intensity* are based on the construction of scales. For example, using the paired-comparison technique developed by Thurstone, raters decide which of the possible pair of intensity indicators ranks higher on a scale of attitudes. The judgments are then used to construct categories into which recording units are placed and measured. The use of measurements will depend upon the purpose of the research and the study's expressed hypothesis.

iv) Testing of Hypothesis and Objectivity

Researchers guarantee objectivity by carrying out their analysis according to explicity set rules that enable different investigators to obtain the same results from the same messages or documents.

Content analysis is aimed at exactness and the elimination of bias in the investigative process. Some methods are employed to decrease the degree of subjectivity inherent in procedures designed to analyze or evaluate the contents of materials.

(a) By determining the proportions of a message that fall into clearly defined, conceptually valid, and mutually exclusive categories, the researcher attempts to in crease unit frequency validity.

(b) In a systematic content analysis, the "inclusion or exclusion" of content is done according to consistently applied criteria of selection;

(c) The analyst is not free to select and report what strikes him as interesting but must methodically classify all relevant material in his sample.

(d) It attempts to eliminate analyses in which only materials supporting the investigator's hypotheses..

(e) It insists on the researchers to have a thorough training in the procedures of coding, identification of content categories, relating them with each other and in drawing inferences objectively.

(f) The procedures used by investigators in a content analysis encompass a carefully prepared and reproducible plan for the collection and quantification of elements of communications.

v) Examples of Content Analysis

As discussed above, content analysis is done through identification of relevant themes/ ideas/focus, their coding and measuring the frequency of their occurrence in a given document. Some of the examples could be editorial emphasis in newspapers, political communication of important leaders, manifestos of political parties, topics dealt with newspapers or magazines, vocabulary analysis of commonly used words, difficult words or easily understood words and so on. All such documents are amenable for content analysis to highlight the important points, focus on any issue of topical relevance and the extent of coverage of these ideas.

The selection and analysis of information, situation, characters, or other elements in a communication can indicate pertinent features such as comprehensiveness of coverage or the intentions, biases, prejudices, and oversights of authors, publishers, creators, or other persons responsible for the content of materials.

8.2.4 Use and Application

There are three major applications of content analysis:

- to describe the attributes of the message;
- to make inferences about the sender of the message and about its causes or antecedents;
- researchers make inferences about the effects of messages on recipients.

Content analysis is most frequently applied in describing the attributes of the message. Harold Lasswell formulated the basic questions which can be raised to describe the attributes of a document. They are "Who says what, to whom, how, and with what effect ?".

A researcher may analyze messages to test hypotheses about characteristics of the text, what inspired the message, or the effects of communication.

Content analysis is also very useful research method for libraries. It is a very useful tool for policy analysis, book selection and communication analysis.

8.2.5 Content Analysis in LIS Research

Although most library science inquiries based upon content analysis have been descriptive in nature, this technique can also be engaged to probe hypotheses pertaining to the contents of materials. For example, assume that a test will be made of the theory that the nes of most contemporary best-selling novels are based primarily upon sex or violence. By clearly identifying activities that can be characterised as "sex" or "violence", and then by quantitatively analyzing the themes (content) of pertinent novels according to the established categories, the investigator could test the proposition. Futhermore, the hypothesis might be statistically tested and the study should be conducted rigorously. Hypothesis testing is especially facilitated when a comparison is sought in a study between or among elements of various documents.

Busha and Harter (1980) have listed the following types of content analysis projects that were completed by librarians in the USA.

1) A comparison of the contents of best-selling novels with those that did not sell well, centering upon themes, readability, settings, and other elements in an attempt to identify the characteristics present in popular books and absent in works having less appeal.

2) A comparison of selected novels with the motion pictures based upon them.

3) An analysis of treatments given to the topic of banks in the US by four major adult encyclopedia;

4) An exploration of the contents of novels listed in supplements to the Junior High School Library Catalogue in an attempt to characterise treatments given to the topic of adolescent culture

5) An analysis of information provided about educational filmstrips in selected periodicals.

6) A examination of the published writings of public library directors to determine the impact of the Great Depression on six metropolitan libraries.

8.3 DOCUMENTARY RESEARCH

Like Content analysis, Documentary research also deals with the documents and the words therein. It usually analyses and quantifies the content in print media.

8.3.1 Definition and Meaning

'Documentary research' is a generic term used to refer to inquiries into the printed tools of librarianship — books, journals, and indexes. Busha and Harter have enumerated a sample of broad research questions falling into this category, which will help us to understand the scope of this research method. The questions are as follows:

1) What are (or should be) the characteristics of a good index ? Classification scheme? What kinds of indexes produce retrieval results ?

2) How can a computer be programmed to do automatic indexing ? Automatic classification ? Abstracting ?

3) Will an analysis of citation patterns associated with journals in various subject literatures reveal anything useful about the importance of individual journals? Authors? or the "hardness" of a subject field?

4) Are there optimal approaches to the design of subject heading lists or thesauri ?

5) How does the depth of indexing affect retrieval effectiveness? The exhaustivity of indexing?

6) What effect does indexer/cataloguer inconsistency have on subsequent retrieval? How prevalent is inter-indexer inconsistency?

7) How quickly do various classes of literature obsolesce, in the sense of "losing usefulness"? How fast are various subject classes of literature growing?

8) How productive are writers? Journals? What proportion of journals ranked in order of usefulness, is required to satisfy a given percentage of potential information requests?

A number of similar related questions can be added to the above list. We can understand from these questions that documentary research basically deals with an objective and quantitative analysis of words in the texts of documents. There are a number of substantive and seminal studies conducted in the areas of bibliometrics, citation analysis and automatic classification/ indexing/abstracting, information storage and retrieval, etc. which could be categorised under documentary research method.

8.3.2 Bibliometrics

Bibliometrics is assuming a major branch of library and information science in recent times. It is concerned with the techniques of measurement and evaluation of communication media. It deals with the quantification of several interrelated aspects of writing and publication. It depends heavily on numbers (statistical data, figures, charts, bar diagrams, etc).

Hertzel defines Bibliometrics as "the analysis of the structure of literature using various tools, counting, rank-frequency distribution, and citation analysis; and although the structure of literature basic to all disciplines, it is particularly important in the area of information retrieval".

Bibliometrics involves in applying various methods of analysis, both qualitative and quantitative. It aims at identifying different characteristic features and behaviour of the documentary sources. It applies various statistical analyses to study patterns of authorship, publication, and literature. The features studied include

i) the bibliographical details (or variables);

ii) factors that are connected with document generation (i.e.,author's interests, publisher's affiliation, subject scatter, publication date, geographic concentration, etc.);

iii) behaviour of the documents
   - analyzing co-authorship pattern
   - understanding the role of institutions
   - role of institutions in research and publishing;

Bibliometric analysis is done by using techniques like, grouping, measurement, observation and mapping the literature. An in-depth analysis and use of these techniques lead several researcher to propound, Bibliometrics laws.

i) *Bradford's Law of Scattering*: Bradford discovered that a comparatively small number of "core" journals in a field yielded a high proportion of all the relevant articles. He also identified "zones" of less-productive journals, each zone providing reduced yield as increased number of marginally productive journals. His observations were generalised as Law of Scattering. This relates the proportion of relevant papers in a field to the proportion of journals required to this quantity, when the journals have been ranked according to productivity. Bradfo d's Law has been further examined by Vickery, Stevens, Leimkuhler, Brookes and many researchers.

ii) *Zipf's Law*: George Zipf's work was concerned with the use of words in natural-language text. He found that, when a large collection of text is analysed, a comparatively small number of words occur very frequently and count for a large proportion of all the word occurrences in the text. This was shown as a characteristic of hyperbolic distribution in a graph and called Zipfian distribution.

8.3.3 Citation Analysis

Thousands of studies have been conducted using the technique of citation analysis. These studies are based on an assumption that act of citing an author is meaningful. *Bibliographic coupling* is a technique that involves searching for groups of papers having common references. After identifying the most-cited papers in a field representing the research front, an analysis can be made of how these papers couple bibliographically.

Citation studies have been used to establish networks of scientific papers, to rank journals by importance, to generate additional documents relevant to a search question, and to evaluate a scientists's productivity. (Read Unit-14 for a detailed study on Bibliometrics and Citation Analysis)

8.3.4 Studies on Automatic Indexing, Abstracting and Classification

Numerous studies were conducted to use computers to prepare indexes and abstracts. In automatic indexing, keywords are selected from the texts of the documentary through a computer processing and they are used as descriptors in information storage and retrieval. Thesauri also compiled using automatic indexing systems. Likewise, key sentences are also extracted from the texts to serve as abstracts. This is called automatic abstracting. Recent studies deal with automatic classification systems, where documents are assigned automatically to subject classes.

In this type of automatic language processing experimental designs basically depend c 1 the decision theory techniques from operations research. Sophisticated statistical techniques such as factor analysis are also sometimes used. The analysis is usually carried out on words from the text through the methods such as count, tabulate, compare and associate.

8.3.5 Studies on Literature Searching and Information Retrieval

Librarians do extensive literature searches to answer reference queries, to provide information services or to compile bibliographies. The results of the literature searches are evaluated by various criteria - coverage, recall, precision, response time and effort. These criteria are equally applicable to any information retrieval system, whether manual, semi-mechanized, or mechanized systems. (We find a detailed study on the methods of evaluating ISAR systems in Course-03: Information Storage and Retrieval).

8.4 OTHER RESEARCH METHODS

8.4.1 Delphi Method

According to Busha and Harter, Delphi method comprises a group of modified survey procedures, designed for use in refining the judgmental data collected from a panel of selected experts". It is a systematic approach to generate consensus opinions among a group of carefully selected and anonymous respondents. The panel of experts are selected based on their knowledge of the subject relating to the investigation. These experts need not have similar backgrounds nor from the same discipline.

Delphi method is based upon a series of questionnaire interchanges often referred to as "rounds". Usually four questionnaire interactions (or rounds) are conducted to determine consensus opinions. In some cases three rounds are enough. The responses are ranked repeatedly by all participants in the study. In the statistical analysis of responses, the opinions that fall outside the interquartile range (i.e., those that are ranked within the first and fourth quartiles) are considered to be aberrant.

Delphi method can be applied in the areas like technological forecasting tasks, ascertaining values and preferences in futurology studies, incentive planning (ie efforts to encourage the development of technical inventions), etc.

Many researchers doubt this method for objectivity as there is a chance for the opinions being contaminated by strong personalities of the dominant respondents and bandwagon effect. This method also has another limitation that it is not amenable to precise analytical techniques.

In library and information science, Delphi method was used by Library Research Unit, University of California, Los Angeles. It was one of the most ambitious projects to study the uture of library education in US.

8.4.2 Participatory Research

It is an integrated activity that combines social investigation, educational work and action. The research process involves an approach to social investigation with the full and active participation of the community in the entire research process. It is used as a means of taking action for development and is an educational process of mobilising the groups of people for development. It emphasises research implications that enable oppressed groups to improve their lives and analyses problems in terms of community, social structure and cultural context of work.

The participatory research rejects the more traditional conceptions of social science research and seeks knowledge which has impact on social systems. On the contrary it is felt that academic methods which are vigorous in scope may undercut effective action and contribute to maintain the status quo. The participatory research involves the participation of respondents on the entire process and defines all participants as researchers. In other words, it seeks information from unwilling sources and hence may contribute to the generation of data which otherwise is not accessible.

Research on participatory efforts whose objective is to increase knowledge for the purpose of stimulating participatory process must involve the participants. It follows a **dialogical approach** where probing dialogues generally shall generate new questions and reveal issues in new perspectives. The distinction between the researcher and the researched is eliminated and importance is given to the participants' critical reflection. It is believed that when people are involved in the research, the action becomes part of their reflections.

The participatory research is solely in response to and for the fulfillment of needs of the less powerful, weaker segments of the social setting. Some are even of the opinion that participatory research break the legacy of objectivity as it goes beyond the boundaries of mere data gathering technique. It involves subjective commitment on the part of the researcher to the people under study which is essential. In other words, it rejects the value of neutrality of the researcher. There is a closer involvement of the researcher with the researched community. It encourages dialogue than manipulativeness and oversimplification of social reality; its approach is problem-centred to resolve the problems by transforming the conditions (social), political (socio-psychological). It is conceived as an educational process for both the researcher as well as the people with whom the research is conducted. It respects peoples' own capability and potential to produce knowledge and analyse it.

Some of the limitations of participatory research are that it lacks objectivity and the results may not be generalisable. However, it gets information which can be considered as qualitatively rich reflecting the felt needs of the people and may help to resolve experienced problems and to formulate development strategies. Since the focus of this method is to encourage people to articulate on general issues and problems, the type of data it will yield is different from what we obtain through other social science research methods.

In the context of library, this method is very significant in understanding the needs of users of library. Many user studies that are routinely conducted come under the category of participatory research. They truly represent the view of the readers in respect of collection of books/journals for the library but also the kind of library services that may be designed and produced.

8.4.3 Comparative Librarianship

Comparative studies are common in many disciplines of social sciences and humanities - economics, law, education, linguistics and literature. Comparative studies in librarianship encompass comparisons of the practices and study of library science in nations and regions of the world. Dorothy Collings defined Comparative librarianship as "the systematic analysis of library development, practices, or problems as they occur under different circumstances (most usually in different countries) — considered in the context of the relevant historical, geographic, political, economic, social, cultural and other determinant background factors found in the situations under study".

The topics generally included in comparative studies are library resources and services, publishing vis-a-vis libraries, book trade and libraries, cross-cultural influences on librarianship, patterns of communication, political control on libraries, and other related topics.

The research methods used in comparative studies are direct observation, historical records, government documents (reports and policy decisions), publications from professional bodies, statistical compilations, compilations of case studies, etc. Sometimes, survey and case study methods are also employed to collect data. Data is from these sources analysed to make qualitative and quantitative comparisons.

The comparative studies in librarianship are often listed under the headings: International librarianship, Librarianship in Developing Countries, Librarianship in Latin American Countries, etc.

Some of the most useful publications in the field of comparative librarianship which are useful for the researchers are:

Harvey, John F. *Comparative and International Library Science.*

Danton, J. Periam. *Dimensions of Comparative Librarianship.*

Simsova and MacKee. *A Handbook of Comparative Librarianship.*

8.4.4 Sociometric Technique

Sociometry deals with social relations among groups and individuals in a group. It is concerned with the measurement of social distance and also the institutional behaviour. The social distance is discovered and measured through attraction and repulsion between individuals and groups. The study of interaction among individuals and groups helps us to understand the social configuration and social phenomena.

The basic instruments used in sociometry in data collection are questionnaires and interviews. Usually, the questionnaires contain open-ended questions to elicit proper view of social interaction among groups.

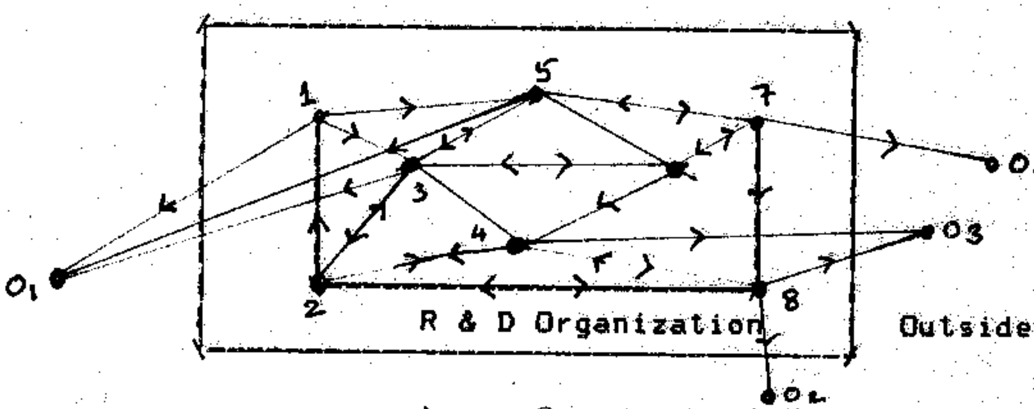
In libraries and information centres, sociometric techniques could be used in library management and reference services. The workers in a library or a section may be asked to give options about the choice of his/her colleagues "with whom he/she is interested to work". Based upon the rank of choices elicited, the staffing for a section or a project/assignment of a library could be fixed and thus conducive environment and productivity could be raised. In referral service, sociometric technique is used to identify the experts from different fields and their communication patterns.

There is a set of other related techniques associated with sociometry, namely, Sociogram and Sociomatrix.

Sociogram

A Sociogram is a geometric figure which indicates each person and his choice or communication pattern through the connecting lines and arrows. For example, the communication pattern among the scientists for discussing about an experiment can be represented through a sociogram. The scientists may be asked to list at least four persons with whom they usually contact for discussing about the current research in their subject field. The choice of scientist from outside organisation and the communication pattern could be shown in a sociogram. (Figure-1)

In the figure the scientists of the organisation are shown as 1,2,3.....8 and the outside scientists as O1,O2,O3 and O4. The scientists No.3 and 4 are much informed on the current trends in their subject field. Similarly the outside scientist O2 has a strong relationship with the scientists of the organisation. The results of the study could be used to identify/locate outside experts for consultation.



----->----- One way communication
---<-->--- Two-way communication

Figure-1 : A Sociogram showing the Pattern of Communication among Scientists

Sociomatrix (Sociometric Data Matrix)

We have learnt that sociometric studies help to identify relations between individuals and groups. Sociomatrix is another related technique, which is used to make sociometric data more amenable to statistical manipulation. For example, a matrix may be developed to show the reader's preferences of author(s) in a given subject. Statistical manipulation may help such data useful to book selection in libraries.

Readers	Preferences to the Authors					
	A B	C	D	E	F	
a	-	II	I	III	-	-
b	I	II	-	III	-	-
c	I	III	-	-	II	-
d	-	II	I	-	III	-
e	-	III	II	I	-	-
f	I	-	-	III	II	-

Table-1: Matrix showing Reader's Choice to a given Author

In this example, six Readers (a,b,c,d,e and f) are asked to give their choice to six authors (A,B,C,D,E and F) and mark them in the preferential order of I,II,III. These are shown in Table-1

Reader's Preference	Authors					
	A	B	C	D	E	F
FIRST	3	-	2	1	-	-
SECOND	-	3	1	-	2	-
THIRD	-	2	-	3	1	-
Total	3	5	3	4	3	0

Table-2: Statistical Manipulation of the Reader's Choice

In Table-2, the preferences scored for individual authors are given. The highest number of preferences (i.e. 5) is scored by B, whereas author-A has received First preference by 3 readers compared to authors C and D who have received two and one First preferences respectively. The author who gets maximum preferences is called "Choice Star" and the one who gets second choice is called "Aristotle Leader". The author who is not chosen is called "Isolate".

Depending upon the number of relations, the choice can be called as 'Mutual' (if there are only two), 'Triangular" (if there are three) and 'Quadrangular' (if there are four links).

8.4.5 Constructive Typology

Typology is analogous to classification. It is a common method of classifying subjects. For example, we classify/categorise libraries into academic, public, scientific and technical libraries depending on certain common characteristics. These 'types' (of libraries) are distinguishable from one another. Oganising the sets of categories into several types is called 'Typology'.

In the terminology of Constructive Typology, a type is a selective and purposive simplification constructed in terms of certain criteria. Thus, the constructive typology is a purposive and planned selection (involving abstraction, combination and sometimes even accentuation) of a set of criteria with empirical observations that serve as a basis for comparison. In most cases it is a mental exercise, but rarely it is an intuitive process. Scientific methodology is used in type construction involving logic, normative rules and procedures. Systems theory, historical research, case study, experimental procedure and quantitative procedure are used in constructive typology.

Constructive typology has been used in many disciplines of social sciences, including Library and Information Science. Historians use chronological and localised types. They periodise (ancient, medieval and modern, etc.) and localize (Indian, European, etc.) the history in their typology construction. On the other hand for the sociologists, the types are undated and non-localised. Social conduct, social organisation, personality, etc are types constructed in sociology.

8.4.6 Projective Techniques

Projective techniques are data gathering techniques used in situations, where the respondents are unable to give responses. These are indirect methods used as an approach to obtain information from the respondents. These techniques are originally used by physicians and psychologists.

Margaret Slater considers projective technique as an aspect of qualitative research which amounts to inviting respondents to play games. There are a number of information gathering techniques used in projective techniques. They are i) role playing; ii) drawing pictures; iii) story telling; iv) arranging show cards; v) imaging what one might do in hypothetical situation; and vi) complete sentences.

Projective techniques are relatively unstructured techniques and hence there is a possibility of getting vague and ambiguous responses from the individuals. However, psychologists trained in projective techniques can easily find out the true nature of the individual.

Some of the important projective techniques are listed below:

(1) *Inkblot Tests*: These tests require the person to impose meaning upon an object shown and interpret it. E.g: Rorschan Test.

(2) *Story Telling Tests*: In these tests a person is asked to provide an imaginative and dramatic account of a picture shown to him. Eg: Thematic Appreciation Test.

(3) *Graphic Techniques*: In these techniques the task is to produce a drawing of a person or house with minimum specification. Eg: Tonkin-Horn Picture Arrangement Test.

(4) *Completion Tests*: These tests are exemplified by sentence completion, word association, etc.

(5) *Social Attitude Tests*: In this scales are constructed on the basis of individual's responses of approval or disapproval on certain careful formulated statements/ propositions about some given issues/ institutions/ groups of people.

Often these indirect tests have been come under severe criticism by many that these do not provide evidence getting more valid data. Rarely these tests have been successful due to lack of expertise in administering, scoring and interpreting data. Projective techniques need thorough understanding.

In library and information science research, inkblot tests and graphic techniques can be used in library schools for selecting candidates for training. Story telling tests are useful in recruiting children's librarians.

8.4.7 Statistical Study

Statistical study deals with numbers, usually by counting or measuring the properties of a phenomena. Statistical methods does not study any one individual but a group of individuals and deal with quantitative rather than qualitative information. The methods facilitate the condensation and analysis of a large body of quantitative data.

Statistical methods can be applied at any stage of research process. You will find the application of statistics largely in data analysis and interpretation (units 10 and 11), however, it is also dealt to a smaller extent in sampling and data collection (unit-9), and presentation (unit-12). In this Section let us examine the use of statistics in measurement and scaling.

Data is collected by a method of counting and such data of finite numbers is called 'discrete data'. Example: Number of books issued in a day, week, month or year; number of male and female borrowers of a library; etc. Scaling is another type of measuring the data. There are four types of scales. They are nominal, ordinal, interval and ratio scales.

Nominal scale: The nominal scale is the lowest level of measurement. It is qualitative in nature and is not quantitative. It is commonly used to determine quality and is of 'yes/no' type; objectives classified into two or more mutually exclusive categories and each category is represented by a different symbol. The nominal scale of measurement is used to classify - gender (1 male and 2 female), place of residence (1 rural and 2 urban), marital status (1 unmarried, 2 married, 3 separated, 4 divorced, etc.), nationality, ethnicity, religion, etc. In LIS research we categorise users (users and non-users), grades of librarians, etc.

Ordinal scale: This scale is used to measure the variables that are not only classifiable but also exhibit some kind of relation. It possesses an inbuilt characteristic of order and enables us to determin greater than, less than, etc types of comparison with reference to quality measurement. Researchers assign rank values to the objects according to order (lowest extreme to highest extreme or vice versa). Example: 1 Strongly agree; 2 Agree; 3 Disagre; 4 Strongly disagree).

Interval scale: Interval scale of measurement also provides a ranking of positions as that of ordinal scale, but the distance between the limits of an interval is equal. One observation can be greater (or smaller) than another. It has also a zero point. Examples of variables measured according to interval scale are income, intelligent quotient (IQ), etc.

Ratio scale: Ratio scales are like that of interval scales having an absolute (natural) zero, below which no value occurs. The ratio scale has also the property of equivalence, relations in which one variable can be greater than another, and a fixed interval. Researchers measure the variables such as weight, time, length and area have natural zero points and are measured with the ratio scale. Eg. collection development in libraries in different time periods.

8.5 LET US SUM UP

* Content Analysis is a procedure designed to facilitate the objective analysis of the appearance of words, phrases, concepts, themes, characters, or even sentences and paragraphs in printed or audio-visual materials.

* Documentary Research is a generic term used to refer to inquire into the printed tools of librarianship - books, journals and indexes.

* Bibliometrics is concerned with the techniques of measurement and evaluation of communication media. Bradford's Law of Scattering and Zipf's Law deal with core journals and the number of words in the text respectively.

* Delphi method uses a group of modified survey procedure to refine the judgmental data collected from a group of anonymous respondents.

* Participatory research, comparative study, sociometric technique, constructive topology and projective techniques are other techniques dealt in this unit.

8.6 REFERENCES AND RECOMMENDED BOOKS

BUSHA, Charles H and Stephen P Harter. *Research methods in librarianship: Techniques and Interpretation*. New York: Academic Press, 1980. pp.171-179.

FRANKFORT-NACHMIAS, Chava and David Nachmias. *Research methods in the social sciences*. 5th ed. London: Arnold, 1996. pp.324-331.

INDIRA Gandhi National Open University. *Research Methodology* (MLIS-E2). New Delhi: IGNOU, 1994.

KRISHAN Kumar. *Research methods in library and information science*. New Delhi: Har-Anand Publications, 1992.

8.7 MODEL EXAMINATION QUESTIONS

I ESSAY QUESTIONS

1) Discuss the major methodological issues in content analysis and how you would go about resolving them.

2) Explain the importance Bibliometrics and Citation Studies in library and information science research.

3) Explain the use of Sociometric techniques in Social sciences, with special reference to LIS.

II SHORT NOTES

a) International librarianship

b) Units and Categories in Content Analysis

c) Delphi method

d) Participatory Research

BLOCK- III : DATA COLLECTION, ANALYSIS AND INTERPRETATION

Statistical methods have many purposes in research process. The statistical methods and procedures are used for summarizing, simplifying, reducing and presenting the results of the research. Inferential statistics help us beyond a mere description of a set of data, they are also useful to make predictions and also to test hypotheses. Traditional methods, which might have taken months of painstaking hand calculation, can now be tackled effectively and instantaneously with the help of computers. Researchers without making any deeper study into the theory of the statistics can now use them with the help of more sophisticated software packages.

The block aims to provide a description of the important methods of data collection, analysis and interpretation used in social sciences. The specific objectives of the block are

Unit-9 provides an over view of data collection methods using questionnaires, interview schedule and observation. It also covers various sampling methods, including the techniques of probability and non-probability sampling.

Unit-10 introduces various statistical methods used in data analysis. It includes statistical measures for one variable and two variables in data analysis, time series and testing of hypothesis.

Unit-11 deals with using computers in data analysis. It discusses the use of computer software and statistical packages used in processing of research data. An extensive description of Statistical Package for Social Sciences (SPSS) is also provided.

Unit-12 explains the various methods of presentation of research results and their interpretation.

UNIT - 9 : DATA COLLECTION METHODS AND SAMPLING TECHNIQUES

Structure

9.0 AIMS AND OBJECTIVES

The aim of the unit is introduce you to data collection methods and sampling methods associated with it. Sampling is an important aspect in data collection that every researcher should have clear about it.

After going through this unit you will be in a position

- explain the terms like 'population', 'sample', and 'sampling frame'
- identify and prepare the appropriate tool(s) for collection of data
- appreciate the need for and utility of sampling and able to work out sample size for a research study
- list out the salient features of some of the important sampling methods
- draw a sample from the given universe following suitable sampling design
- prepare a questionnaire/schedule for collecting data
- conduct sample surveys in an objective and efficient way.

9.1 INTRODUCTION

In social science research, the scholar has to identify the problem and formulate hypotheses as a part of the research design. Researchers need to collect data in order to test hypotheses and to provide empirical support for explanations and predictions. For this investigators are required to construct their measuring instruments in order to collect sufficient data pertinent to the research problem. They use any of the basic instruments of data collection such as questionnaire, interview and observation, depending upon the problem chosen. Once these basic steps are completed the data needs have to be spelt out. The relevant data are to be collected from the population either by complete enumeration or by sampling.

Collecting data from the complete population is called *Census method*. It is usually time and resource consuming process and problem. They use any of the basic instruments of data collection such as questionnaire, interview and observation, depending upon the problem chosen. Once these basic steps are completed the data needs have to be spelt out. The relevant data are to be collected from the population eith collected from the population eiths use a relatively small number of cases (a sample) as the basis for making inferences about all the cases (a population). It is a widely used method and sampling theory guides us to select a representative sample from the universe and sample survey research can yield best results at minimum cost. The relevant tools can be employed to elicit data from the units of the sample.

9.2 DATA COLLECTION METHODS

To seek answers to the questions/issues considered for investigation and also to test the hypotheses, data are to be collected from the sampling units (or respondents if the population under consideration is human population) on the items/ aspects already identified by the investigator. Most commonly used tools are : questionnaire, interview schedule and observation.

9.2.1 Questionnaire

Questionnaire is most popularly and widely used tool in survey research. In a questionnaire, the data will be filled-in by the respondents himself/herself. To elicit information from the respondents care has to be taken in the design of the questions so that the respondent feels comfortable in understanding and interpreting the questions included in the questionnaire. This method has been found to be relatively cheaper and has been widely used in the developed societies where (sample) the respondents are knowledgeable enough to go through the set of questions and can provide answers without external assistance. It is also presumed that the respondents are willing to cooperate with the research activity and provide reliable, valid and accurate data.

It is evident that the questionnaire is nothing but a layout or a set of questions arranged in a systematic and sequential way. Since drafting of a questionnaire requires special skills, one has to be familiar with the steps involved in the construction of a questionnaire. The following aspects have to be taken cognisance of in the development of a questionnaire.

(1) *Ordering the theme*: The researcher has to identify the various themes/issues on which he/she would like to collect information from the respondents. Ordering the themes is a very important task and separate blocks can be earmarked for each theme. For each aspect/issue, the appropriate questions have to be identified and formulated in such a way that based on the responses (to these questions) the above issue or aspect can be examined/answered. Normally, the sequencing of questions is from general to specific.

(2) *Simple, direct and unambiguous questions*: It is suggested that the questions should be drafted in such a way that they are simple, direct and unambiguous. They should clearly indicate the purpose of enquiry. Leading questions (with suggestive answers) and personal matters should be avoided. Questions should be arranged in logical order. It is imperative that the questionnaire should be brief so that the cooperation of the respondent is ensured.

(3) *Types of questions*: When several types of questions to be asked, such as factual, opinion and attitude, information, self-perception and standards of action factual questions comes first.

i) *Factual questions* normally pertain to respondent's age, education, experience, or any other pertinent personal data needed in the study is included. These factual questions help us to examine the relationship between the respondent's characteristics (normally, the independent variables) and the object of the study (the dependent variable).

ii) *Opinion and attitude questions*: Opinion about the topic or attitude can be ascertained from the respondents. See the following example:

Librarians should critically examine the policy of acquiring the detective novels in public libraries.

a) strongly agree b) agree c) undecided d) disagree e) strongly disagree

iii) *Information questions*: It is an attempt to determine how respondents know about a given topic and how or when they gained certain knowledge. For example,

What do you like or dislike most about your library's information service ?
When and how did you know about the OPAC ?
Who is the president of the ILA ?

iv) *Self-perception questions*: These allow individual respondents to compare their ideas or actions with those of other persons.

If you were asked to evaluate your success as a librarian in motivating your staff, how would you rate your performance?

a) very successful b) successful c) unsuccessful d) very unsuccessful e) uncertain

v) *Standards of action questions*: These questions are helpful when researchers attempt to determine how respondents will act in certain circumstances.

The university would like to know your feelings about the library services provided at your study centre.

What is your personal evaluation?

a) excellent b) good c) adequate d) barely adequaate e) bad f) very bad

(4) *Unstructured and Structured questions*: The type of questions to be asked can be structured or unstructured or open ended or closed ended.

Unstructured questions allow respondents to reply freely without having to select one of several provided responses. These questions could be described as open-ended questions. The responses to these questions tend to produce wide range of answers, which are very difficult to analyse. Example:

What, do you think, are the main causes of present unemployment in the library profession ? What suggéstions do you offer to remove it ?

Structured questions are characterised by a group of already provided fixed alternative responses. The analysis and tabulation of answers to these close-ended questions are very easy.

In the case of closed ended questions, all the possible responses have to be *apriori* determined. The responses may be dichotomous (yes/no type), trichotomous (yes/no or do-not-know or no-opinion type) or polytomous. For example,

How much do you feel that you really need computerising your library ?

(a) very much (b) some (c) not much (d) not at all (e) don't know

By structuring the questionnaire the task of filling-up of data may become easier to the respondent. However, while seeking opinion oriented questions it may still be advisable to make it an open ended questions. Otherwise, provision should be made for unknown responses. For psychological questions a three-point or a five point scale can be administered so that the responses can be coded and analysed easily.

(5) *Scaling Fixed Responses*: The structured pattern questions have multiple-choice items and the respondents chose the most appropriate item. The researcher carefully assigns numerical values to these precoded replies. Many questions with fixed responses are similar to items in Likert-type scales. Each item is assigned a numerical ranking based on a continuum that contains predetermined units of measurement. These scales are designed to show a differentiation among respondents who have a variety of opinions about an attitude object. For example:

i) Suppose the staff of your library go on strike for a month, how much would you miss the library?

   a) a very great deal (weight = 4)
   b) quite a lot (weight = 3)
   c) not very much (weight = 2)
   d) not at all (weight = 1)

ii) On the whole, how would you rate the quality of your library's services ?

   a) very poor service (weight = 1)
   b) poor service (weight = 2)
   c) adequate service (weight = 3)
   d) good service (weight = 4)
   e) very good service (weight = 5)

(6) *Semantic Differential Scale*: In this type of structured questions pairs of antonyms and synonyms, together with seven-step rating scales are provided. The attitude from the respondents help us to check the position on each continuum between the most positive and the most negative terms. This type of scale is often called a 'semantic differential scale'(Busha and Harter). For example, a questionnaire to evaluate the performance of a chief librarian, we may design a question and a semantic differential scale as follows:

How would you rate your performance as the chief librarian in your library during the past year?

Evaluation of performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Bad	-	-	-	-	-	-	-	Good
Disorganised	-	-	-	-	-	-	-	Organised
Unfair	-	-	-	-	-	-	-	Fair
Uncertain	-	-	-	-	-	-	-	Certain
Erratic	-	-	-	-	-	-	-	Systematic
Preoccupied	-	-	-	-	-	-	-	Observant
Harsh	-	-	-	-	-	-	-	Gentle
Outspoken	-	-	-	-	-	-	-	Reserved
Withdrawn	-	-	-	-	-	-	-	Outgoing
Submissive	-	-	-	-	-	-	-	Dominant
Totals	0	0	6	12	10	18	0	

The respondent's overall rating of the performance of a single instrument is calculated as sum the total values of each column and divide by the number of continuum rows (6+12+10+18 = 46/10 = 4.6).

(7) *Pre-testing of questionnaire*: The questionnaire should be pre-tested for deletions, modifications and inclusion of new ones. The researcher should be convinced of inclusion of each and every question in the schedule so that the questionnaire contains an irreducible minimum number of questions.

Advantages and Limitations of Questionnaires:

Advantages: The following are the advantages of questionnaires:

Questionnaires can be administered when the population is widely scattered. It saves time and cost. The environment of the respondent is not conditioned and the responses to the questions can be viewed as reliable. A request letter seeking the respondents' co-operation is necessary. And informing the respondents about the need for and utility of the research study may motivate the respondent to cooperate with the study under consideration. It is beneficial to indicate her/him that her/his contribution to the study will be highly valued since it may ensure a better response from the sample respondents.

Limitations: The following are the limitations of questionnaires:

i) It is obvious that the quality of response depends upon to what extent the researcher could motivate the clientele group and also on the general attitudes of the respondent class.

ii) There is no objective way of verifying the reliability of the data. If scales aiming at a higher degree of precision are used then respondents should be adequately explained about the salient features of such scales, otherwise loss of precision will be severe. Field experiences indicate that the rate of response to mailed questionnaire will be poor perhaps due to change of address or lack of interest of respondents in the particular research study. So, a large number of questionnaires will go waste.

9.2.2 Interview Method

It is a more effective method of data collection since the researcher comes in direct contact with the respondents. He has an opportunity to build personal rapport with his client and make him appreciate the utility of the study. He can use either a schedule of questions or the modern electronic gadgets to record the information. The probing and also elaborating a point of view is possible under this mode. In social science research, the investigator generally uses a schedule of questions and approaches the sample respondents for recording of information. It has been suggested that the investigator should be polite, courteous and should not be emotional at the time of conducting the interviews. Some of the limitations of the questionnaire method mentioned above can be avoided in this particular approach.

9.2.3 Observation

You may notice that in the administration of questionnaire or conducting of interview, the behaviour of the respondent is partly conditioned which may affect the quality of the data. When we would like to observe the dynamic behaviour of the sampling unit i.e. respondent in its natural settings and record the response to certain stimulus, the above methods will be of little use. In such situations, observation becomes necessary for obtaining accurate information. Generally, the choice of observation technique is made on the basis of the nature of the problem, skills of the investigator and the characteristics of the observed. 'Observation' method fails if the observed are intelligent enough to prevent observation of their behaviour by the outsider / investigator. To use observation method as a tool for data collection, it requires certain degree of skills on the part of investigators, which partly depends upon the social background of the observer.

Purpose:

The three main purposes of observation are :

i) to capture and document the actual human behaviour;

ii) to provide graphic description of the social reality and to trace out the factors/ events affecting social relations; and

iii) to explore and identify themes/areas of research or alternative ways of conducting research.

Types:

Based on the actual role played by the investigator, two types of observation methods are identified: (1) participant observation, and (2) non-participant observation.

(1) Participant Observation

The researcher / investigator assumes the role of a participant in a social situation without herself/himself emotionally being involved in the process. It implies that the society should accept her/him as a part of it, and then only the investigator will be in a position to observe the interaction of various social groups in their natural surroundings. By observing and by talking to people he can redefine the social relations or discover new interpretations of the events.

Pre-conditions:

i) Investigator should be acceptable to the observed community

ii) Investigator should not become a part of the social situation and should not impose her/his value system. She/he should be neutral; and

iii) Investigator should always try to seek information but not resort to questioning as it develops unequal relationship between the investigator and the observed. He should possess the skills or capabilities to sense changes and under stand events.

Advantages:

i) Being a member of the community, she or he can observe the group dynamics and social behaviour more closely and thus avoid these biases that creep into in the conduct of interviewing; and

ii) Scope for collecting a lot of data is very high; indeed it may help identifying new hypotheses and get new insights.

(2) **Non-participant observation:** A method of investigation in which the researcher observes the social behaviour of people in their natural settings without becoming an actual participant in the system. In other words, the identity of the researcher is maintained and the implicit assumption is that the presence of an outsider does not significantly influence the behaviour to such an extent that the information accuracy suffers. When compared to the participant-observation method, it has certain advantages in terms of selecting the situations to ensure representativeness.

9.3 SAMPLING TECHNIQUES

After constructing the basic instruments of data collection the researchers use them on the subjects. The subjects may be respondents, events, cases or units defined by the research problem. It is often impossible, impractical, or extremely expensive to collect data from all the potential units of analysis covered by the research problem.

9.3.1 Sample and Sampling

Researchers select a relatively small number of units (a subset) when they accurately represent all the relevant qualities / attributes of the whole set. The whole group from which the sample has been drawn is technically known as *universe* or *population* and the group (subset) actually selected for study is known *sample*.

Universe: Though the terms universe and population are used synonymously, there exists a small difference them. Conceptually speaking the universe represents the whole group from which the units or samples are selected. There is no operational definition for the term universe, the definition of population is equally applicable to it. The universe of a research project is often understood by its area and scope. For a research problem like 'Impact of information services on Research' the scope of research activity is not necessarily connected to an institution. However, we may select R&D institutions in India from which a sample is drawn.

Population: The concept of population is fundamental to survey research method. A population is the "aggregate of all cases that conform to some designated set of specifications". For example, the specifications are "people" and "residing in India" (ie., a population consisting of all the people who reside in India).

Generally, the population of a research problem has to be defined in terms of (1) content, (2) extent, and (3) time. For example, "College librarians", "Andhra Pradesh" and "as of April 1, 1998" (ie., a population consisting of all the librarians of college libraries in Andhra Pradesh as on April 1, 1998). Other examples are: All life members of ILA by the end of calendar year 1997; All female librarians employed in Hyderabad city during 1991 and 1995; All librarians of agricultural libraries in Andhra Pradesh as on June 1, 1997; All science books purchased by DrBRAOU Library during the period 1988 to 1997; All articles based on case studies published in LIS journals in 1996.

Sampling Unit: A single number of a sampling population (e.g., a city, a citizen, a household, a library, a librarian) is referred to as a *sampling unit*. Usually sampling units have numerous attributes, one or more of which are relevant to the research problem. For example, "Attitude of College Librarians in Andhra Pradesh towards Computerization". The sampling units have many traits (variables), including, opinions, attitudes, expectations, academic qualifications, computer training, etc. The research project may examine only one variable or relationship among different variables.

Finite and Infinite Population: A population may be finite or infinite, depending up whether the sampling units are finie or infinite. A *finite population* contains a countable number of sampling units. For example, all college librarians who are rolls as on April 1, 1998. An *infinite population* consists of an endless number of sampling units, such as an unlimited number of coin tosses.

Sampling Frame: A complete list of all sampling units, from which a sample is selected, is called a sampling frame. Ideally, the sampling frame should include all the sampling units in the population. But in practice, a physical list rarely exists and hence researchers compile a substitute list. For example, in a large national study of public libraries, it is impossible to obtain a complete and accurate listing of all public libraries in India. In small scale research studies, the sampling frame is generally based on telephone directory, membership list of associations, voters list of a village, etc. The accuracy of a sample depends on the sampling frame. Hence researchers should be careful about the typical problems in sampling frames, such as incomplete frames (real sampling units excluded or missing), clusters of elements (more than one individual in a household), blank foreign elements (new units not included in outdated lists), etc.

9.3.2 Advantages and Limitations of Sampling

Sampling studies are becoming more and more useful in social research. Recent developments in sampling techniques have made these more reliable and valid. However, there are certain advantages as well as disadvantages for the sampling method. Let us examine them.

The advantages of sampling method are as follows: (a) It saves time and money, (b) It is possible to conduct minute observations and detailed study, (c) It is possible to achieve a very close range of accuracy with a properly selected sample, (d) It is more convenient to manage a small sample.

The limitations of sampling method are as follows: (a) It may sometimes lead us to biased selection and thereby lead us to draw false generalisation, (b) Selection of a completely representative sample is very difficult draw, especially in case of too small or too heterogenous population, (c) Researcher need to specialised knowledge of sampling techniques, (d) The cases selected through sampling method are sometimes widely dispersed and inaccessible.

9.3.3 Sample Designs

A number of sampling methods are used for drawing representative samples from the population under research study. Every researcher must be sufficiently familiar with these sampling techniques. The sampling techniques are basically grouped into (i) Probability sampling, and (ii) Non-probability sampling.

9.4 PROBABILITY SAMPLING

There are various methods of sampling that can be used in drawing samples from a given population to draw inferences about the population parameter (s). These methods have been classified into two categories viz., probability sampling methods and non-probability sampling methods. To ensure representativeness and also to eliminate biases, the probability sampling is of very much use. The theory of probability will help us in finding the chances of a particular sampling unit being included in our sample. In the estimation of population parameters the theory of probability can be made use of in respect of probability sampling but such an estimation of population value is not feasible in the latter. A few popular methods of sampling are discussed below.

9.4.1 Simple Random Sampling (SRS)

In Simple Random Sampling (SRS), every individual unit and every combination of units have equal chances of being chosen. Selection of a unit is independent of selection of any other unit. There are two types of random sampling viz., SRS with replacement (SRSWR) and SRS without replacement (SRSWOR). In the former, the probability of selection at any stage is the same and is equal to 1/N where N is the size of the population. So, in the sample same units may get included more than once. In the case of latter, the chances of a unit being selected increase with successive draws and on any given draw the probability of remaining units being selected is equal regardless of units previously included.

A number of methods are used for selection of the sample units randomly and use of random number tables is more widely practised. Prepare the sampling frame after taking due care to avoid repetitions and omissions i.e. the list should be complete in all respects. Failure to prepare an exhaustive sampling frame implies that SRS is not the right one for adoption. The units have to be assigned numbers. For instance, let the size of the population be 853. Then the units will be given numbers starting from 001, 002, ..., 852 and 853. Take random number tables with three digits and choose either columns or rows. If the first random number is 230 then the unit with the number 230 is included. Let the second random number be 962 since there is no unit bearing the same. So, you have to move to next random number. Repeat this till you get the required sample size.

If population (N) is very large, use of SRS WOR may not violate the assumption of independence significantly.

9.4.2 Stratified Random Sampling

It is a method which can help to prevent biases and also to assure a more representative sample. Stratification of population into certain strata (groups) is in the direction of experimental control. The principle of stratification operates with sub-groups of more homogeneous composition within a large population. In other words, the population is partitioned in such a way that the units within each stratum are homogeneous with respect to the characteristic under consideration and inter-strata variation is high. This will also reduce the sample size as well. A good example is the classification of libraries into libraries in professional colleges, degree colleges, universities or based on management such as government and private. Similarly users of libraries can be grouped into students, research scholars, faculty members and others. Once the population is stratified, sampling units can be drawn from each stratum following either simple random sampling or systematic sampling. Allocation of sampling units to each stratum can be proportional or disproportional to its size.

i) **Proportional Stratification:** The number of units to be sampled from a stratum is directly related to its size. So more and more units will be drawn from the strata which are (relatively) larger in size.

ii) **Disproportional Stratification:** Allocation criterion is independent of the size of the stratum. When the strata are homogeneous, a small sample itself can capture the variability and also is more cost-effective. However, this type of stratification does not permit inter-strata comparison. Under certain circumstances, the adoption of stratified random sampling improves the overall efficiency of the sampling design. The higher the intra-strata homogeneity, the greater will be the increase in efficiency. A given degree of accuracy can be obtained with a smaller sample under stratified random sampling.

9.4.3 Systematic Random Sampling

This is one of the popular sampling techniques and its efficiency lies in the selection of the sample when the sampling frame is complete in all respects. The basic principle of drawing the sample is more or less similar to simple random sampling.

In this method instead of using random number tables for selection of all the units, only the first unit will be selected randomly from the universe and the latter units are picked up from the universe using the sampling fraction. For instance, if the first drawn is unit 11 and the sampling fraction is (k = N/n) 30 then 41, 71, 101, ... units will be selected. So, it is obvious that the selection of the first unit should be done carefully using random number table or such other device. It is very much essential that there should not be any specific pattern or trend while assigning the numbers to the various units in the population. In the absence of such care, there is a chance for some type of categories being unrepresented in the sample. As mentioned earlier, the list of all units should be complete in all respects and duplications and omissions should be avoided in administering this method.

9.4.4 Multi-Stage Random Sampling

It is a method of drawing the ultimate units for data collection from a given population, when the population is very large. To reduce the sample size without sacrificing the precision and also to ensure representativeness, the multi-stage random sampling techniques is resorted to. At each stage sampling units are drawn randomly.

In a large scale study on the functioning and problems of libraries in a state, the state can be divided into geographical regions viz., districts. Here, the first stage units are districts. A few districts can be selected randomly from this set of all the districts of the state. Within the selected district (s) the libraries can be grouped into three categories: public libraries, libraries attached to educational institutions and private libraries. From these second stage units, one can draw randomly a few units from each category and the survey can be carried out either by personal interviews or questionnaires etc. This method of sampling is found to be very effective in large scale surveys where preparation of the sampling frame itself is either very difficult or costly in terms of time and resources.

9.4.5 Cluster Sampling

You are aware that in the stratified sampling we divide the entire universe into certain groups or classes known as strata and we randomly choose sample units from each stratum. The assumption is that each stratum is homogeneous and the inter-strata variability is high. But in certain situations enumerating the samples from all the strata may turn out to be time consuming and costly. In such situations the population is partitioned in such a manner that each group will contain more heterogeneous elements and inter-group variation will be small. This is conceptually different from the stratified random sampling and in fact diametrically opposite to the stratification concept. These groups are called clusters.

For example, in conducting large scale surveys of households living in cities, it will be economical to divide the city into certain clusters where one may find people of different socio-economic and cultural background in each of these clusters. In the cluster sampling all the units in the selected clusters will be completely enumerated. If the clusters are formed, with proper care enumeration of even a few clusters will provide us the required data with the expected degree of accuracy. Selection of clusters may be based on principles of random sampling. The cluster can also be treated as multi-stage sampling with the first stage units being the clusters and final / ultimate stage units being the units in the selected clusters. Once the clusters are identified, the researcher can prepare the list of sampling units and carry out the survey.

It is evident that in the case of a single stage cluster sampling there is no scope for sampling error since every unit within the cluster is being enumerated. In the case of selecting a few clusters our sampling error depends upon the variability between the clusters. If the clusters are highly heterogeneous, cluster sampling will be cost-effective and also yield highly desirable results.

9.5 NON-PROBABILITY SAMPLING

Non-random or non-probability sampling is one in which the probability or chance of inclusion of a unit in the population is unknown. You may recall that in random sampling, the probability selection mechanism allows application of statistical theory to examine the properties of estimates. The actual precision of the estimates can also be worked out. But in non-probability sampling such applications of theory is not possible and no valid estimates of risks of error can be made. Thus, there is no legitimacy to the inferences drawn and is entirely a subjective evaluation. However, non-random sampling methods are widely used and more so in exploratory studies which will provide valuable insights and also to identify a hypotheses for testing in future research. Its popular use can be attributed to convenience and relatively lower cost.

9.5.1 Accidental Sampling

Also known as haphazard or convenience sampling. Sample selection is made based on the (easily) available information e.g. the users of a library on the day of the enquiry, people visiting panchayat office during the period of investigation. The selection of such samples is justified on the grounds of the convenience.

Given the biases in drawing such samples, the risks of generalising the results beyond the sample are very high. Generalisations are justified only when one is sure and establishes that the sample is a good replica of the defined population.

9.5.2 Quota Sampling

This type of sampling method is more frequently used in public opinion surveys. The quota sampling appears to be similar to the stratified sampling. The interviewer will be given some 'quotas'. For example, in a sample survey to elicit the views of users of a university library, the investigator may be asked to cover 10 professors, 15 readers, 30 lecturers, and so on.

Within each category, the selection of the individual / person is left to the discretion / convenience of the investigator.

In situations where sampling frames are not available or incomplete quota sampling is found to be effective.

9.6 SAMPLE SIZE AND ERRORS IN SAMPLING

There are several misconceptions about the appropriate size of a sample. A sample may include only one sampling unit, or all units or any number in between. To estimate the adequate size of the sample properly depends on the expected level of accuracy and the acceptable level of the standard error.

9.6.1 Determination of Sample Size

Many persons believe that the number of units in the sample (n) should be related to the size of the population (N). In a few studies, you might have seen that a five or ten per cent of the population has been taken as the sample size. If the population is more homogeneous (heterogeneous), you require small (large) sample. In fact, the answer to the question depends upon what one would like to do with the result and also a number of factors have to be taken into account for determination of sample size.

In any standard text book, you will find sample size is a function of the degree of accuracy of the estimate and cost. The sample size is derived from the formula for confidence intervals. So, the size of the sample depends upon

i) confidence level chosen for estimating the population parameter (α)

ii) degree of accuracy with which we would like to estimate the parameter (k)

iii) some reasonable estimate of the variability of the parameter (s)

$$n = (\alpha)^2 \frac{(s)^2}{(k)^2}$$

The measure of variability (s) can be obtained from pilot studies, the earlier works undertaken in that region, etc. The standard deviation of the characteristics reported in those studies or the value obtained from the pilot study can be used. The value of K depends upon the choice of the researcher.

α = 1.96 for 95% confidence level.

The sample sizes for few hypothetical values of s and k are given below :

s	k		
	10%	20%	30%
1.0	384	96	43
2.0	1537	384	171

From the table, it is evident that the optimal size of the sample (n) varies with changes in the values of s and k. However, it does not mean that cost of conducting survey has no role in the determination of sample size. In fact, the optimal value of n should minimise the survey cost per unit and maximise the precision.

In multi-subject integrated surveys, the optimal sample size has to be worked out for each parameter / characteristic and then maximum value of 'n' among this set has to be chosen for drawing the sample.

9.6.2 Reliability and Bias in Sampling

Reliability of a sample is defined as the proportion of its (observed) variance to the true variance. As mentioned earlier, that the main purpose of sampling is to draw inferences about the population parameter or the relationships of the variables without resorting to complete enumeration. The validity of the generalisation or meaningfulness of the inferences depends upon the nature of the sample. The reliability of the sample hinges upon the accuracy of the estimate of the parameter one obtains from the sample. So, for ensuring reliability (of the parameter), the sample selection should be unbiased. In other words, the reliability of the results is related to the representativeness of the sample vis-a-vis the population at large.

The representativeness is decided by the adequacy of the sample size but it does not necessarily imply that one should opt for a large sample. What is needed is that the sample selected is a replica of the population i.e. the extent of heterogeneity with respect to the characteristic under study in the population should also be reflected in the sample. For instance, if you want to collect the views of users of a mobile library facility in a town / city, you should not use telephone directory or list of four wheeler registrants as a sampling frame for drawing the sample since 'middle income class' happens to be the major user group of mobile libraries. Such samples need not be representative and the results cannot be made use of for deriving valid inferences. Further, the term reliability connotes that the results should be more or less same even if different persons follow the (same) procedure.

The following precautions will help ensure reliability of sample surveys: i) obtain the details regarding the nature of universe / population and this will help ensuring representativeness of the sample ; and ii) determine the sample size scientifically. Reliability, to some extent, increases with the size of the sample and it also varies inversely with the variability (of the parameter).

Biases in sampling arises due to several factors which include:

a) *faculty selection of the sample* due to either inappropriate choice of sampling method or non-adherence to the randomisation process;

b) *substitution* of the unit in the sample by another owing to the problems of non-availability, non- cooperation etc. The inclusion of neighbouring household is not an uncommon practice. The discrepancies in the characteristics of the substitute from the origin will result in errors. In a few studies, more number of units will be sampled and these excess units can be made use, if necessary.

In fact, by minimising the biases, the reliability of the sample will increase.

9.6.3 Sampling and Non-Sampling Errors

Sampling errors (SE) are those which occur by virtue of the fact that only one sample is drawn from the universe and inferences are made on that sample. Since samples vary from one to the next and the estimates are likely to vary from one sample to the other an error may be introduced by taking decisions based on one sample. However, the probability theory helps us to estimate these errors. By choosing an appropriate sampling design, the sampling errors can be minimised.

Non-sampling errors (NSE) are referred to as the errors of measurement. These occur on several counts which include faulty definitions, investigator's bias, ambiguity in the questions, interviewer's bias, errors in coding, processing, analysis and interpretation of data. They can not be estimated. Even when utmost care is taken in the design of the study taking all precautions for drawing the sample, certain non-sampling errors are bound to occur. The non-sampling errors can be controlled to some extent by 'pre-testing' of the instrument, imparting training to field staff, effective supervision etc.

There is a trade-off between sampling errors and non-sampling errors. As the sample size increases, sampling errors can be reduced but non-sampling errors are likely to increase and vice-versa.

9.7 LET US SUM UP

One of the important stages in Social Science Research is data collection. The two major steps involved in this stage are: identification and design of the instrument for data collection and method to be adopted for drawing the sample from the population.

The merits and demerits of three instruments for data collection including questionnaire are discussed. The two major branches of sampling vize., probability sampling and non-probability sampling are discussed at length. The appropriateness of these sampling techniques is also pointed out. The optimal sample size depends upon a number of factors and a formula for working out the sample size is mentioned. Steps to be taken for minimising the bias and thus increasing the reliability of sampling are presented.

9.8 REFERENCES AND RECOMMENDED BOOKS

BLALOCK, H.M. *Social Statistics*. New York: McGraw-Hill, 1972.

GOODE, W.J. and Hall, P.F. *Methods of Social Research*. New York: McGraw-Hill, 1952.

GUILFORD, J.P. *Fundamental Statistics in Psychology and education*. New York: McGraw-Hill, 1956.

KALTON, Graham. *Introduction to Sampling*. London: Sage, 1983. (Sage Publication Series, No.35)

MURTHY, M.N. *Sampling Theory and Methods*.Calcutta: Statistical Publishing Society, 1967.

ORENSTEIN, Alan *et al. Understanding social approach: an introduction*. Boston: Allyn & Bacon, 1978.

9.9 ASSIGNMENT

(i) You are asked to find out the percentages of books lost in a library. What sampling techniques you will use and why? Please furnish the details of the approach you are going to adopt?

(ii) You have been asked to conduct a survey of the socio-economic conditions of the staff of public libraries in Andhra Pradesh. Prepare a questionnaire to collect data also on their views on the functioning of the libraries and suggestions for improvement.

9.10 GLOSSARY OF TERMS

Population or *Universe*: Aggregate of objects, persons or events; set of all the units from which samples are selected and studied. E.g.: In a study on the growth of libraries in Andhra Pradesh, all types of libraries in the state constitute the population.

Sample: A part of the population and so, only a few units of the population are included.

Sampling: Study of a part or sub-set of the population with a view generalise the results obtained from the sample to the entire population.

Sampling unit: Unit to be sampled e.g. person, library, book in a library, household, user of the library services, village etc.

Sampling frame: List of all sampling units.

Sampling fraction (k): Ratio of units in the population to the units in the sample (N/n).

Sampling error: difference between a population value (parameter) and the corresponding value derived from the sample.

Standard error: measure of variability of the population parameter. It indicates the accuracy of the estimate It measures the stability of the estimate obtained from the sample. It is the standard deviation of the sampling distribution.

Parameter: Any characteristic of the population such as variance, mean and proportion. E.g.: average number of visitors per day to the public libraries in Hyderabad.

Statistic: Characteristic of the sample ; used to estimate the parameter value.

9.11 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) What are the instruments used for data collection and discuss the merits and demerits of each of them ?

2) A well designed 'sample survey' is better than poorly planned and conducted 'Census' - Substantiate this view .

3) Compare and contrast the following probability sampling methods: i) Simple Random Sampling; ii) Stratified Random Sampling; iii) Cluster Sampling

4) What factors determine the size of the sample? State the steps to be taken to increase the reliability of sampling

II. SHORT NOTES

a) Cluster Sampling

b) Sample Error

c) Universe, Population and Sample

d) Census method

UNIT - 10 : DATA ANALYSIS - STATISTICAL METHODS

Structure

10.0 AIMS AND OBJECTIVES

Data Analysis is an important aspect in research methodology. The present unit introduces you to Data Analysis using statistical methods.

After going through the unit you will be able to

- classify the variables and apply appropriate statistical techniques for analysing data
- use the right measure of association between two variables (for both) quantitative and qualitative category types of variable
- apply the regression method to quantify the effect of independent variable on dependent variable
- delineate various components in a time series data and carry out significance tests for testing the validity of hypotheses.

10.1 INTRODUCTION

In the previous unit we have discussed the development of an instrument for data collection and also the choice of the sampling method to draw a representative sample. In this unit we shall introduce you to some selected important techniques of statistics useful for analysing the data obtained from the primary and secondary sources. As a prelude, we will classify the types of variable and discuss the statistical measures like measures of central tendency and measures of dispersion to analyse the information contained in the variables. Attempt is also made to suggest the appropriate measure in the context of various types of variables.

In the second part of the unit the measures of association will be discussed at length. The correlation coefficient, the chi-square test etc., will be introduced to find out the nature and strength of relationship between the variables. Further, to assess the likely effect of an independent variable on the dependent variable, regression method will be presented along with the technical part of estimating a simple regression equation.

The concept of time series and also the components of time series are explained. Steps to isolate and estimate the various components are also suggested.

Testing of hypothesis is an important theme under statistical inference. Initially the key terms associated with the hypothesis testing are discussed followed by the application of t-test in three important situations.

10.2 PROCESSING OF DATA - EDITING AND CODING

After the collection of data using either questionnaire or schedule it has to be processed for statistical investigation. The first step for the processing of data is *editing and coding.*

10.2.1 Editing of Data

In editing, the schedules or questionnaires which are also known as returns collected from the respondents will be scrutinised carefully to detect the errors that creep in due to various factors. Some of the common errors are mentioned below :

1) Factual Errors

These types of error occur by wrongly entering the information in the schedule. For example, the names of the district of one state may be reported against some other state which can be checked and corrected.

2) Irrelevant Information

It is not unusual to find such type of responses which have no relation with the particular aspect being investigated. These schedules/questionnaires can be sent back, if possible, for correction.

3) Arithmetical Errors

These arise due to confusion with respect to units of measurement. Sometimes abnormal figures will also be reported in the returns which can be corrected by cross- checking with the other data provided in the return. For example, total production of the cropped area values may be reported against production per acre. So, provision of built-in checks in the computer programme will help identify these types of errors.

Wrong Classification: A few subjects are wrongly classified and are shown in categories to which they do not belong. The information relating to student of a university may be shown against the category earmarked for professors. This can be corrected by referring to the field investigators or by cross- checking the data on parameters like age.

Some of these errors can be edited by recanvassing the schedule to the respondent. However, given the time and cost constraints, such methods are less likely to be employed in the normal practice. By, using the information from the rest of the schedule or questionnaire, some of these errors can be checked and edited. So, in editing, one has to check for internal consistency and where internal consistency is not ensured one may have to opt for a 'call back' for correction or rejection of the return. In the case of large scale surveys, rejection of a few returns may not effect the outcomes or inferences of the survey. However, as the number of incomplete and / or defective returns increases, the reliability of the data will be reduced. If the percentage of returns which are either incomplete or defective is more, then it is advised to scrap the survey and go for a fresh study.

10.2.2 Coding of Data

To carry out the analysis of data either manually or with the aid of computers, the various answers given by the respondents to the set of questions administered will be assigned codes. Such conversion of responses into a set of numbers is called coding. These numbers assigned should not be interpreted as values or weights to the various possible responses. Assignment of codes merely facilitates further analysis. The code is fed into the computer and thereafter each code corresponds to a particular response to the question under consideration. These codes can be assigned before data collection or afterwards but the former will reduce the burden of the investigators during the survey. It is self evident that the exhaustive and mutually exclusive principles have to be applied in the assignment of codes. All possible answers to a given question should be predicted and should be allotted a distinct code and there should be one-to-one correspondence between the codes and responses.

In the pre-coded schedules/questionnaires, the researcher has to predict the possible range of answers to the question. For example, the answers to the question on the 'marital status' of the respondent fall into one of the following categories: married (1), unmarried (2), divorced (3), separated (4), widowed (5) and remarried (6). The numbers in the parentheses are the codes. So, code '5' for a respondent indicates that he / she is a widow / widower. These codes do not possess any ordinal or cardinal value.

In the post-coding situation, the number of categories may be less than 6 if some of these are not relevant for the sampled population.

Open-ended Questions: As you have seen in the construction of schedule/questionnaire, a few opinion-oriented questions may be included. In some situations the researcher due to incomplete knowledge cannot visualise the range of answers. In those cases, post-coding appears to be more appropriate.

In reality, we may find that both pre-coding and post- coding operations will be carried out in the same survey.

10.3 STATISTICAL MEASURES FOR DATA ANALYSIS - ONE VARIABLE

Before learning about the statistical measures used in data analysis, we should first understand about different types of variables.

10.3.1 Types of Variables

The social scientist will collect data on various socio-economic, cultural and political variables either directly from the respondents or collect relevant data from published and unpublished sources. These variables are classified into quantitative and qualitative variables.

The *quantitative variable* takes different values which are measurable and objectively verifiable. The *qualitative variables* are those which are basically categorical in nature. For instance they refer to various disciplines on which books are classified and maintained in the libraries. For instance, the classification of books belonging to sociology, economics and political science etc., is a good example. The other examples are the type of users — professors, readers, lecturers, research scholars, students and others.

The qualitative variables are again classified into nominal variables and ordinal variables. The *nominal variables* are those which partition the data into certain classes or categories as mentioned. In other words, any element in the data set should belong to one and only one category. So, these categories or classes should be exhaustive and mutually exclusive. If there is an explicit or implicit order in the nominal variable it is referred to as *ordinal variable*. The classification of employees working in the libraries as per the scales of pay (in ascending order) is an example of ordinal variable. Similarly the users can be classified according to the level / years of research or teaching experience. The statistical techniques have been evolved overtime to treat these two types of variables mentioned above.

The quantitative variables can be further sub-classified as continuous and discrete variables. The *continuous variable* is one which can (theoretically) take any value in the range of the variable. For example, the heights and weights of students can be measured to the last digit (depending upon the precision of the instrument employed). On the contrary, *discrete variable* is one which can take certain selected values of the variable eg. the number of library books issued to the students will only take an integer value.

10.3.2 Simple Methods for Data Analysis

The purpose of data analysis is to condense the large volume of data into certain meaningful values. In other words, statistical analysis helps in extracting information that is contained in the data relating to the various types of variables. It also helps us in pointing out the latent relationships and trends in these variables. It is evident that human eye cannot delineate the patterns and trends among the various variables/characteristics when the number of observations and also the number of variables are large. So, the problem becomes all the more challenging and difficult when the data are collected from large population. So, we need certain summary measures derived by condensing the data to answer the questions selected for this study.

1) Frequency Distribution

Frequency is the number of times a particular value of the variable repeats itself in the data. If the female employees in library are found out to be 18 out of a total strength of 51 the frequency of female employees is shown as 18. The frequency distribution is a series of observations which are similar or closely related are pooled together and classified into separate groups/classes. The clustering of observations is made in the order of magnitude of the variable either individually or as a group. The frequencies are juxtaposed to the respective group or values of the variate. If the values of the variable are considered individually it is referred to as discrete frequency distribution when presented as groups it is known as continuous frequency distribution. In the table below, the data on age group-wise classification of employees in a library are shown.

Table-1 : Age Group-wise Classification of Employees

Age group	Frequency (No.of employees)	Age group (years) (Inclusive type)
18 - 21	3	18 - 20
21 - 25	8	21 - 24
25 - 35	15	25 - 34
35 - 45	16	35 - 44
45 - 60	9	45 - 59
	51	

In the above example there are five classes and the first class contain all those employees who have completed 18 years but having age less than 21 years. In other words, the employees with age 21 years will be included in the second class.

Class Limits: The smallest and the largest possible value in each class are called as lower limit and upper limit of the class respectively. In the above example, the class intervals under the third column include the upper limit of the class.

Class Mark: The class mid point is also known as class mark and it is obtained by taking the average of lower and upper limits of the each class. For instance, 19.5 is the class mid-point for the first class and this is taken as a representative value of the first class.

Class widths: The difference between the upper and lower limits of the class is known as class width. It is suggested that the number of classes should not be too small or too large and to derive the optimum number of class intervals one may follow the Sturtge's rule.

$$n = 1 + 3.32 \ \log N$$

where n = no.of classes

N = total no.of observations

2) Proportions and Percentages

For analysing nominal variables proportions and percentages are frequently used. The number of cases falling in each category of the variable are counted (frequency) and its relative size *vis-a-vis* the total number is considered. For instance, if x1, x2, x3 and x4 represent the number of books belonging to four different disciplines and x (x1 + x2 + x3 + x4) is the total number of books in the library, then the proportions of books in various discipline are given by x1 x2 x3 and x4 respectively.

It is evident that the sum of the proportions should always be equal to 1. The categories or groups should be mutually exclusive and exhaustive. These proportions will be useful in comparing two or more groups when the absolute values differ substantially. In the table given below data are given regarding type of visitors to two university libraries during the first week of January, 1994.

Visitor	Library1	Library2	Proportions	
			Library1	Library2
a) staff	124	200	0.62	0.57
- teaching	56	150	0.28 (.45)	0.43 (.75)
- non-teaching	68	50	0.34 (.55)	0.14 (.25)
b) Others	76	150	0.38	0.43
Total	200	350	1.00	1.00

Note: Figures in brackets are proportions worked out with staff size as the base.

The absolute figures are not easily amenable for comparison as the bases (or total number) are different. By looking at proportions, it can be inferred that relatively a higher proportion of users belong to 'staff' in the case of library-1. In relative terms, the user community contains a higher proportion of 'others' and a lesser proportions of staff in the latter. The teaching staff account for a higher share in library-2 when compared to library-1. Calculation of proportions with the staff as base indicate that the situation in university library-1 is obverse to that of library-2.

It is to be noted that sum of the proportions should be always unity but in a few cases the total need not be exactly unity because of rounding off errors. The convention is to change the figures for the categories having largest proportions (of cases). The assumption is small change in the decimal place in a larger proportion is relatively less significant than the same change in a smaller figure.

Proportions when multiplied by 100 give percentages. In fact, the term 'per cent' means per hundred. It is one way of standardisation which will be useful for easy and direct comparisons. If the number of cases are small, the use of percentages will be misleading and it is suggested that the researcher should make use of the absolute figures only. However, the percentages can be shown along with the absolute values to avoid confusion and wrong interpretation.

Exercise : Data relating to the college students belonging to different disciplines and using the library facilities are collected during the reference month of a survey and are presented below.

Frequency of use of library	Economics	Sociology	Political Science	Others	Total
Regular	20	30	20	15	60
Occasional	10	20	25	20	100
Total	30	50	45	35	160

Find out i) proportion of regular users who are students of sociology

ii) proportion of occasional users among students

iii) are there relatively more regular users among political science students than among 'others' and

iv) you prefer to compute percentages row-wise or column-wise and why ?

3) Averages

It has been observed that there is a tendency on the part of the data to centre around certain values. This tendency of data to get centred around at some values is called Central Tendency. Mean, Median and Mode are some of the popular measures of central tendency.

Mean ($\bar{X}$) : In general terms, mean of any variable is the average value of the observations.

$$\bar{X} = \frac{(x1 + x2 + ... + xn)}{n}$$

Let daily wage (Rs.) of five persons be : 13, 14, 15, 16, 17

The average wage : (Rs.) 75/5 = Rs.15/-

Weighted mean: If some values of the variables have special importance, then the mean should be taken as the weighted average

Category	No.of persons	Wages paid (Rs./day)
Females	100	13/-
Children	200	7/-
Males	50	15/-

The average wage rate per worker is not $\frac{13 + 7 + 15}{3} = \frac{35}{3}$ = Rs.11.67/-

but $\frac{100 \times 13 + 200 \times 7 + 50 \times 15}{100 + 200 + 50} = \frac{3450}{350}$ = Rs.9.85 /- (approx.)

Usually, the observations will be processed and frequency distributions are prepared for further analysis. Calculation of mean, for ungrouped (discrete case) and grouped (continuous case) data is as follows.

Discrete case: Frequency distribution of marks obtained by 24 students is given below :

Marks (x)	Frequency (f)	fx
49	3	147
60	2	120
71	5	355
52	5	260
63	6	378

$$\bar{X} = \frac{\Sigma fx}{\Sigma f} = \frac{1260}{21} = 60$$

Continuous Case: The frequency distribution is prepared by forming certain class intervals. The mid-point of the class interval is assumed to be a representative value of the class. The weekly wages of 23 workers are given below :

Monthly Salaries Rs.	Mid point (m)	No.of workers (f)	fm
99.5 - 104.5	102	5	510
104.5 - 109.5	107	3	321
109.5 - 114.5	112	8	896
114.5 - 119.5	117	2	234
119.5 - 124.5	122	2	244
124.5 - 129.5	127	3	381
		$\Sigma f = 23$	$\Sigma fm = 2586$

$\bar{X} = \Sigma fm / \Sigma f = 2586 / 23 =$ Rs. 112.4 /- (approx.)

1) The extreme values influence the value of the mean
2) The open ended class intervals create problems in estimation of mean.
3) The mean may not be equal to any of the observed values of the variable.

Median (M):

Median(M) is the value of the variable which divides the series in such a manner that the number of items below it will be equal to the number of items above it. M is a location measure or positional average.

Arrange the items in ascending order and the value corresponding to mid-item is the Median.

Let 8, 10, 1, 3, 5 be the set of five observations (n).

The observations are arranged in the ascending order of magnitude : 1, 3, 5, 8, 10

Since n is odd, the 3rd observation is the median value (M) and is equal to 5.

In case of even number of observations, the median is the average value of middle observations. For the series of observations : 1, 3, 4, 7, 9, 10

Median(M) is given by the average of the 3rd and 4th observations is $\frac{4+7}{2} = 5.5$

The values of other items need not be known to us and only their ordering is needed. e.g. Arrange the boys according to their intelligence, then boy at the middle position is the boy with median (average) intelligence. M is not influenced by extreme values.

Example: Find out the median size of the town libraries (in terms of book collection) in a district from the following frequency distribution

No. of books (X)	No. of libraries (f)	Cumulative frequency C.F.	
100	5	5	
150	6	11	
120	7	18	— No.120 is the median
130	8	26	
200	10	36	

Cumulative frequency (CF) is obtained by adding the preceding totals to the class frequency i.e. sum of the frequencies of upto and including the class under reference. Calculate the cumulative frequency distribution and find out the library size corresponding to N/2 th observation i.e. 36/2 = 18th observation. The number of books in the 18th town library is 120.

Median value or size of the town libraries is the size in which the Median / item falls. Median size of the town libraries is 120.

The formula for working out the median for continuous case is

$$M = L + \frac{(N/2 - C.F)}{F} \times i$$

where L = lower limit of median class*
N = No.of observations
C.F = Cumulative frequency of the class preceding the median
F = Frequency of median class
i = Interval of the median class

* The class interval in which the median (i.e., N/2 th observation) falls.

Example:

Find out the median value of the weekly wages of 200 library workers.

Weekly Wages (Rs.)	Mid point	No.of workers (f)	Cumulative frequency
099.5 - 104.5	102	15	15
104.5 - 109.5	107	30	45
109.5 - 114.5	112	80	125
114.5 - 119.5	117	25	150
119.5 - 124.5	122	20	170
124.5 - 129.5	127	30	200

$$M = 109.5 + \frac{55}{80} \times 5 = 112.9$$

Mode (Mo):

Mode is the value of the variable having maximum frequency.

Let daily wages (Rs.) be 12, 13, 14, 15, 14, 14 and 16 for seven workers. The Mode is Rs. 14/-

Modes can be many.

Example: If weekly wages (Rs.): 200, 250, 300, 350, 400, 400,250, 250, 40, then Modes are Rs.250 and Rs.400 (bi-modal distribution).

Mode is one having heavy concentration and it is a typical value of the series. Useful for comparison of two groups of students in their performance. There are several methods for calculation of mode yielding generally different (modal) values of the series.

Grouped series: Mode is given by

L : Lower limit of modal class

fo : frequency of the class preceding the modal class

f2 : frequency of the class following modal class

i : interval of modal class

$$Mo = L + \left[\frac{fo}{fo + f2}\right] \times i$$

Modal value of weekly wages (the problem referred to in the earlier)

Modal class : 109.5 - 114.5

$$\text{Mode, } Mo = 109.5 + \left[\frac{25}{30 + 25}\right]$$

$$= 109.5 + (25/55) \times 5 = 109.5 + 2.27 = \text{Rs.}111.77$$

The formula / equation connecting the three measures of central tendency is

Mode = Mean - 3 (mean - median)

For the above example,

Mode = 112.4 - 3 (112.4 - 112.9) = 112.4 - 1.5 = 113.9

Characteristic of a good average :-

- should be rigidly defined
- should be easy to compute and understand
- should use all observations
- should be insensitive to (extreme) single observation
- should be amenable for further algebraic manipulation treatment
- should have sampling stability

Nature of variable and relevant 'average' :

Variable	Average		
	Mean	Median	Mode
Quantitative	✓	✓	✓
Qualitative			
- Ordinal		✓	✓
- Nominal	✓		

4) Measures of Dispersion

The averages or measures of central tendency (mean, median and mode) help us to calculate the levels of the characteristic under consideration which can be used as a representative of the data. The representativeness of this measure depends upon the nature of the data or the distribution of the data around these measures of central tendency. If there is skewness in the distribution of values of the variable or if there are a few extreme values then the averages may not be representative of the data. For instance, the per capita income of India is a misleading average to know the (average) standard of living of the people since a large proportion of the people are poor. Further, a single measure may not be adequate to represent the entire series or to extract the information contained in the data as these measures may fail to reveal the hidden trends in the distribution of the variable. Thus, there is need to supplement the averages with measures of dispersion or spread or variability.

These measures essentially capture the scatter of observations around the averages and reflect the extent of variation in the values of the characteristic. Lesser the variability, greater is the homogeneity in the series of observations and better is the average as a representative of the series. And conversely higher the variability, greater is the heterogeneity and the averages fail to be representatives of the series. The measures of variability will also help us in comparing two groups or two situations with reference to the variable or characteristic under study.

Range: It is one of the simplest measures of variability and is measured by taking the difference between the maximum and minimum values in the series. So, higher the range, greater is the variability and vice-versa. In the case of extreme values, range is a misleading measure of variability.

Quartiles: These are the values of the variable which split the series into a number of (equal) parts. Percentiles, deciles and quartiles are measures which partition series into hundred, ten and four equal parts respectively.

Quartile Deviation (QD) : Quartiles are those values which partition the series into four equal groups. The first or lower quartile (Q1) refers to the value of the variable below which 25 per cent of the observations lie and above which 75 per cent of the observations lie. The highest / upper quartile (Q3) splits the series of observations in such a way that there will be 75 per cent of the observations below this value. The quartile deviation

(QD) is obtained by $\dfrac{Q3 - Q1}{2}$

In other words, the QD measures the extent of variability between the lower and upper quartiles. Lesser the QD value, greater will be homogeneity in the set of observations. The second quartile which bisects the series is also known as Median.

Standard Deviation(SD) : It is the square root of variance(S). Variance is a measure of variability with reference to the mean. Standard Deviation is given by

$$SD \text{ or } S = \sqrt{\frac{\Sigma(X - \bar{X})^2}{n}}$$

since the sum of the deviations from the mean value will always be equal to zero, a standardised measure of variability is obtained by squaring these deviations and then taking the square root of their sum. Lesser the value of standard deviation, greater will be the homogeneity of the observations.

The following example helps us in working out the above measures and also the interpretation of the results.

Exercise:

Measures of Variability

The monthly expenditure data (X) of 20 village libraries are given below : Compute the various measure of variability and also coefficient of variation of the expenditure of the village libraries. Interpret the results.

X :	3000	2500	3500	1500	1000	3000	1800	3000	2000	3500
	1500	1700	2200	2500	1600	3100	2700	1900	2300	3300

Array:	1000	1500	1500	1600	1700	1800	1900	2000	2200	2300
	2500	2500	2700	3000	3000	3000	3100	3300	3500	3500

Range: Minimum value of X : Rs.1000
Maximum value of X : Rs.3500

Range of X: Rs.3500 - Rs.1000 = Rs.2500

Quartile Deviation (Q.D.) N = 20 (even)
Lower quartile (Q1) : Rs.1,750
Upper quartile (Q3) : Rs.3,000

$$Q.D. = \frac{Q3 - Q1}{2} = \frac{3000 - 1750}{2} = 1265/2 = \text{Rs. } 625$$

Standard Deviation (SD):

$$\bar{X}, \text{ mean expenditure } = \frac{\Sigma X}{N} = \frac{1}{20}(47{,}600) = \text{Rs. } 2380$$

$(X - \bar{X})$:	1380	-880	-780	-680	-580	-480	-380	-180
$(X - \bar{X})^2$:	1904400	774400	608400	462400	336400	230400	144400	32400

$(X - \bar{X})$:	- 80	120	320	620	620	720	920

$(X-\bar{X})^2$: 6400 14400 102400 384400 384400 518400 846400

$(X-\bar{X})$: 1120 1120

$(X-\bar{X})^2$: 1254400 1254400

$\Sigma(X-\bar{X})^2$: Rs. 8561800

SD or S = $\sqrt{428090}$ = 654.286 = Rs.654/- (approx.)

Coefficient of Variation (COV):

Measures of central tendency and measures of variability will help us to provide estimates about the level and distribution of the series of observations for each characteristic under consideration. In general, we come across various situations where comparison across groups becomes difficult. The following two-way table gives us the four situations.

Level or average		Variability: Low (L)	Variability: High (H)
	Low (L)	L — L	L — H
	High (H)	H — L	H — H

There is a need for having measures which can take care of both level and distribution simultaneously. One such measure is coefficient of variation (COV) which is given by

$$COV = \frac{SD}{Mean} \times 100$$

The lesser the coefficient of variation, greater is the homogeneity of the observations and vice-versa. The coefficient of variation for the above exercise is given by

$$COV = \frac{S.D}{X} \times 100$$

$$= \frac{654}{2380} \times 100 = 27.5\ \%$$

The overall variation is 27.5 per cent which is not very high. The COV is useful for comparing a number situations at one time.

10.4 STATISTICAL TECHNIQUES FOR DATA ANALYSIS - TWO VARIABLES

In the preceding section you have learnt about statistical techniques for data analysis using one variable. In this section let us see statistical techniques for data analysis using two variables.

10.4.1 Measures of Association - Correlation

i) Pearson's Correlation Coefficient (r)

It measures the degree of association between two variables which are quantitative in nature i.e. ratio or interval type. If X1, X2, ... Xn and Y1, y2, ... Yn are observations on variables X (say, investment in a scheme) and Y (say, net income from the scheme), then correlation coefficient (r) is given by

$$r = \frac{\text{Covariance (X, Y)}}{\text{S.D (X)} \quad \text{S.D (Y)}}$$

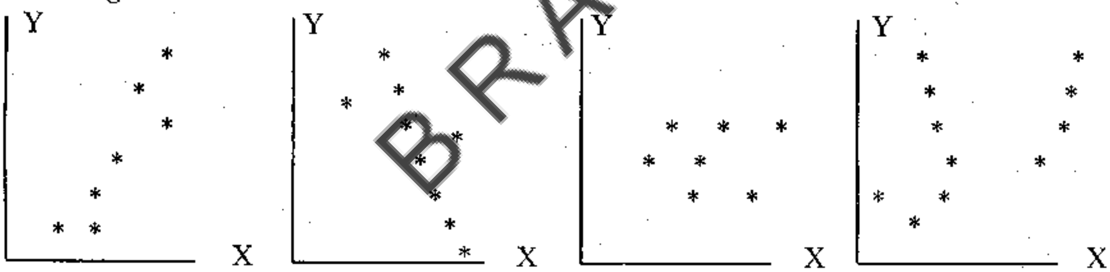
$$= \frac{\Sigma(X - \bar{X}) \; (Y - \bar{Y})}{\Sigma(X - \bar{X})^2 \;\; (Y - \bar{Y})^2}$$

The covariance indicates r = how X and Y vary together. If it is negative, then r will be negative and vice-versa. The value of r lies between -1 and + 1.

r = (X, Y)	= + 1	association is perfect and positive
	= - 1	association is perfect and negative
	= 0	no association i.e. X and Y are independent

The nature and also the strength of association (to some extent) can be observed by Scatter diagram i.e. by plotting the observations on graph.

fig:



a) Positive association   b) Negative association   c) No association   d) No association

Properties:

i) 'r' is unaltered by adding or subtracting a constant to or from X or Y.

$$r = (X, Y) = r\,(X + K, Y) = r\,(X, Y\text{-}K) = r\,(X\text{+}K, Y\text{+}K)$$

ii) 'r' is unaltered for changes in the scale. In other words either multiplying or dividing either of the variables by a constant, the value of 'r' does not change.

Example: Compute the correlation coefficient between socio- economic status (X) and the attitude score (Y).

X = 5 8 9 7 6

Y = 3 7 8 8 5

Solution : Correlation coefficient

$$r = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\Sigma(X-\bar{X})^2\,(Y-\bar{Y})^2}$$

$(X-\bar{X})$	:	-2	1	2	0	-1
$(X-\bar{X})^2$	:	4	1	4	0	1
$(Y-\bar{Y})$	:	-3	1	2	2	-2
$(Y-\bar{Y})^2$	:	9	1	4	4	4
$(X-\bar{X})(Y-\bar{Y})$	:	6	1	4	0	2

$\Sigma X = 35$ $\quad \Sigma Y = 30$ $\quad \Sigma(X-\bar{X})^2 = 10$ $\quad \Sigma(Y-\bar{Y})^2 = 22$

$\bar{X} = 7$ $\quad \bar{Y} = 6$ $\quad \Sigma(X-\bar{X})(Y-\bar{Y}) = 13$

$$r = \frac{13}{(10)\ (22)} = 0.88$$

The correlation coefficient is positive and high.

i) Rank Correlation

This measure provides intensity of association between the ordinal or rank-ordered sets (pair) of observation. Sometimes, the ranks are substituted for the quantitative data for reducing (arithmetic) labour or due to errors in measurement or different methods of data collection e.g. the income for different states are estimated by adopting different procedures. So, ranks are substituted for actual income estimates. The limitation of this approach, however, is that the distance between 1st and 2nd is treated as same that between 2nd and 3rd and so on. Such as assumption may not be t enable.

The rank correlation (R) is given by :

$$R = 1 - \frac{6\ \Sigma d^2}{N\,(N^2-1)}$$

where d : difference between paired ranks

N : number of observations

$R = 1$ when $\Sigma d^2 = 0$

R lies between -1 and +1

Example: The ranks based on awareness score of a RD programme and incomes realised by 10 farmers are given below. Using rank correlation method, find out the strength of association between them.

Farmer	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Awareness score	6	3	7	2	1	8	4	9	5	10
Income	1	2	3	4	5	6	7	8	9	10

Solution: The difference in ranks is calculated and are given below

d :	5	1	4	-2	-4	2	-3	1	-4	0
d^2 :	25	1	16	4	16	4	9	1	16	0

Rank correlation $$R = 1 - \frac{(6)\ (92)}{10\ (100 - 1)}$$

$$R = 0.44$$

The association between the two variables is positive but the strength of the association is not high.

10.4.2 Contingency Measures

In this sub-section, Yule's co-efficient (Q), a measure to find out the nature and extent of association between two attributes, will be introduced. Attribute is a qualitative characteristic and possession or non-possession of the characteristic is used as a criterion for classifying the units/cases in the population. The students can be classified into regular users of library services and occasional users. Capital letters A, B, C are used to denote the presence of an attribute and small letters a,b,c ... for its absence. The frequency or number of cases / units falling into these categories are represented (A), (a) .... The number of units possessing the two attributes are indicated by (AB) and (ab) denotes the number of cases not having both the attributes.

In the association of attributes we will find out whether a given attribute is associated with the other attribute. If so, what is the nature (positive or negative) of association and also the intensity or strength of the association.

Positive association: If both the attributes are present or absent together in the data, then attributes are positively associated. Two attributes A & B are said to be positively associated when the actual number of cases of (AB) is more than the expected number of cases.

Negative Association: If the presence of one attribute is associated with the absence of other attribute, it is called negative association. Two attributes A & B are negatively associated if the actual number of cases of (AB) is less than the expected number of cases.

Independence: Two attributes are said to be independent if the presence of one attribute does not cause the presence or absence of the other. Under independence condition the actual number of cases will be exactly equal to the expected number.

The formulae to calculate the expected number of cases and also the nature of association are given in the table below :

Attributes Association	Positive Association	Negative	Independence
A & B	$(AB) > \frac{(A)(B)}{N}$	$(AB) < \frac{(A)(B)}{N}$	$(AB) = \frac{(A)(B)}{N}$
a & b	$(ab) > \frac{(a)(b)}{N}$	$(ab) < \frac{(a)(b)}{N}$	$(ab) = \frac{(a)(b)}{N}$

Where N is the total number of cases. The Yule's coefficient (Q) is given by

$$Q = \frac{(AB)(ab) - (Ab)(aB)}{(AB)(ab) + (Ab)(aB)}$$

The value of Q lies between - 1 and + 1

when Q = 0 , the two attributes A & B are independent
Q > 1 the two attributes A & B are positively associated
A < 1 the two attributes A & B are negatively associated

Example: The number of employees of a research institute is 600 and the employees have been classified based on : i) the possession of Technical Qualification (A) or otherwise (a) and ii) also on the basis of usage of library services - regular user (B) and occasional user (b). Find out whether there is any association between the two attributes i.e. possession of technical qualification and regular use of library services using the information given below :

	Use of Library services	Qualification		
		Technical (A)	Non-technical (a)	
(B)	Regular	175 (AB)	100 (aB)	275 (B)
(b)	occasional	75 (Ab)	250 (ab)	325 (b)
	Total	250 (A)	350 (a)	600 (N)

$$Q = \frac{(AB)(ab) - (Ab)(aB)}{(AB)(ab) + (Ab)(aB)}$$

$$= \frac{(175)(250) - (75)(100)}{(175)(250) \quad (75)(100)}$$

$$= +0.7$$

So, it can be concluded that the two attributes viz., possession of technical qualification and regular use of library facilities are positively associated.

10.4.3 Chi-square (χ^2)

This measures the association between two nominal variables or between nominal and ordinal variables. On the basis of observed frequencies in the sample, the theoretical frequencies in the population are generated. If the difference between the observed and expected (theoretical) frequencies is significant in statistical sense and not due to chance, then it will be concluded that there is no association between the variables.

In the tossing of the coin, one expects "Heads" and "Tails" to occur with equal chance. Thus, if a coin is tossed 100 times, one expects 'Heads' and 'Tails' to occur 50 times each. These are the theoretical frequencies. In an experiment if 45 'Heads' and 55 'Tails' occurred one can formulate a hypothesis that there is no difference in the occurrence of the two events. Do the observed frequencies constitute evidence for rejection of this hypothesis ? The χ^2 statistic is given by

$$\chi^2 = \frac{\Sigma(O - E)^2}{E}$$

Event Observed	Frequencies Expected	O - E	$(O - E)^2$	$(O - E)^2 / E$
Head	45	50 - 5	25	0.5
Tail	55	50 - 5	25	0.5

$\chi^2_{cal} = 1, \ \chi^2 (5\%) = 3.84$

$\chi^2_{cal} > \chi^2$ tabular value $\longrightarrow$ Reject the hypothesis

$\chi^2_{cal} < \chi^2_{tab}$ $\longrightarrow$ Do not reject the hypothesis

since $\chi^2_{cal} < \chi^2_{tab}$ the hypothesis is not rejected

χ^2 is used extensively in social science research e.g. to test whether a sample drawn is representative of the population. In other words one can examine whether the expected frequencies follow certain relation (ratio) as per theory.

Example: In a sample survey of 200 households, the observed frequencies and percentages by racial origin are given below. The distribution of various races in population is also provided. Test the representativeness of the sample.

Racial origin	Obs. frequency	Distribution (%) Sample	Population	Expected frequency	O - E	$(O - E)^2/E$
French	95	47.5	62.5	125	-30	7.2
English	67	33.5	19.4	39	28	20.1
Others	38	19.0	18.1	36	2	0.1
Total	200	100.0	100.0	200		

$\chi^2_{cal} = 27.4 \quad \chi^2_{tab} = 9.2$

$\chi^2_{cal} > \chi^2_{tab}$ $\longrightarrow$ Reject the hypothesis

The sample is not representative of population.

Chi-square can be used to test the independence of two nominal variables. Then the paired observations will be presented in the bi-variate frequency table, and based on the observed frequencies the association between the variables is studied. These tables are called contingency tables. To generate the expected frequencies the following procedure is adopted.

		Y		
		Y1	Y2	
X	X1	a11	a12	r1
	X2	a21	a22	r2
		c1	c2	n

Let X and Y be two nominal variables with two categories each let r1 and r2 represent the row totals and c1 and c2 represent the column totals and n is the total number of observations.

$$r1 + r2 = c1 + c2 = n$$

The expected frequency in the 1st row and 1st column i.e. (X1, Y1) is given by (r1) (c1) / n.

Similarly, the expected frequency for (X1, Y2), (X2, Y1) and (X2, Y2) respectively are (r1) (c2) / n, (r2, c1) / n and (r2, c2) / n

The expected frequencies are those which one should expect when the variables are independent.

The null hypothesis (Ho) is that there is no association between the two variables.

Two attributes A & B are independent when $(AB) = \dfrac{(A) - (B)}{n}$

If the observed cell frequencies and expected cell frequencies differ significantly, then we reject the hypothesis of independence.

Example: Data obtained from 180 farmers and 120 non-contact farmers on the sources of their information are given below. Test the association between type of farmer and source of information.

Source of information

		S1	S2	S3	Row total
Type of farmer	C	45	110	25	180
	NC	65	35	20	120
Column Total		'110	145	45	300

Expected cell frequencies

	S1	S2	S3
C	66	87	27
NC	44	58	18

Obs.Fr (O)	Exp.Fr (E)	O - E	$(O - E)^2$	$(O - E)^2 / E$
45	66	-21	441	6.68
110	87	23	529	6.08
25	27	-2	4	0.15
65	44	21	441	10.02
35	58	-23	529	9.12
20	18	2	4	0.22

$\chi^2_{cal} = 32.27$ $\quad S^2_{tab} = 9.2$

$\chi^2_{cal} > \chi^2_{tab}$ ———> reject the hypothesis

i.e. the type of farmer and source of information are associated. In carrying out test the cell frequency should be more than 5 and in case it falls less than 5 then few categories have to be merged

10 4.4 Simple Regression

In statistics, the term regression indicates the average relationship between / among the variables. Variables are classified as dependent variable and independent variables in the context of regression.

Dependent variable: It is that variable whose values are influenced by the values of the other variable (s). This is also known as explained variable or 'effect' or 'response' variable. eg. in the case of a demand equation, demand for a product or service is the dependent variable. (Dependent variable is usually denoted by Y'). The yield of a crop becomes the 'effect' variable when we are dealing with crop production functions.

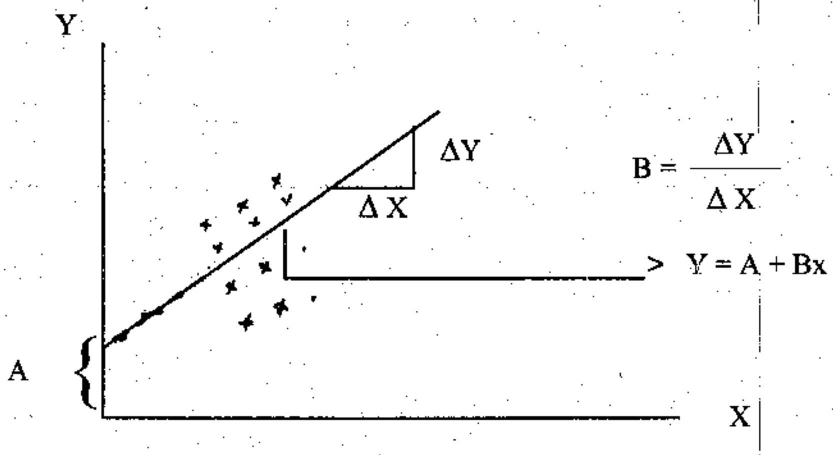
Independent Variable(s): The variable (s) which influence the dependent variable or which causes variations in the dependent variable. It is also known as explanatory variable or predictor variable or 'cause' variable. This is denoted by 'X'.

If there is only one explanatory variable it is called simple regression. Otherwise it is known as multiple regression. A regression can be linear or non-linear in variables depending upon the nature of relation between the dependent and independent variables.

The diagram given below explains the concept of regression. The values of the independent variable 'X' are plotted on 'X' axis and the dependent variable 'Y' values are depicted on the 'Y' axis. The scatter of observations shows the values of the dependent variable for given values of 'X' observed in a population. It is evident from the graph that against each value of the independent variable a number of 'Y' values are observed.

For instance, if 'X' is fertiliser consumption and 'Y' is yield per hectare of a crop, it is expected that farmers using same level of fertiliser may obtain different yield levels because the other cultural practices and soil fertility etc., do vary from farmer to farmer. Find the average value of 'Y' at each level of 'X' and then draw a line / curve connecting all these mean values, to get what is called regression line or curve. The shape of the curve tells us whether the relationship is linear or non-linear.

Fig:



Intercept: It is the expected value of 'Y' when the independent variable attains the value of '0'. For instance, in the above example the average yield of a crop in a particular area when the fertiliser application is '0' is referred to as intercept. Graphically this can be observed by extending the regression line towards 'Y' axis. Value of 'Y' which corresponds to the point of interception by the regression line is the intercept. It is to be noted that in certain situations the estimated value of intercept cannot be interpreted meaningfully. For instance, a negative value of the slope in the consumption function cannot be explained. In the regression equations intercepts are usually represented by 'A'.

Slope: In an ordinary regression, the slope refers to the expected change in the value of the dependent variable for a unit change in the value of independent variable (s) it will be equal to Y/X. The slope is also referred to as marginal value in the field of economics. In the case of consumption function it denotes the marginal propensity to consume i.e. whenever the income goes by one rupee, what proportion of it will be spent for consumption is known through slope. It refers to marginal productivity when we deal with production functions. For instance, when the fertiliser consumption goes up by one kg. what is the expected change in the production / productivity of a crop can be estimated by slope.

Slope is also known as (partial) regression co-efficient. In the interpretation of regression co-efficient one should use the '*ceteris paribus*' condition. In other words, other factors remaining constant, what is the likely change in the dependent variable for one unit change in the independent variable is given by the regression co-efficient. usually slope is denoted by 'B'.

Simple regression model:

$$Y = A + Bx$$

where Y = dependent variable
X = independent variable

A & B : Parameters defined above as intercept and slope / regression co-efficient respectively. Using the data obtained from the sample, the values of A & B will be estimated using the principle of least squares. For the estimation of the parameters the normal equations are used.

$$\Sigma Y = nA + B\ \Sigma x$$

$$\Sigma Yx = A\ \Sigma X + B\ \Sigma X^2$$

by solving these normal equations A & B are estimated.

$$\hat{B} = \frac{\Sigma(X - \bar{X})(Y - \bar{Y})}{\Sigma(X - \bar{X})^2}$$

$$A = Y - \hat{B}\bar{X}$$

Co-efficient of Determination (R^2) : Which denotes the percentage of variation in the dependent variable that is being explained by the independent variable(s). It is also known as explanatory power or predictive power of the model. In the case of simple regression R^2 is nothing but the square of the ordinary correlation co-efficient. The R^2 value lies between '0' and '1'. A value closer to '1' indicates that the model chosen is very appropriate. For instance, $R^2 = 0.6$ means that the independent variable could explain only 60 per cent of the variation in the dependent variable and the remaining 40 per cent of the variation is unexplained by the model.

t-ratio: To know whether the regression co-efficient is significant or not we need some measures which will enable us to say something about the relation between variables in the population at large based on the sample considered. The t-ratios are presented below. The regression coefficients in the regression equation. For large samples (i.e. sample size exceeding 30), a t-ratio value of 1.96 or more indicates that the regression co-efficient is significant i.e. the relation between dependent and independent variable observed in the sample will also hold good in the population and we can conclude the same with 95 per cent confidence. If the t-ratio is very small then it implies that the chosen variable has no significant impact on the dependent variable under study.

Interpretation of the regression equation: When we obtain the results, the following aspects need to be considered in the interpretation of the results.

a) magnitude and sign of the intercept

b) the t-ratio of the regression co-efficient

c) sign and magnitude of the regression co-efficient, and

d) R^2

If the R^2 is very low, it cannot be used for prediction / forecasting purpose. If the t-ratio of the regression co- efficient is low, the chosen independent variable has little influence on the dependent variable. In other words, it can be dropped from the model.

The sign of the regression co-efficient shows whether the dependent variable value increases or decreases whenever the independent variable increases.

If the regression equation is considered in terms of 'log y' and 'log x' instead of 'y' and 'x', the coefficient of log x is called the elasticity. i.e. $\log y = A_1 + B_1 \log x$, and

B_1 is the elasticity i.e. it gives the percentage changes in y for one per cent change in the value of x.

Example: The data on the average number of hours spent per month in library by 10 (sample) students and their overall score are given below. Fit a regression equation and find out what was the effect of using library on the scores obtained by the students.

Y :	80	35	55	60	45	81	75	40	72	60
X :	60	10	25	35	20	55	50	15	45	30a

The proposed regression equation is $Y = A + Bx$

where Y : the overall score
X : average number of hours spent in library

The normal equations are

$$\Sigma Y = nA + B\,\Sigma X \qquad — (1)$$

$$\Sigma YX = A\,\Sigma x + B\,\Sigma X^2 \qquad — (2)$$

substituting the values, we get

$$603 = 10A + 345\,B \qquad — (a)$$

$$23370 = 345A + 14625\,B \qquad — (b)$$

Solving these equations, we get the estimates of A and B

$B = 0.9427$ and $A = 27.78$

So, the estimated regression equation is

$Y = 27.78 + 0.9427\,X \qquad R^2 = 0.9$

Ceterus paribus, students can obtain a minimum of 28 marks on the average if they do not make use of library facility. For every one hour use of library per month the aggregate score is expected to go up by 0.94. The R^2 value of 0.9 is very high which implies that the chosen independent variable could explain 90 per cent of variation in the dependent variable.

10.5 TIME SERIES

Time series a set of values of a variable or characteristic arranged chronologically i.e. it is a series observations or ordered values over successive points in time. In the parlance of regression analysis, when the independent variable is time, and the dependent variable is assumed to vary with time, it is called as 'time series'. In our daily life we come across several instances and the time periods can be hours, days, weeks, months and years e.g. the monthly additions to the stocks of books in a library.

The number of books purchased in a library during the period 1984-94 is given below :

Years:	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
No.of: books	1120	1350	1295	1608	3119	3277	3090	4028	3163	5315	5625

Time series analysis (TSA) deals with systematic analysis of for (a) to identify factors causing the changes in the variable and (b) to forecast the future values (making certain assumptions).

It is evident from the above example that during the reference period the collection of books has gone up considerably and the raising / increasing trend indicates that purchases may go up further in the post 1994 period.

Components of Time series: A time series is a net result of several factors which are responsible for the fluctuations in the variable and these are

i) secular trend
ii) seasonal component
iii) cyclical component and
iv) irregular or random component

Secular trend: It is a smooth long-term movement indicating the overall tendency of the data. It may be an increasing or decreasing or stagnant tendency. The causes of secular trend can be several.

Seasonal component: Variations that occur in the data systematically within a span of one year. The climatic factors, festivals, examinations etc., cause such fluctuations e.g. the borrowing of books in the college libraries increase during the periods of examinations.

Cyclical component: These are variations in data of a longer period, say 4 to 5 years. Several factors are responsible for these long-term periodic movements of time series. The cycles may vary in length and size. e.g the fluctuations in industrial production, share markets etc.

Random component: Concerned with the unexplained variations in the data and factors responsible for these irregular patterns are either unknown or sporadic. The natural and other disasters can be the sources of such aberrations. In time series analysis, these four components will be separated systematically by assuming either an 'additive model' or 'multiplicative model'. In additive model,

$$Y = T + S + C + R$$

and in multiplicative model, $Y = T \times S \times C \times R$

when Y : (dependent) variable
T : secular trend
S : seasonal component
C : cyclical component
R : Random component

In additive model, all the four components are assumed to be independent and in the latter interaction effect is assumed.

a) Measurement of components of Time Series:

The calculation of the components of time series involves tedious calculations. You may refer the books suggested below for the details. A number of computer packages are also available which can be made use of. The following steps are involved in the analysis of time series.

(1) Moving average method: Estimate the trend component using moving average method (MAM). MAM is a method for smoothening the series and a series of averages will be calculated.

Apriori, the researcher has to decide the period. Let it be 3. The average of the first three reservations is worked out $\frac{(y1+y2+y3)}{3}$ and placed at the midpoint of the first three observations i.e. at second observation. The second moving average is calculated taking the 2nd, 3rd and 4th observations. If there are n observations and the period chosen is 3, then we will have n-2 moving averages. So, as the period increases the number of moving averages reduces and information loss increases.

Regression method is also employed to workout the trend component with time as independent variable.

(2) **Isolation of seasonal component:** A number of techniques are available for measuring seasonal component. The 'ratio to trend' method is usually followed.

(3) **Measurement of cyclical component:** A long-term series data is a prerequisite to estimate this component. If the length of data is short, usually it is assumed that cyclical component is absent. The cyclical component is estimated using harmonic analysis or spectral analysis.

(4) **Random component:** After elimination of the three components from the time-series, the residual is known as erratic or random component.

10.6 TESTING OF HYPOTHESIS

'Tests of Significance' are important tools for decision making in both social sciences and natural sciences and forms part of inferential statistics. Decisions will be made about the population parameter based on (sample) value of the sample statistic. Essentially, it deals with some of the following aspects:

i) whether the observed relationship in the sample(s) holds good in the population at large ;

ii) whether the sample drawn comes from a population with specific parameter value ; and

iii) whether the different samples are drawn from the same population or from different populations.

In decision analysis, Tests of Significance will help in providing some degree of accuracy to the decision taken. In other words, the generalisation on the basis of the sample data can be made along with the probability of taking a wrong decision. For better understanding of the concept and mechanics of tests of significance, familiarity with the following terms is necessary.

Population: Includes all the units defined by the researcher having specific region.

e.g. a) All libraries in a state
b) All rice fields in a Taluk
c) SC & ST households in a village etc.

Parameter: The characteristic under study of a population. e.g. a) Percentage of students availing of library facilities b) Average income of the employees of public libraries in a state.

Sample: A part of the population whose basis inferences are made about the population parameter.

Sample statistic: The counter part of the population parameter in the sample.

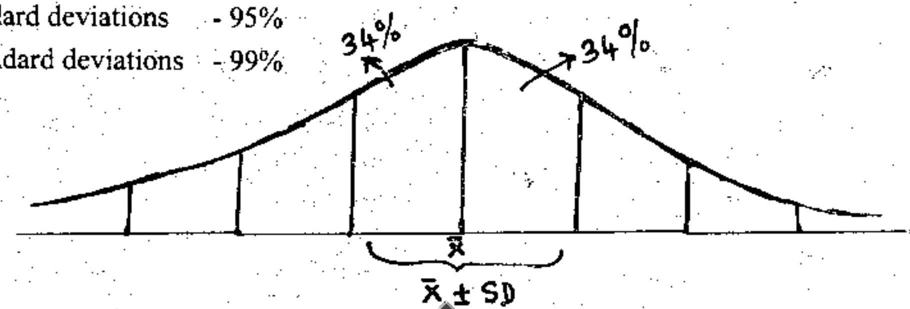
Sampling Distribution: From a (large) population, a number of (random) samples can be drawn and accordingly a large number of sample statistic values can be worked out. The distribution of these sample statistic values is known as sampling distribution. The knowledge about the form of sampling distribution is very important. In fact, 'tests of significance', based on the exact form of sampling distribution, are known as parametric tests. In other words, sampling distribution refers to the distribution of the parameter.

If the sampling distribution follows normal or t or binomial, the properties of those respective distributions can be made use of, in carrying out the tests.

Sampling distribution is normal: Normal distribution is a symmetrical bell shaped curve and all the measures of central tendency coincide with one another. So, the area below the mean is 50 per cent of the total area.

The percentage of observations that lie between
mean ± one standard deviation - 68%
mean ± two standard deviations - 95%
mean ± three standard deviations - 99%

Fig:



In other words, a value of $\bar{X}$ less than ($\bar{X}$ - 2SD) or more than ($\bar{X}$ + 2SD) is less likely to occur and the probability of such occurrence is hardly 5 per cent i.e. a sample, with the statistic value lying outside the two 'S.D.' limits (from mean), to set into our selection process has 5 per cent chance.

The standard normal variate (SNV), obtained by subtracting mean from each observation and dividing the resultant by the standard deviation, also follows

$$\text{normal distribution (SNV)} = \frac{X - \bar{X}}{SD}$$

and the above results hold good for SNV as well.

For large sample (i.e. sample size greater than 30), the sampling distribution for most of the parameters is normal. If it is small sample, (is less than 30), then sampling distribution follows t-distribution and as sample size increases, t-distribution tends to normal. The Tests applicable for small samples are valid for large samples also but not vice-versa.

Standard Error (SE): It is the standard deviation of the sampling distribution i.e.the SD of all the possible statistic values . It is a crucial variable in conducting the tests of significance. It can be estimated with the help of sample statistic value. The smaller the SE greater will be the precision of the estimate (of the parameter).

Hypothesis: In conducting statistical tests,the researcher makes a few assumptions about the population and the sampling procedure. Those assumptions which are more or less valid and those of which researcher is willing to accept constitute the model. The assumptions which the researcher wants to test or verify are called hypotheses.

Hypothesis also comes close to a theory. Hypothesis after repeated trials and verifications will lead to the formation of a theory. It is a statement about the population parameter which needs to be tested. It may be an assumption. Research hypothesis (also known as alternative hypothesis, H_1) is one in which the researcher is primarily interested. Since there are no tests for acceptance of hypothesis, we take a diametrically opposite hypothesis which we would like to reject on the basis of empirical evidence. This hypothesis is called 'Null Hypothesis' (H_o).

For example, research hypothesis can be that regular use of library will lead to improved / better performance in examinations, since, the libraries have been designed and expanded to serve this purpose. The null hypothesis in that case is : There is no significant improvement in the performance of the students availing of library services regularly.

Errors in Decision Making: Two types of errors can creep in the decision-making.

Rejecting the null hypothesis (H_0) when it is true - Type 1 error. Accepting the null hypothesis (H_0) when it is false - Type II error.

The former is also known as level of significance. Statistical tests are constructed to minimise Type II error, once the type I error is fixed.

		H_0 is true	H_1 is true
DECISION	Do not Reject H_0	Right decision	Type II error
	Reject H_0	Type I error	Right Decision

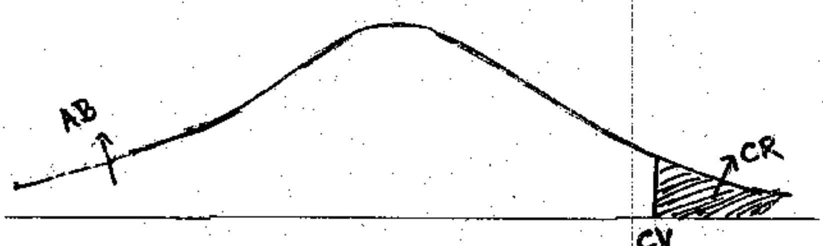
The researcher should specify the value of 'α ' i.e. maximum degree of error of Type I nature he / she is willing to tolerate in decision making.

$\alpha = 0.05$ —> only in 5 per cent of cases, the decision maker is prepared to go wrong. It is also referred to as 95 per cent confidence level, since in 95 per cent cases the researchers would like to take correct decisions.

In general, the 'α ' values will be taken as 0.01, 0.05 and 0.10 and this should be decided before actually carrying out the test.

Critical Value (CV) and Critical Region (CR)

The critical value refers to the table value of the test statistic for a chosen level of significance. The region (s) of the curve beyond the critical values is called as Rejection Region or critical region.



The area of the critical region corresponds to the probability of drawing samples whose sample statistical values are beyond the 2 SD or 3 SD limits. Whenever the test statistic value is more than the critical value (in absolute terms), then it falls in the critical region leading to rejection of the null hypothesis. Thus, the size of the critical region depends upon level of significance (α) / Type I error. The complement to the critical region is acceptance region (AR). So, decision making is done on the basis of the location of test statistic value.

The critical values are given in the normal distribution and t-distribution tables for various 'α' and degrees of freedom.

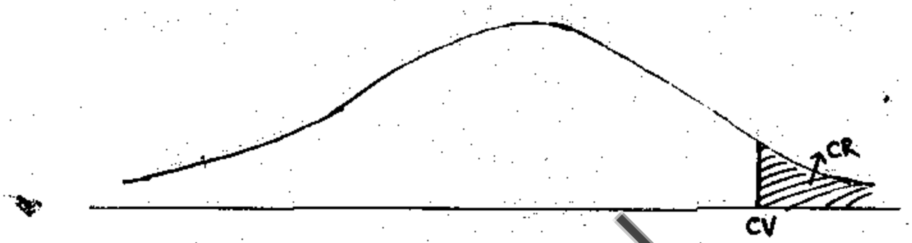
Degrees of freedom: It is a function of sample size (s) and the number of parameters to be estimated. This has to be worked out given the test statistic, number of observations etc.

2-Tailed or 1-Tailed test: The specification of the alternate or research hypothesis (H_1) determines whether it is one-tailed or two-tailed test.

a) **1-Tailed test**

H_o : There is no significant difference in the mean incomes of the two samples $X_1 = X_2$
H_1 : The average income of sample 1 is higher than the average income of the sample 2.

$H_1 : X_1 > X_2$



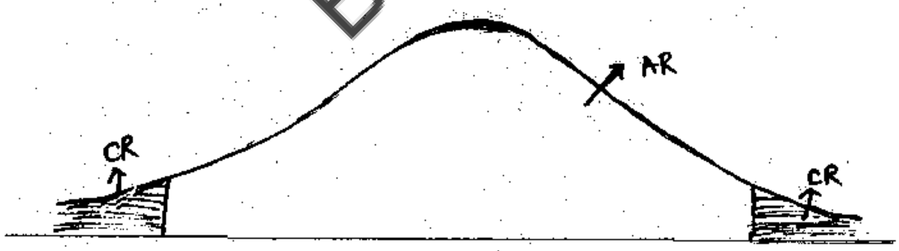
The critical region falls on only one side of the curve of the sampling distribution..

b) **2-Tailed Test :** H_o : There is no significant difference in the mean incomes of the two groups.

$H_o : X_1 = X_2$

$H_1 : X \neq X_2$

H_1 : There is significant difference in the mean incomes.



Both the ends of the distribution constitute the critical region.

Some selected critical values

Level of significance (α)

	0.01	0.05
Large sample ($n > 30$)		
2 - T test	2.58	1.96
1 - T test	2.33	1.64
Small sample ($n < 30$ & $df = 20$)		
2 - T test	2.53	2.09
1 - T test	2.84	1.72

Test - statistic

Tests of significance can be carried out for

a) Parameter *Vs* sample statistic, then the test statistic is

$$\left\{ \frac{\text{Parameter - sample statistic}}{\text{Standard Error}} \right\}$$

b) Sample statistic *Vs* sample statistic, then the test statistic is

$$\left\{ \frac{\text{sample statistic - sample statistic}}{\text{Standard Error}} \right\}$$

NB : If the sample size is small, then t-test will be carried out. If observations are made at two points of time / for two type programme etc., from the sample respondents, paired t-test will used.

If the sample size is large, (i.e. the sampling distribution is normal) both the tests can be used.

Steps in conducting tests of significance

(1) Formulate the research and null hypotheses
(2) Decide whether it is 1-tailed or 2-tailed test
(3) Fix the level of significance (α)
(4) Compute the test-statistic
(5) Find the degrees of freedom
(6) Find out the critical value (from the appropriate tables)
(7) Decision Rule:

(a) If calculated value(absolute value) of test-statistic > critical value (CV) Reject the null hypothesis (H_0) in favour of research hypothesis H_1

(b) If calculated value of test-statistic _ critical value (CV) Do not reject H_0

Exercises:

Paired t-test : This test is applied when observations are made from the same group at two time points.

Exercise: Knowledge scores of a sample of 10 participants who have attended a programme in library sciences at an institute are given below : Test whether the programme has made any significant impact on the knowledge levels of the participants.

X (Pre-)	30	20	25	26	30	25	26	30	35	36
X (Post-)	50	40	30	32	25	35	30	32	40	40

H_0 : No significant difference between final and initial scores
H_1 : There is significant differences between the scores

The test statistic $t_{n-1} = \{ \bar{d} / \sqrt{S^2 / n-1} \}$

where d is the differences in the scores for each participant and S^2 is variance of such differences

d :	20	25	5	6	-5	10	4	2	5	4

$$\bar{d} : 7.1; \; S^2 = \frac{\Sigma (d - \bar{d})^2}{n} = 54.29$$

$$t_9 = \frac{7.1}{\sqrt{54.29/9}} = 2.9$$

C.V. t_9 (0.05) = 2.26 (2-Tailed test & α = 0.05)

$t_{cal} > t_{tab}$ —> Reject H_o

Interpretation: There is significant difference in scores and the training programme benefited the participants significantly.

T-test: Population parameter *Vs* sample statistic

Example : The sample mean (m) and population mean (X) yield of a crop are 35 and 40 quintals respectively.

m = 35 qtls. n = 50 x : 40 qtls. s = 5 qtls.

α = 0.05 H_o : X = m

H_1 : X $\neq$ m

$$\text{Test - statistic } Z = \left| \frac{m-x}{S/\sqrt{n}} \right| = \left| \frac{35-40}{5/\sqrt{50}} \right| = |-7| = 7$$

C.V. (2-T test) (.05) = 1.96

$$\left| \frac{m - x}{SE} \right| \leq CV$$

Reject H_o in favour of H_1

For one tailed test, H_o : m = X, H_1 : m < X

C.V. 2.58 test statistic value > C.V.

Reject H_o

So, the sample mean and population mean are significantly different. In other words, the sample is not a representative of the population whose average yield is 40 quintals.

Sample statistic *Vs* sample statistic

T-Test : Test for mean differences

		Sample 1	Sample 2
Mean income (Rs.)	: X	310	285
S.D. of income (Rs.)	: S	15	18
Sample size	: n	21	25

H_o : No significant difference in the population means
H_1 : Significant difference in the population means

test statistic :

$$t_{n1+n2-2} = \frac{X_1 - X_2}{\sqrt{\dfrac{(n-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}} \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

$$t_{44} = \frac{25}{\sqrt{\dfrac{20 \times 225 + 24 \times 324}{44}} \sqrt{\dfrac{1}{21} + \dfrac{1}{25}}} = 4.9$$

for $\alpha = 0.01$ and 2 T-test, CV = 2.71

So, reject H_o —> the two samples are drawn from two different populations.

10.7 LET US SUM UP

In this unit, we discussed several statistical methods which are useful for data analysis. The variables have been classified into quantitative and qualitative (ordinal and nominal) variables and the appropriate methods to deal with these two types of variables are discussed. The popular measures of central tendency viz., mean, mode and median are discussed and formulae are given to work out the averages for both discrete and continuous cases. Since averages can not explain the total phenomenon, they are supplemented with measures of dispersion.

To deal with bivariate (and multi-variate) cases, a number of measures of association are suggested — the Pearson's correlation coefficient for quantitative variables and chi-square, rank correlation and Yule's coefficient for the qualitative type variables.

To assess the effect of independent variable on the dependent variable, the simple regression method is introduced. If the number of independent variables is more than one, it becomes multiple regression. Using normal equations, estimation of regression equation is demonstrated.

The term 'Time Series' refers to a set of observations arranged in a chronological order. The basic components of time series — secular trend, seasonal component, cyclical component and random component are elaborated and their estimation procedures are mentioned. The cyclical component is usually assumed to be absent if the length of data is short.

Testing of hypothesis is an important aspect in statistical inference. The key terms associated with hypothesis testing are explained. This is followed by a few applications using t-test. The t-test is applied when sample size is small (less than 30).

10.8 ASSIGNMENT

1) Using the data available in any research report or any other publication fit a regression function and explain the steps involved. Also interpret the results.

10.9 REFERENCES AND RECOMMENDED BOOKS

BLALOCK, H. M. *Social statistics*. New York: McGraw-Hill, 1972.

BUSHA, Ch. and Harter, S. P. *Research methods in librarianship: Techniques and Interpretation*. New York: Academic Press, 1988.

GOON, A.V.M.*et al. Basic Statistics*. Calcutta: World Press, 1966.

HENKERL, Ramon E. *Tests of Significance*. New York: Sage publications, 1976. (Series No 4)

HILDEBRAND, David K. *et al. Analysis of ordinal data*. New York: Sage publications, 1977. (Series No.8)

LIEBETRAU, Albert M. *Measures of Association*. New York: Sage publications, 1983. (Series No.2).

OSTROM, Charles W. Jr. *Time series analysis*. New York: Sage publications, 1978.

RAO, I. K. R. *Quantitative Methods for Library and Information Science*. Bombay: Willey Eastern, 1983.

RESEARCH Methods in Library and Information Studies/ ed by M. Slater. London: Library Association Publishing, 1990.

REYNOLDS, H. T. *Analysis of Nominal Data*. New York: Sage Publications, 1977. (Series No.7)

STEPHEN, Peter and Susan Hornby. *Simple Statistics for library and information professionals*. London: Library Association Publishing, 1995.

10.10 MODEL EXAMINATION QUESTIONS

I ESSAY QUESTIONS

1) Mention the steps to be taken while editing and coding the data.

2) What are the measures of central tendency ? Discuss the assumptions and limitations of some selected measures

3) To assess the extent of heterogeneity in a given set of observations what summary measures will you apply ?

4) Explain the concept of association and discuss the measure appropriate in the context of attributes

5) What is regression ? Describe the steps involved in estimating a regression equation.

6) What do you understand by the term 'Time Series' and how the different components are estimated ?

7) What is the difference between a null hypotheses and a research hypotheses? Discuss the errors in decision making ?

II SHORT NOTES

a) Data coding
b) Quantitative variables
c) Frequency distribution
d) Steps in conducting tests of significance

UNIT-11: DATA PROCESSING - USING COMPUTERS

Structure

11.0 AIMS AND OBJECTIVES

This unit aims to introduce you to the data processing methods using computers. It also introduces data preparation for computerised analysis using the statistical packages.

After studying the unit, you should be able to

- discuss the need for using computers in data analysis
- explain the differences between the statistical packages and the statistical languages
- describe the salient features of SPSS package.

11.1 INTRODUCTION

In the previous unit we have learnt how the data is analysed using statistical methods. Data is collected for analysis is coded, stored, retrieved and analysed. Researchers assign numeric codes to their data. The number assigned to an observation is called a code. Coding increases the ability to use computers to store, retrieve and analyse data.

Use of computers in research has increased over the years. Computers are simply tools that help us to store, process, access and analyse data sets more quickly and easily. Comput rs calculate the statistics and provide printouts of the results. However, the researchers have to supply correct and reliable data, choose statistics tools that are appropriate for the level of data, and interpret the results properly.

Traditional methods, which might have taken months of painstaking hand calculation, can now be tackled effectively and instantaneously. The developments in the accessible computer software has led us to use the many statistical tools without completely depending on the statisticians. Researchers without making any deeper study into the theory of the statistics can now use them. Thus, the statistical packages are becoming more and more user-friendly now-a-days.

11.2 DATA PROCESSING - PREPARATION AND ANALYSIS

Data processing is a link between data collection and data analysis. Here the observations are transformed into codes that are amenable to analysis. There are two stages in data processing - data preparation and data analysis.

11.2.1 Data Preparation

At the first stage of data processing, researchers prepare the data so as to make it suitable for analysis. You have to acquaint yourself with some of the methods of preparation of data for coding and codebook construction.

i) Coding and Coding Schemes:

Numerous individual observations are to be classified into a smaller number of meaningful categories to simplify their description. The categories identified from the data are assigned some numbers. These numbers may be purely arbitrary (like, 1 for male and 2 for female as in nominal-level variables) or they may reflect the ranking of ordinal or interval variables (like Income level of respondents are coded as '1' for Rs.1000-5000, '2' for Rs.5001-10000, '3' for Rs.10001-15000, and so on). Thus, the number assigned to an observation is called a Code. Systems of categories used to classify the responses or acts that relate to a single item or variable, are referred to as Coding Schemes.

ii) Rules for Coding:

Let us examine some of the guiding principles involved in constructing the coding schemes.

(1) The responses should be classified into meaningful categories and assign a number to each of the categories.

(2) Code numbers should make intuitive sense for variables that can be rank ordered. In other words, higher scores should be assigned higher code numbers according to their order/level.

For example, when assigning numbers according to age groups a higher age should be associated with a higher grouped code. A 50-year old person should get a higher code than a 25-year old person.

(3) When confirming code numbers to those starting with 0 or 1, each category in the group may be increased by 1 over each category.

(4) Coding decisions depend on several factors, viz., theory, mutual exclusivity, exhaustiveness and detail. Researchers doing quantitative analysis generally test their hypotheses derived from theory. As many issues are of multidimensional nature, each dimension requires a separate category.

(5) The actual categories developed must be mutually exclusive and exhaustive. Each unit of response analysis should clearly fall into one and only one category (mutually exclusive). The coding scheme must be exhaustive - every response must fit into a category (exhaustive).

(6) The enumeration of categories must be sufficient to exhaust all the relevant categories expected of respondents. A few responses not fitting into expected categories may be classified as "other".

(7) The detail of categories in a coding scheme depends on the research question. When in doubt, add another category, but do not aggregate many responses coded to a more general level.

(8) In deductive coding, categories should be linked to the theory from which the research hypothesis was derived. Deductive coding is most common with quantitative research.

(9) Deductive coding often pre-tests the instrument on a small sample of population of interest (Pilot study). Here response categories are directly classified and constructed (close-ended questions) and modified after the pilot study.

(10) Qualitative researchers usually design the coding scheme inductively from the data in their effort to develop grounded theory. The researcher designs the coding scheme on the basis of a representative sample of responses to questions. The responses mentioned most frequently are included in the coding scheme and they are used to analyse the data. "Other category" is included for less frequently mentioned responses.

iii) Codebook Construction:

A Codebook is constructed with codes assigned for each of the variables used in the questionnaire of the research project. They contain information regarding each variable's name and number, the coding scheme and coding for missing data. A brief form of the actual survey questions is also included.

The Codebook serves as a guide for the coders who will translate the raw data and feed that onto a spreadsheet or input device for later use in computerised statistical analysis.

A Codebook format contains each variable identified by its name, brief question content, coding scheme employed (values), column numbers, etc. (A model of Codebook format is given below).

Variable Name	Questions in brief form Code Numbers(001-500) Max.999	Column Nos (01-80)
ID NUMBER	Code Actual Number of Respondent	1-3
Q1 (Questions)	Gender (1=male; 2=female)	4
Q2	Educational qualification (1=UG; 2=Graduate; 3=Postgraduate; 4=Doctorate; 5=Others)	5
Q3	Occupation (1=Student; 2=Research Scholar; 3=Lecturer; 4=Reader; 5=Professor; 6=Non-teaching staff; 7=Others)	6
Q4	Salary Grade (1=<Rs.5000; 2=5001-10000; 3=10001-20000; 4=20001-30000; 5=>30000)	7
Q5	(1=Strongly agree; 2=Agree; 3=Somewhat agree; 4=Somewhat disagree; 5=Disagree 6=Disagree strongly; 9=DK (Don't Know/ NA Not Applicable)	8
Q6	(1=Excellent; 2=Good; 3=Fair; 4=Poor; 9=DK/NA)	9

iv) Data Coding:

Researchers use the Codingbook and the coding schemes to translate data into a format that allows computer processing to get desired results. First, the coded data need to be transferred to a form acceptable by the computer programmes for storage and analysis. The researcher himself/herself or his/her research assistant translates each response on the questionnaire into a code. Raw data can be coded in a number of ways. However, to increase the coding reliability, researchers need to keep the coding schemes as simple as possible. It is a standard practice to verify a sample of each coding work to ensure reliability. Frankfort-Nachmias and David Nachmias (1996) have given General Coding Instructions, which are summarised below.

GENERAL CODING INSTRUCTIONS :

A. Coding must be done in red *pencil.*

B. Never erase any interviewer-circled codes or comments. If the questionnaire must be corrected, draw a line through the code circled in error. Do not make it impossible to read what was done originally. Please note that green marks are field department corrections.

C. Every column must have a code, and no column may contain more than one clearly circled code.

D. For those questions noted *FLAG*: this instruction applies to questions where special coding problems have been anticipated. Coders will be provided with a supply of little clips/

flags to attach to the page where the problem occurs. Coders may also flag all other areas in the questionnaire where information is incomplete or unclear and requires the supervisor's attention.

E. For those questions noted *LIST*: record on an "OTHER" list form the questionnaire identification and all verbatim comments relating to the response to be listed. A separate list should be kept for each question. If, however, a question has more than one listed code, keep a separate list for each code. Record at the top of each form the study number, the question number, the deck and column number, and the listed code. In most surveys, all "Other (Specify)" codes are listed.

F. The "no answer" (NA) and "refusal" code for this questionnaire is '9' in a one-column field, '99' in a two-column field, etc. NA is coded when the respondent does not give an answer, when the interviewer fails to ask a question or to record the answer, when the written information is contradictory or too vague to code, and when the coder needs to supply a code in order to resolve a tricky skip pattern. NA is allowed for every question except those specifically excepted in the codebook, such as race and sex.

G. The "not applicable" (NAP) code is "R", which means "reject" or "blank" to our keypunchers. NAP is coded when a question was not supposed to have been asked (i.e., because of directions to skip it)

H. If "don't know" (DK) is not a pre-printed code, then DK is coded '8' in a one-column field, '98' in a two-column field, '998' in a three-column field, etc. If DK has been listed along with other responses in one question, edit out (or do not code) the DK response.

v) Coding and Entry Devices:

A variety of coding devices are used to organise the raw data. These include transfer sheets, edge coding, optical scanning, and direct data entry.

(1) *Transfer sheets*: Keypunch cards were used in the 1980s. Later these were replaced by transfer sheets. The transfer sheets are used to record data in columns specified by the codebook. Now coders use spreadsheet forms to organise responses in rows and values of the variables across the columns.

(2) *Edge coding*: Edge coding eliminates keypunch cards and transfer sheets. The coders transfer questionnaire responses directly into spaces at the outside edge of the instrument. The information is keyed directly into the input devices.

(3) *Optical scanning*: These are like computer-graded multiple- choice examination answer sheets. Scanning machines read black pencil marks and produce data files automatically. This eliminates keying errors associated with manual data entry.

(4) *Direct data entry*: The responses from the questionnaires are entered directly into the computer. However, the data must be edited to ensure missing responses also have a designated code for the input.

vi) Data Editing and Cleaning:

Data Editing involves checking data for errors and omissions. This is done both during and after coding phase. In large research projects, where the services of project assistants are hired, the research supervisor reviews the responses and checks the coding. Corrections are carried out before the data is sent for analysis.

Data Cleaning is the proof reading of the data to catch and correct errors and inconsistent codes. The computer spread sheets are designed to test for logical consistency set up in the coding specifications. Illegitimate and wild codes are checked during data cleaning.

11.2.2 Data Processing *vis-a-vis* Analysis

Processing of data refers to concentrating, recasting and deling with data so that they are as amenable to analysis as possible. In the preceding sections we discussed what is precisely called as 'Data processing'.

Analysis of data involves a number of closely-related operations that are performed with the purpose of summarizing the collected data and organizing these in such a manner that they will yield answers to the research question. This also refers to verifying the data in relation to the suggested hypothesis and prevailing theories. Data analysis help us in drawing conclusions that are as amenable to theory formation. Statistical methods help us in many ways in data analysis.

11.2.3 Use of Statistics in Data Analysis

Statistics, as a scientific discipline, is evolving rapidly. Statistical methods are more widely used in any discipline. The developments in electronic data processing using more sophisticated computer systems have also provided impetus for the invention of entirely new statistical methods which are otherwise impracticable or inconceivable with the manual use of statistical techniques. As a result of new kind of statistical tools that are coming up in recent times deeper and more fundamental changes are likely to occur in traditional systems.

Lunneborg and Hand (1996) have given some examples of such new methods which are coming up in recent times. They are

1) Resampling methods, such as jackknife, bootstrap and cross-validation methods. Such tools repeatedly analyse resamples drawn from the original sample, and so get an idea of the variability which results from sampling.

2) Non-linear methods. These are often analytically intractable and need rapid optimisation methods to estimate the parameters of the models.

3) Stochastic optimisation methods permit global optima to be found, even in the presence of many local optima. Simulation allows one to explore the properties of estimators or models which cannot be solved using analytic methods.

4) Recursive partitioning (tree) algorithms are becoming popular for classification and prediction problems. These are very different from traditional methods based on distribution theory.

5) Non-parametric smoothing and curve estimation methods, such as spline and kernel approaches, are becoming increasingly important.

6) New kinds of statistical models, such as graphical models, based on a combination of graph theory and distribution theory, and neural networks, effectively a flexible non-linear model, are being developed.

11.3 USE OF COMPUTERS IN DATA ANALYSIS

Social scientists have been using computers to organise the research data and process it for many decades. As we know that computers are tools that help us to store, process, analyse and access data sets more easily and quickly. With the expanding powers of processing computers are becoming more useful to calculate the statistics and provide the results almost instantaneously. For this computers need suitable software compatible with the system and also to serve the user's purpose.

11.3.1 Computer Software for Data Analysis

Several manufacturers have developed and marketed software for various purposes. The software may be categorised as below.

Types of Software:

i) **Word Processing:**

Write, Works, MS Word, Word Perfect, AmiPro, WordStar

ii) **Text layout and Desk Top Publishing:**

Ventura, PageMaker, Timeworks

iii) **Spreadsheet:**

Excel, Lotus 1-2-3, Quattro, Supercalc, Works

iv) **Presentation graphics:**

Harvard Graphics, CorelDraw

v) **Data collection:**

EpiInfo

vi) **Statistical analysis:**

SPSS, SAS, Stata, Splus, Minitab

vii) **Database programming:**

Access, Clipper, Paradox

viii) **Bibliographic database:**

End Note, Papyrus, Reference Manager

ix) **Networking:**

Kermit, Procomm, Telcon

x) **System maintenance:**

Norton Utilities, PC Tools, OEMM, Stacker

11.3.2 Statistical Packages and Statistical Languages

There are two types of statistical software: Statistical packages and Statistical languages.

Statistical Packages: Statistical packages contain a group of programmes, collected together so that they can be used from a common command language. They have provision for constructing a sequence of commands according to some built-in template. They are aimed more at the potential researcher from another discipline than at a professional statistician. Examples: BMDP, SPSS.

Statistical Languages: Statistical languages are more like programming languages. They have simple commands for basic statistical operations. They usually requires user's intervention to put the commands/instructions together to undertake more statistical analyses. These are more flexible than statistical packages. They also often have an emphasis on graphics. Examples: SPIDA, GAUSS.

The differences between the statistical packages and statistical languages can be examined in the following tabular form:

Statistical Packages	Statistical Languages
1) The statistical packages are evolved from the statistical systems in the days of punched cards.	The statistical languages are evolved from general purpose programming languages (such as FORTRAN, Pascal, C, etc) hence these look like programming languages.
2) These consist of suites of programs. These can be used from a command language.	These consist of simple commands and requires user's operation to put the commands for statistical analyses. Allows more user interactivity.
3) Specific sequence of commands to be followed and hence these are somewhat fixed.	The statistical languages have simple instructions and hence these are more flexible.
4) These are well suited to structured forms of analysis such as multivariate analysis of variance.	New statistical algorithms can be assembled quickly from the existing languages. These also provide local expansibility through macro-commands.
5) These may or may not have emphasis on graphics.	These have often emphasis on graphics.
6) These are aimed more at a potential researcher.	These are used by professional statisticians and researchers any discipline.

Often the distinction between the statistical packages and the statistical languages are not clear. Some packages permit internal programming languages or an interface to FORTRAN commands. Likewise the statistical languages have attributed a considerable user base and these have built up libraries of macros for performing sophisticated and structured analysis.

11.4 SOFTWARE FOR STATISTICAL ANALYSIS

You have learnt that computer software used in data analysis is available in two modes - statistical packages and statistical programming languages. In this section let us study some examples of statistical software used in research data analysis and presentation.

11.4.1 Statistical Packages

Though most of the statistical packages have their origin in batch mode, now they have been moving towards interactive mode.

(1) BMDP

BMDP is a suite of various statistical programmes. The programmes have uniform format. Each programme in the suite groups together commands into paragraphs comprising sentences. It has its own syntax. Each programme is called and run separately as per the requirement.

For example, BMDP2V is a programme, which performs analysis of variance and covariance with repeated measures. It consists of paragraphs like /INPUT, /VARIABLE, /GROUP, /DESIGN, /PRINT and /END. Let us examine the following sample programme:

```
/INPUT FILE = '\bmdp\eg5.dat'
 VARIABLES = 6.
 FORMAT = FREE.

/VARIABLE NAMES = GROUP, TIME0, TIME1, TIME2, TIME4, TIME6.

/GROUP CODES(GROUP) = 1,2.
 NAMES(GROUP) = PLACEB, LECITH.

/DESIGN GROUPING = GROUP.
  DEPENDENT = TIME0 TO TIME6.
  LEVEL = 5.  NAME = TIME.
  ORTHOGONAL.
  POINT(1) = 0,1,2,4,6.

/PRINT LINESIZE = 80.

/END
```

The /INPUT paragraph specifies the file containing the data and describes the number of variables and the format of data. The /VARIABLE command names the variables. The / GROUP command describes the codes used to describe the groups in the data and gives the groups' names. The /DESIGN paragraph represents the core of the analysis, specifying the dependent variables.

The package covers a wide range of analysis, including simple data description, analysis of variance and covariance, multiway frequency tables, log-linear models, factor analysis, step-wise discriminant analysis and regression analysis, Box-Jenkins times series analysis, life tables and survival functions, cluster analysis, non-linear regression, unbalanced repeated measures models with structured covariance matrices, correspondence analysis and many others.

(2) SPSS (Statistical Package for the Social Sciences)

SPSS was designed especially for the analysis of social sciences data. It is a package of programmes covering a wide range of analyses. They include frequency distribution, univariate and multivariate analysis of variance, loglinear models, correlation analysis, regression analysis, factor analysis, cluster analysis, reliability analysis and so on. A particular programme is invoked by specifying its name. The sub-commands direct precisely what analysis is to be performed. For example, the GROUPS sub-command specifies which groups are to be discriminated between and the VARIABLES sub-command specifies which variables are to be used.

The package comes in different versions - mainframe, PC and Windows. According to the computer environment at the user's end the package is to be purchased. Internet Discussion Group for SPSS Users has been formed to interact and share each other's experiences.

(3) IDAMS PACKAGE

IDAMS stands for Internationally Developed Data Analysis and Management Software. The package is distributed free-of-charge by the Unesco Secretariat upon requests from institutions. It can also be obtained from 55 official distributors (41 national and 14 specialised).

IDAMS has been developed for professionals who, in scientific and/or administrative environments, have to analyse numerical data and handle numerical data bases. Users will find in IDAMS the classical range of univariate and bivariate statistics, and multivariate statistical techniques, plus some recent advanced techniques. IDAMS also provides a series of data management programmes, and a powerful language for transformation of data. From the release 3.0 (1992) for microcomputers it is equipped with performant editors, graphic facilities and online help messages. The software can be used without programming skills or knowledge.

IDAMS was designed for IBM compatible mainframe (IDAMS-MF) and microcomputers (micro-IDAMS). IDAMS-MF runs on IBM 360/370 or equivalent with 150 K fixed or virtual memory, 12 mb disk space and on OS operating system. micro-IDAMS runs on PC/ XT, AT, PS/ 2 or compatible computer, 640K memory, 3mb hard disk space(minimum) and MS-DOS or PC-DOS ver 2.1 or above.

IDAMS allows downloading of numeric data from CDS/ISIS to IDAMS, uploading results to CDS/ISIS, easy retrieving of information for management, calculations/statistics as required. It also provides facilities to build standard pre-defined analysis procedures for a given application. The statistical techniques of the package are used for various analyses - univariate and bivariate distributions and statistics, non-parametric tests, Pearson correlations, analysis of variance, factor analysis (including correspondence analysis), multidimensional scaling, partial order scoring, rank-ordering of alternatives, multiple linear regression, etc.

The Internet e-mail address: idams@unesco.org

Web address: http://www.unesco.org/webworld/idams/

(4) MICROSTAT

Microstat is an application software exclusively for statistical analysis. It covers most of the statistical procedures that are frequently used in academic research. The options are accessed through a menu. It works on IBM compatible PCs either with hard disk or with two floppy drives. It needs a minimum of 512k RAM.

The options in the main-menu are listed in alphabetical order (A to P). They are

A Data Management System
B Descriptive Statistics
C Frequency Distribution
D Hypotheses Tests: Mean
E Analysis of Variance
F Scatterplot
G Correlation Matrix
H Regression Analysis
I Time Series Analysis
J Non-Parametric Statistics
K Crosstab/ Chi-Square Tests
L Permutations/Combinations
M Probability Distribution
N Hypothesis Tests: Proportion
O [Installation]
P [Terminate]

Each option is Querry-Answer type. By pressing alphabet shown against the option procedures are invoked. There are sub-commands for the options. For example, different types of probability distribution can be calculated through the sub-menus:

A Binomial;
B Hypogeometric;
C Poisson;
D Exponential;
E Normal ;
F Distribution;
G Students;
H Chi-square
I [Terminate]

(5) Electronic Spreadsheets : Lotus 1-2-3 and MS Excel

Electronic Spreadsheet is a computer programme that enables the user to setup a display of rows and columns in which some entries are manually entered and others are calculated automatically using formulas supplied by the user.

Spreadsheets are valuable for displaying information. They transform the PC screen into rows and columns and allows the user to work with a series of interrelated values. When a user changes one value on an electronic spreadsheet the programme automatically recalculates any other affected values it reads.

Although spreadsheets are originally intended for accountants, these programmes have become quite popular with management analyses, numerical and financial data. There are a number of spreadsheet programmes, for example, VisiCalc, SuperCalc, Lotus 1-2-3 and Microsoft Excel which are very popular.

i) Lotus 1-2-3:

The 1-2-3 spreadsheet was developed from Lotus Development Corp. It combines the good features from several other spreadsheets along with the most extensive use of menus and help to offer great power and convenience. The 1-2-3 software offers an enormous 255 x 2048 matrix that allows to create sheets with elaborate levels in detail.

ii) Microsoft Excel:

MS Excel is one of the powerful software for electronic spreadsheets. It has very advanced features in spreadsheet and data analysis for statistical and engineering purposes. This data analysis facility is called 'Analysis Tool Pack'.

In Analysis Tool Pack, the following data analysis options are available:

1 ANOVA - Single Factor
2 ANOVA - Two factor with Replication
3 ANOVA - Two factor without Replication
4 Correlation
5 Covariance
6 Descriptive Statistics
7 Exponential Smoothening
8 F - Test for Two samples for Variances
9 Histogram
10 Moving Averages
11 Random Number Generation
12 Rank and Percentile
13 Regression
14 t-test paired two sample means
15 t-test two samples assuming equal variances
16 t-test two samples assuming unequal variances
17 Z-test two sample for means
18 Fourier Analysis
19 Sampling

MS-Excel allows import/export facility from external/remote data sources. It analyses data from a list or database. We can cross-tabulate or summarise data and then apply various mathematical functions to the summary. The software can automatically translate formulas from Lotus 1-2-3 file.

MS-Excel allows to create a compound document. It transfers tables and charts from MS-Excel into MS-Word document or vice-versa. Graphic objects can be created using MS-Excel software.

11.4.2 Statistical Programming Languages

As noted earlier the statistical languages are like programming languages. They are designed to meet the need for doing non-standard analysis. They are also helpful in developing new statistical techniques. For this purpose statistical computing platforms have been developed from the general purpose programming languages, such as FORTRAN, Pascal, and C. These programming languages recognise the expressions given as matrix and do multiplications as required. Modern programming languages operate in computing environments have three important features. They are

i) *Interactivity*: The requested computations are done instantaneously. The results are available for inspection and further computations. The statistician-programmer interacts with the data, making choices as the analysis proceeds.

ii) *Graphical capability*: As the interaction between statistician and data occurs through a reasonably high-definition computer screen, data can also be described graphically as well as numerically in the course of analysis.

iii) *Local expansibility*: Using the commands from the computing language, users may write programmes of their own and such programmes are called 'Macro-commands', user function, or external procedure. The macro-commands can be executed repeatedly exactly as if it were one of the programme's more basic commands.

Examples of Statistical Programming Languages:

The most widely used statistical programming languages on statistical computing platforms are:

1) Splus (Bell Labs, USA: Statsci Europe, Oxford)
2) GAUSS (Aptech Systems Inc., Maple valey, USA)
3) SPIDA (MacQuarie University, Sydney, Australia)
4) SC Statistical Calculator (SC Mole Software, N Ireland)

(1) **Splus** : S or New S is the commercial version of the Splus language, developed by and for scientists at the Bell Laboratories in the USA. It is available for Windows as well as Unix platforms. Its built-in algorithms are made object-oriented and solve different statistical tasks. There are also provisions for down-loading of the external procedures. Its features are also suitable for graphics.

(2) **GAUSS** : It is a statistical programming language for either Unix or DOS platform. It has a set of highly optimised linear algebra and related matrix handling routines. Add-on external routines allow computations such as linear regression, curve fitting and time series models. It is easy to learn. Internet discussion group for GAUSS users are also formed.

(3) **SPIDA** (Statistical Package for Interactive Data Analysis) : SPIDA is available for DOS and UNIX computers. It includes both basic and applied commands. An extensive manual is provided for programming of SPIDA macro-commands.

(4) SC (Statistical Calculator): SC provides a large number or procedures and functions, both built-in and external. Each of these procedures helps in carrying out the required statistical task by the user. SC is an ideal language where new forms of analysis are to be tried. It is a flexible language.

11.5 *SPSS* PACKAGE AND ITS USE IN STATISTICAL ANALYSIS

The SPSS (Release-4) is one of the most popular and widely used software packages by the social scientists for computerised data analysis. The package operates in three main environments - mainframe, on a PC and in Windows. There are two versions - Professional (or full service) programme and Studentware version. Both are available for the PC (SPSS/PC+) and Windows (SPSS/Win) environments. Student version tends to provide fewer subprogrammes and allow for a smaller data matrix and fewer options.

11.5.1 Preparing Data for Analysis

The data for analysis can be prepared and stored on tape or diskette or in hard disc depending on the computer system. There are two formats for feeding the data - from a standard textfile or from a spread sheet.

(1) **Standard Textfiles**: Data entered from a standard textfile can be imported from a word processed or edited file or entered directly into the runstream using the SPSS internal editor. In SPSS (PC ver,) we can use the *Review Editor* to build the data sets. The easiest way to format data is to use a *fixed format*. In the simplest example of fixed format data entry, the information for each case is contained on one line or row of text. SPSS refers to this line as a *record*. In fixed format, each variable's value for every case is contained in the same column locations. When the data for a case includes more information than one record, the second record may be taken with a new variable (for example, ID2 for the second record, ID3 for the third record and so on), so that all records associated with each case are together in order (case-1:record-1; case-2:record-2 and so on to case-n:record-n).

(2) **Spreadsheet Data**: SPSS has adapted to the widespread use of spreadsheets by allowing them to be input into the setup runstream. The data have to be organised so that each row contains a particular variable's value. For example, Column-A would contain each case's ID, while Column-B would contain each case's age and so on. Generally, variable names are placed in row-1 and the data for the first case in the row-2 onwards. SPSS will read and assign both the variable's names and their associated values for each case.

11.5.2 Setting up on SPSS System File

The language of SPSS is logical and quite simple. The software contains a number of files for its operation.

(1) *Command Files* contain the commands to execute in any particular run. These commands are known as the *runstream.*

(2) *Data Files* files contain the raw data. They are stored in any readable form, but generally coded in 80-column form.

(3) *Listing Files* contain the results or output associated with executed procedures that are stored in a command file. Output is viewed on screen or printed.

(4) *SPSS System Files* contain the information about the data from the input file defined uniquely by the series of commands executed in the command file used to set up the system file. The system file contains a *dictionary* that defines variables and also contains information concerning lables, missing values, etc. when a system file is in use, it is called an *active file.*

11.5.3 SPSS System File Set up Commands

Before running the software, the system file set up commands are to invoked. They contain

(1) *Title* : This is an optional feature that allows to lable the run with a title of upto 60 characters. The title is printed on the top of each page of the output.

(2) *Data List*: This command describes the variable names, location, type of data, and number of records associated with the input file.

(3) *Missing values, Variable lables and Value lables*: This is also an optional command to describe variables and their values further.

11.5.4 Running SPSS

After input data has been read into the SPSS system file, it determines that the data has been fed exactly and if any stray or illegal data codes are keyed in, they have to be corrected. During the initial run, SPSS displays the information contained in the data list lines in a table. The variables and their values in the columns could be further checked and corrected before proceeding to insert the required procedures for cleaning, manipulation and output.

The procedure commands enable the system to determine the extent to which cleaning of the data to be done.

FREQUENCIES Procedure checks whether there are any illegal responses for each of the variables.

SELECT IF and LIST commands are used to identify and list the errant cases.

SAVE command incorporates modifications or permanent changes to the system files

11.5.5 Use of SPSS Package in Data Analysis

The SPSS package is used in the data analysis to get various statistics.

(1) SPSS package allows generate summary of statistics including mean, standard error, median, mode, standard deviation, variance, kurtosis, skewness, range, minimum and maximum.

(2) The Graphic portion of the package allows to get any graphic presentation from the values derived from the data analysis. The Windows version has wider range of choices in graphics modules.

(3) The package allows univariate distributions as well as bivariate distributions. Bivariate analysis allows to examine the association between two variables. The CROSSTABS procedure allows to generate bivariate tables and statistics for nominal and ordinal measures. Interval measures can be obtained from the CORRELATIONS subprogram. The three major functions of multivariate analysis - control, interpretation and prediction (ie regression/ multiple regression) are also achieved through equations and subprograms.

SPSS for Windows (Version 6) is a well integrated windows program that adheres to many typical windows conventions and commands, hence it is easy to use. File, Edit, Help, New, Open, Save, Exit, etc are the typical windows commands used in the package. Preferences dialog box, Define variables dialog box, Utilities - Variables box, Data editor window, etc allow the user to control settings, select the menu, manipulate the data with required procedures and get the output easily. *SPSS User's Guide* provides information on the sub-programmes and their use.

11.6 COMPUTER SOFTWARE FOR PRESENTATION

After data analysis using statistical software the research results need to be presented in an effective way. Some of the software packages are specifically designed for this purpose. Researchers can make use of them irrespective of their data analysis methods.

11.6.1 Harvard Graphics (HG)

Harvard Graphics is a high-end chart production application programme developed by Software Publishing Corporation. It is DOS-based and menu-driven programme. It is an intelligent and easy-to-use application package with many features.

Some of the important features of HG include -

i) It can create presentation quality charts and then print, plot or produce slides of them.

ii) It has the capability to add drawings and annotations to charts.

iii) It allows importing and exporting of charts and data.

iv) It can build computerised presentations and create customised symbols.

v) It can graphically display numerical relations, representation of symbols through lines, markers, pie slices, bars and areas.

vi) It has options for tailoring the chart to specific needs and it allows to modify some options in order to customise the chart's appearance.

Users can select a variety of chart styles from the three categories - text charts, organisation charts and data charts. The main menu contains a vertical list of choices, like Create New Chart, Enter/Edit chart, Draw/Annotate, Get/Save/Remove, Import/Export, Produce output, Slide show menu, etc. When an option, for example Create a New Chart, is selected, then appears a second level menu - Text, Pie, Bar/Line, Area, Multiple charts, etc. Second level menu in turn produces another menu. The Text chart styles menu leads to Title chart, Simple list, Bullet list, etc.

11.6.2 MS Powerpoint

Microsoft Powerpoint is basically a Window-based presentation software. The software allows (a) easily create an effective presentation, (b) On-screen electronic presentation, (c) Preparation of 35 mm slides, (d) Printing of handouts, (d) Easy to add graphs, tables, organisation charts, clip art, special text effects using wordart. These features can be used in presenting the research results from the data analysis.

11.7 LET US SUM UP

Let us recapitulate briefly what has been discussed so far in this unit.

* Social scientists have been using computers to organise the research data and process it for many decades.

* Computers are tools that help us store, access, process and analyse data sets more easily and quickly.

* The processing power of computers are increasing and they are becoming more useful to calculate the statistics and provide the results almost instantaneously.

* Computers need suitable software compatible with the system and also to serve the user's purpose.

* The software used in data analysis can be categorised into statistical packages (eg. SPSS) and statistical languages.

* There are a number of software available for processing of research data (eg. IDAMS, Microstat, Spreadsheets, etc) besides presentation software (eg. Harvard Graphics).

11.8 ASSIGNMENTS

Survey and list out the software packages available in India, which can be used in statistical analysis of research data. Evaluate their use in LIS research.

11.9 REFERENCES AND RECOMMENDED BOOKS

ALRECK, Pamela L. and Robert B Settle. *The survey research handbook*. Chicago: Irwin, 1995. p,236-263.

FRANKFORT-NACHMIAS, Chava and David Nachmias. *Research methods in social sciences*, 5th ed. London: Arnold, 1996. p.503-545.

JENSEN, Cary and Loy Anderson. *Harvard Graphics: the complete reference*. Berkeley: Osborne McGraw-Hills, 1990.

LUNNEBORG, C E and David Hand. "Data analysis by computer". (IN *Research methods: guidance for postgraduates*/edited by Tony Greenfield. London: Arnold, 1996). p.195-203.

11.10 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) Explain the need for using computers in data processing and analysis.

2) What are statistical packages and statistical languages ? Bring out the differences between them.

3) Describe the salient features of SPSS package.

II. SHORT NOTES

a) Spreadsheet

b) Codebook construction

c) Data editing and cleaning

d) IDAMS package

UNIT-12: PRESENTATION AND INTERPRETATION OF RESEARCH RESULTS

Structure

12.1 AIMS AND OBJECTIVES

The unit aims to introduce you to the methods of presenting and interpreting research results.

After going through this unit you will be able

- to present the results in a more pleasing and effective manner with the aid of charts and other geometrical devices; and
- to take the necessary precautions in interpreting the results of the study.
- to prepare tables effectively for presentation.

12.2 INTRODUCTION

Presentation and interpretation of the findings of the study are very important since in most of the cases the users of the research study are non-professionals. So, it is rather the duty of the researcher to present the information in a cogent manner and all the inferences relating to the issues / objectives of the study and the tested results of selected hypotheses are reported in an unambiguous manner.

For this, the tabulation and graphic mode of presentation of results have been found to be handy and also attractive. While conducting the study a number of problems may be encountered at various stages and the reader should be cautioned by the researcher about the likely impact of the same on the results while presenting and interpreting the results. It is binding on the part of the researcher to educate the reader regarding the magnitude of non-response problems and also the factors which are beyond the control of investigators and having some bearing on the research outcomes.

The validity and also the acceptability of the results will increase when the researcher is not only objective in carrying out the research study but also in sharing the information with the user. Keeping these in view, the various aspects dealing with presentation and interpretation of the data are discussed in the next few sections.

So, it is imperative that the various findings of the study should be succinctly summarised and the essence of the various pieces of information / results should be presented for the benefit of the user. Mention may be made in the report if some of the biases provided out have creeped in and their likely effect on the findings

12.2.1 Textual Presentation

Textual presentation is considered as an informal presentation of data. There is no defined set of rules for the elementary forms of presentation. It is usually made as a part of the running text i.e. data are reported as a part of the text. This method should not be used when huge data are to be reported.

"Taking 1977 cyclone as a benchmark during the period of 1977-78 to 1991-92, nearly 13,116 human lives-lost, (with maximum number 9,840 in a single year 1977), 338,611 cattle perished (excluding sheep, goat and poultry birds) and nearly 3,318,716 houses were damaged. The total value of losses were estimated at about Rs.19,451.5 million of which Rs.2,515.3 million was accounted by houses, Rs.7222.9 million by crops, and Rs.9,713.3 million by infrastructure ......"

Semi-tabular presentation consists of reporting data in the text discussion by either changing the type (size) or by indenting. As per the Rural Labour Enquiry Report, the details regarding the conditions of rural labour households for the year 1987-88 are as follows :

Some Important Findings of Rural labour Enquiry : 1987-88

S.No.	Item	Value
1.	Estimated No.of Rural households (M)	108,42
2.	Percentage of Agricultural labour households	30.70
3.	Percentage of all Rural labour households to Rural households	39.70

Compared to textual presentation, the semi-tabular mode is better in the sense that it reports all the figures or statistics at one place and facilitates easier comparisons.

12.2.2 Tables

A systematic organisation or classification of data into rows (horizontal arrangements) or columns (vertical arrangements) is an efficient mode of presentation of complex information in the simplest possible form. The reader can easily grasp the essence of the data and also the trends or patterns. The user can quickly locate the desired information. The utility of tabular presentation is that the researcher can drive the point home by this systematic arrangement and also facilitate easy comparison.

Tables are of two types :

(a) general purpose tables or reference tables ; and

(b) special purpose or text tables.

The tables included in the population census reports are good examples of the former type. They contain large volume of data which can be processed further to draw meaningful conclusions. The tables inserted in the main text of the research reports can be classified as text tables. These tables help answering one or a few issues under consideration or provide important information on the selected dimension (s) of the problem.

You will realise that construction of a good table is not an an easy task and requires skills. A good statistical table should be complete and be self-explanatory. The following are some of the key parts / components of a table.

Table Number: Is useful for identification and future reference.

Title: Should be placed above the table and it should be an unambiguous statement about the nature of the data furnished in the table. If the title of the table is too lengthy it can be split into a main title and a sub-title. A good title should indicate the nature and scope of the data and also help us in identifying which aspect (s) of the study are examined in that particular table.

Captions: Refer to the columns and there will be a main heading indicating the type of information contained in the various columns. This is followed by a sub-heading for each column. The main caption may be the cropping pattern and the sub-heading may be the name of specific crop or area under each crop.

Stubs: Referred to rows and the sub-entries describe the contents of the rows. For eg. the various farm size classes like marginal, medium, small and large farmers may be included in the rows.

Body: Quantitative information will be presented against each row and column in the table.

Footnote: Symbols like *, + and @ will be used in the body of the table and also in the stubs and captions whenever there are deviations or discrepancies. This will be explained in the footnotes along with the symbols. For instance * may indicate the information relating to small farmers relate to the year 1991-92 while for the rest, the data relate to the year 1993-94.

Source: When data are compiled from the field observations there is no necessity to give the source. However, when it is taken from another document, published or unpublished, it is the responsibility of the researcher to give the name of the author (s), title of the work, the volume number, the publisher's name, year of publication and page number. This customary practice will help the user to cross-check, if necessary, with the original source.

Units: Units are usually placed below title and above the columns.

The typical structure of a table is given below.

STRUCTURE OF A TABLE

(TITLE) Units

Stub heading	Caption			Total
	Column heading	Column heading	Column heading	
Stub entries	<——	Body	——>	
Total				

Foot note :

Source :

Tables can be two dimensional, three dimensional etc. As the dimension of the table increases the presentation becomes a bit more complex and the ordinary user may not be in a position to read and digest the information contained in the table. The number of rows and columns should not be too many as this may make inter- column or inter-row comparisons more difficult. The column and the row totals and the grand totals should be provided wherever it is necessary. In a few tables, the row-wise or column-wise percentages along with numerical data are also shown, so that the reader can appreciate the trends or patterns of the data. It is the general practice that the independent variable is usually presented in the rows and the dependent variable in the columns. The percentages have to be computed along with the columns so that the effect(s) of independent variable on dependent variable can be easily studied.

12.2.3 Graphic Representation

Data under this mode of presentation is shown through various geometrical figures and devices. These may include line graphs, bar charts and other pictorial representations. The vantage of the pictorial presentation is that quantitative data can be shown in a non-quantitative mode which can be easily understood by even non-technical persons. In these graphic modes the shape and size will be proportional to the magnitude of the characteristics. In other words, magnitudes are understood through relative sizes of the picture. Some of these pictures are drawn not necessarily according to the scale. Depending upon the type of information the dimension of the diagram will be varying and two types of diagrams are widely used — single dimensional and two dimensional. These diagrams help the reader to see the pattern and appreciate the complexity of the situation without virtually going through the numerical figures.

The advantages of diagrams over the tables and numerical quantities are as follows :

i) attractive presentation of information leading to a better visual impact;

ii) make data simple and intelligible and also facilitating easy comparisons; and

iii) easy way of depicting the characteristics of the data which also saves labour and time to convey the message/information

Despite the usefulness, this mode has some shortcomings and these include:

i) understanding the multi-dimensional diagrams is not easy;

ii) loss of precision;

iii) incorporation of more details reduces the utility of this mode; Comparison of more than two or three cases will be rather cumbersome; and

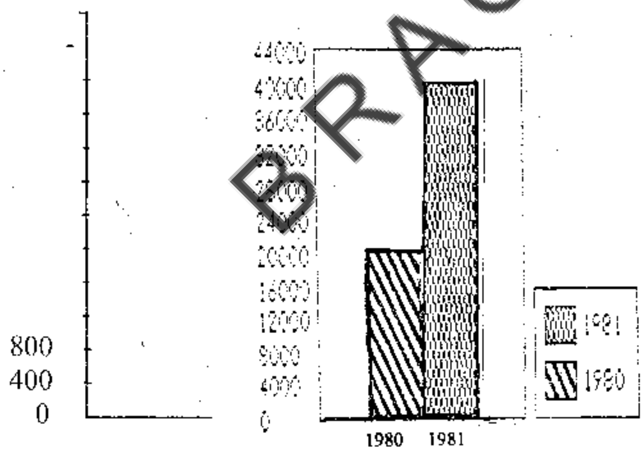
iv) not useful for further statistical processing.

Among the popular modes of graphic representation are bar charts, pie charts, histograms and x-y graphs.

12.2.4 Bar Charts

Most of the commercial statistics are usually presented by bar diagrams. The libraries can display information regarding the procurement of books and journals over years through bar charts. In these bar charts the magnitude of the data is represented by the length of the bar. So, in a bar chart we will have a set of equally distant rectangular bars shown either vertically or horizontally. The length of the bar will be proportional to the magnitude of the characteristic. For instance, if a library purchases 2000 books in 1980 and 4000 books in 1981 the length of the bar corresponding to 1981 should be twice that of 1980. The width of the bar has to be decided taking into account the number of bars which can be accommodated in the earmarked available space. The above information is depicted in the adjacent diagram through bar chart.

Fig - 1

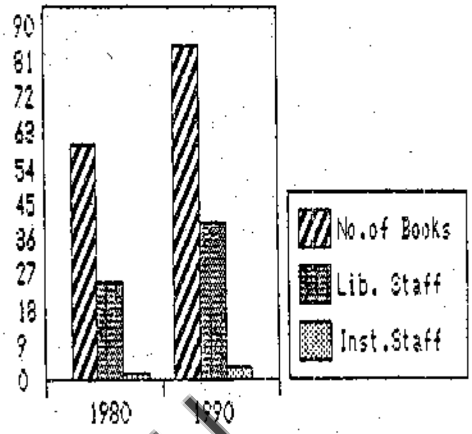


The following steps need to be considered in the drawing of a bar chart. (i) The bar chart can be drawn either in the ascending order or descending order of the magnitude of the data and this will give a pleasing impression ; (ii) the length of the bar should be proportional to the magnitude of the variable it represents ; (iii) width of the bars should be uniform and its size should be determined depending upon the number of bars to be drawn. The quantitative values can also be shown against the bar for the convenience / benefit of the reader.

Multiple bar diagrams are drawn when the researcher wants to compare more sets of statistical data / variables at a time. For instance, the libraries can display information regarding the number of books purchased, the number of employees of the institute and the library staff etc., overtime. It will help the reader to find the type relationships among the growth of library (in terms of books), size of the library staff and the strength of the employees of the institute. This can be shown by marking the bars separately with multiple / different colours or with designs. These multiple bars should be shown on the chart after keeping some space in between one set and the other to ensure clarity.

	1980	1990
i) Books	60000	85000
ii) Institute staff	250	400
iii) Library staff	20	30

Fig - 2

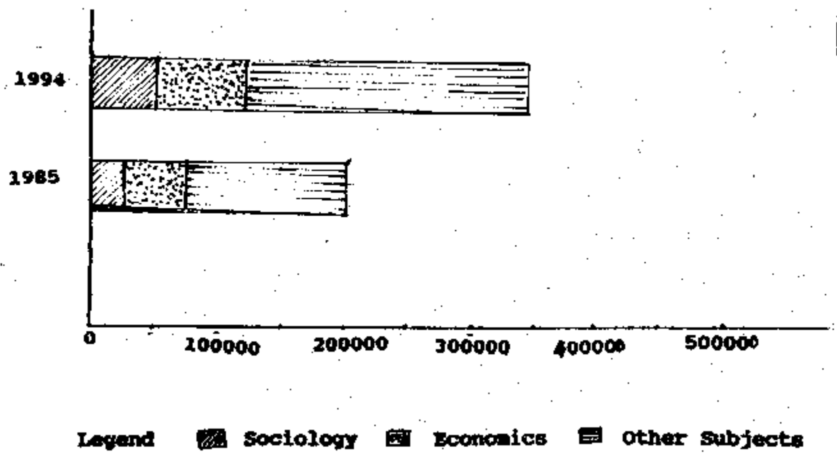


Component Bar Chart

The component bar diagram is also known as sub-divided bar diagram. It is useful for comparing the changes in the component parts among themselves overtime and between each component and the whole. For example the allocation of library funds for various disciplines in a social science research institute be shown through the component bar diagram. Each component will be identified by either a distinct colour or design and the length of that component will be proportional to the fund allocation for that subject. The component bar chart shown below depicts
the broad allocations of money to the various subjects for some selected time points.

	1985	1994
Sociology	25,000	40,000
Economics	35,000	70,000
Others	1,40,000	2,10,000

Fig - 3



The percentage bar chart is a variate of the component bar chart with a total allocation or the whole is equated to 100 and the component shares are represented in percentage terms. For instance, the total expenditure of a library in the year 1993-94 was Rs.3.5 lakhs and the cost towards journals may be Rs.70,000/-. So, in the percentage bar charts the total expenditure of Rs.3.5 lakhs is equated to 100 and the expenditure on journals is shown as 20 per cent.

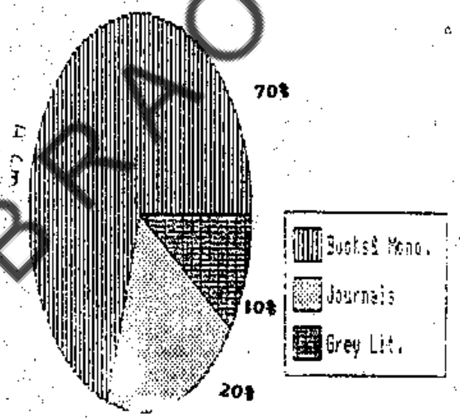
Pie Chart

This is another graphic mode of presentation and the circle is partitioned into various segments and each segment represents the magnitude of the variable under reference. In a bar chart information is depicted in terms of the length of the bar while in the pie chart it is shown through the size of the segment of the circle. Since the full angle of the circle is 360 degrees and this is equated to 100 per cent. In other words 1 per cent is equated to 3.6 degrees. In the construction of pie diagram the total magnitude is equated to 360 degrees and the respective angles of each segment will be worked out. In the following example which is self-explanatory, the expenditure pattern i.e. the various items procured by a library during 1991- 92 is expressed in terms of percentages and depicted through pie chart.

Books and monographs	-	70%
Journals	-	20%
Grey literature	-	10%

When converted into degrees they will be equal to (70x3.6) 252, (20 x 3.6) 72 and (10 x 3.6) 36 and the resultant pie diagram will look as follows:

Fig - 4



12.2.5 Histogram

It is another way of depicting the data through geometrical figures. The class widths or class intervals are shown on the horizontal axis (x - axis) and the class frequencies are represented on vertical axis (y -axis). Histogram is a frequency diagram in which information is shown in terms of rectangles where the area of rectangle is proportional to class frequency. So, it is a set of vertical rectangles which are placed adjacent to one another on the horizontal axis. The class intervals need not be the same. So, the class boundaries will be taken into account while drawing the rectangles. The upper end of the vertical lines are joined together.

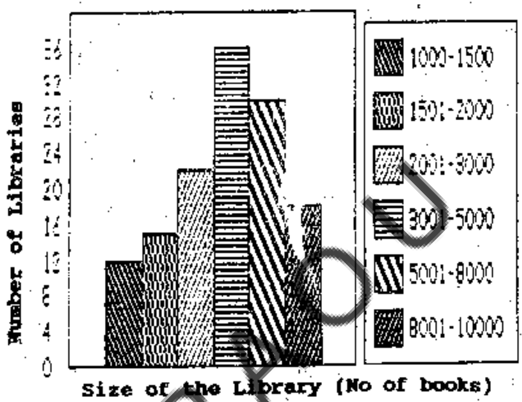
The histogram and bar chart are two qualitatively different diagrams and some people fail to differentiate them. Some of the salient differences between the vertical bar chart and histogram are : (i) The consecutive rectangles are placed adjacent to each other in histogram while in the bar chart there can be space in between the consecutive rectangles. (ii) In the bar diagram, rectangles are of equal width while in the histogram they need not be of equal width and are proportional to the width of the class. (iii) The histogram is a two-dimensional figure while the bar chart is uni-dimensional. (iv) Bar chart is a frequency curve while histogram is area diagram.

Exercise:

The frequency distribution of libraries of different sizes (number of books) is given below. Draw a histogram depicting the information.

Size of the Library (No.of books)	Number of libraries (F)
1000 - 1500	12
1501 - 2000	15
2001 - 3000	22
3001 - 5000	36
5001 - 8000	30
8001 - 10000	18

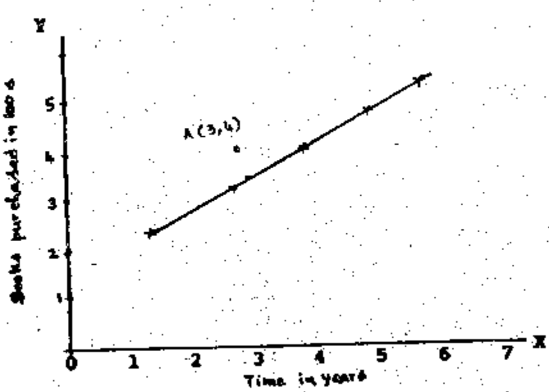
Fig - 5



12.2.6 Graphs

Graphs are another mode of depicting information between two or three variables and useful for representing patterns and trends in data. Usually information on two variables is represented through X- Y graph. The horizontal axis is called x axis and the vertical axis is called as y axis and the point of intersection of these axes is called origin represented by 'o'. A scale has to be defined to represent the data on x-axis and y- axis. A point (3, 4) on the graph indicates that when the value of x is 3, the corresponding value of Y is 4. This point (3, 4) will be represented by a dot on the graph. In the graphic representation the variables can take positive as well as negative values. The collection of points also known as scatter, will indicate the nature and intensity of relation between the variables. The number of books purchased in the library over-time are shown in the graph below :

Fig - 6



Connecting the points on the graph will give us the nature of relationship. If it is a straight line, it is called line graph otherwise it is known as a curvilinear graph. Graphic presentation is found to be more effective in communicating the results with the reader. The (latent) trends and patterns can be displayed through this presentation very effectively.

12.3 INTERPRETATION

The purpose of scientific research is to explain the nature and causes of a phenomenon by proposing relationships between the observed phenomenon with one or more variables and then establish the relationships. A number of statistical methods will be made use of to investigate the phenomenon under consideration and eventually the researcher either establishes the relationship or fails to do so. This implies that one should be judicious in selecting an appropriate method and also be careful in interpreting the results. For careful interpretation of results researcher has to have a thorough understanding of the environment in which the investigation has been carried out and also be aware of the limitations of the methods employed and quality of the data used for statistical analysis.

In social science research as mentioned in the previous sections the researcher cannot completely control the other factors (intervening variables) which will also influence the phenomenon under consideration. So, it is all the more desirable to be cautious in generalising the results. It is found in a number of cases that researchers have a tendency to generalise the results beyond the universe from which the samples have been drawn. Such type of sweeping generalisations may reduce the reliability of the statistical investigation.

In the presentation of data either in terms of tables or diagrams or through statistical measures, the aim of the researcher is to facilitate interpretation of the data by the user / reader. Interpretations automatically lead to inferences and in most of the studies these two terms have been used more or less synonymously. It is evident that mere presentation of facts and figures in one place does not complete the research activity and it is the duty of the researcher to integrate these various disjointed facts with the supplementary information, if necessary, for drawing meaningful inferences. So, the last stage in the social science research is interpretation of the results and then draw the policy implications of the investigation, if any. Further, the researcher should also state the precautions to be taken by the reader in making use of the results. One may even suggest areas or issues for further investigation as derivatives of the current study.

12.3.1 Biases in Interpretation

In a few research studies you might have come across a situation where the findings of the research study are contradictory in nature. These contradiction can be in terms of the results being contra-intuitive and / or some of the findings are mutually in consistent. A number of factors are responsible for such situation and the biases of the researcher, the field investigator and the respondent could also be responsible for the emergence of such situations.

(1) *Biases of the Researcher*:

It has been noticed that the socio- economic-cultural background of the researcher will have a strong influence on the identification of the research theme and also in the formulation of research hypotheses. These biases may influence in the collection of data and also omission of information so that the pet notions / views of the researcher are supported by the empirical evidences. To minimise these biases the researcher has to be objective in his thinking and action and should have an open mind while eliciting information (including probing) from the respondents and also in understanding the environment / social reality in which the study is conducted.

(2) *Investigator Bias*:

In most of the studies the field investigator and researcher may not be the same and the investigator may be a well trained or an untrained person. The gaps in understanding of the problems, differences in perceptions and in the views on the utility of the research study between the researcher and the investigator will influence the accuracy of data. Further, the lack of poor communication skills of the investigator, failure to build rapport with the respondent will also affect the quality of data.

(3) *Respondent Bias*:

In the rural social science research, it has been noticed that there are some built-in biases of the respondent while providing information on some key characteristics like income, age, debts and efficiency of delivery system. The field studies have shown that women usually underestimate their age and men tend to under report their incomes; and while some of the problems and data relating to debts or expenditure are being exaggerated. This could be partly due to the socio-cultural values of the respondents and also due to anticipation of benefits by under-reporting of income etc.

The biases can be minimised by taking precautions in terms of

(a) developing a rapport between the research team and the respondent group;

(b) cross-checking data on some selected items from other;

(c) editing and coding of data with thorough understanding of the environment; and

(d) interpreting the results cautiously in view of the problems mentioned above so that the user can be well equipped with the limitations of the study.

12.3.2 Non-Response Problem

The problem of non-response is increasingly felt by social scientists both while administering the questionnaire and also in conducting the interviews with the help of the schedule. In the case of former, there is no direct contact between the researcher and the respondent except that the researcher through the covering letter seeks the co-operation of the respondents in furnishing the required information. The response depends to what extent the researcher could motivate the respondent in sharing the views and facts with the research team. It is noticed that in many questionnaire based studies, non-response is of the order of 60 to 70 per cent which implies that cost of employing this instrument is very high. Further, non-response has also been noticed when the respondent avoids and evades reacting to certain questions.

In some situations users do not answer the crucial questions. While trivial information is provided, data relating to crucial questions are un-answered by some of the respondents. This situation could be due to the ambiguity in the question(s) framed or due to reluctance on the part of the respondents to provide the necessary details. Pre-testing the questionnaire and consulting the previous studies will help us in refining the instrument and seeking the information from the client. However, there are no well established methods of eliciting information from a reluctant respondent and recontacts and personal visits, if necessary, may help in obtaining the required information to some extent.

In the case of an interview schedule, the field investigator and the respondent will be in direct contact with each other and the former can explain and elaborate the intentions and connotations of the terms used so that the non-response can be minimised. However, a good rapport building before the commencement of (actual) data collection may help reducing evasive responses as well as non-responses.

At the time of data analysis, it is wise to exclude the non-response items of the schedules and carry out the exercise with the remaining responses. If the percentage of non-response is very high, then it is advisable to revisit the field or contact the respondent through mail, as the case may be, to elicit the correct information. If it is not practicable then item (s) under consideration may be deleted for further scrutiny.

12.4.3 Objectivity in Interpretation

Objectivity is the essence of any scientific social science research. The researcher should aim at ensuring objectivity at each and every stage of research process so that he / she will be successful in constructing the social reality. The first and foremost condition is that the researcher should not impose his or her values at any stage of the investigation. The various sections of the universe should be well represented in the sample and their participation in the conduct of research study should be encouraged. For this the sample size should be adequate and the sampling design should be such as to ensure representativeness of the sample. However, the principle of randomisation should be adhered to in toto for the same.

It is advisable to have a 5 to 10 per cent of the sample schedules rechecked and correction factors may be developed so that it will guide the editing of the schedule. To ensure objectivity, it is necessary for the research team to develop an equation with the sample respondents and make them an integral part of the research team. Special features of the environment or universe should be highlighted so that cautious generalisations will be attempted.

It will be useful if conditions under which field investigations are carried out are also reported for the benefit of the user. Supplementing the analysis with a few case studies may help in prescribing policy measures to improve the situation.

Some of these precautions will help not only in ensuring objectivity of the entire research work but also will enable the user to make his own inferences from the results presented. Such a process will also help in identifying a few more hypotheses for further research.

12.5 LET US SUM UP

Presentation of research results is an important activity for any research study. A number of devices and methods are available to make the presentation of results attractive and understandable to even non-professionals. The tabular way of presentation and the graphic modes are discussed.

Interpretation of the results and inferences are more or less used synonymously. The problem of bias has been discussed at length and measures to control different type of biases are also suggested through proper interpretation of the results. The veracity of hypotheses and establishment of relationships and even identification of new issues will be possible.

12.6 ASSIGNMENTS

i) Take a research study or Ph.D dissertation from the library and depict the data on crucial variables by means of charts, and graphs.

ii) "Objectivity in social science research is a myth" - Give your comments.

12.7 RECOMMENDED BOOKS

BLACK, J.A. and D.J. Champion. *Methods and Issues in Social Research*. New York: John Wiley, 1976.

EDWARD, A.L. *Statistical Methods for Behavioural Sciences*. New York: Rinehart & Co., 1954.

12.8 MODEL EXAMINATION QUESTIONS

I ESSAY QUESTIONS

1) Write a note on the data presentation methods.

2) Explain the problem of 'bias' in interpretation of data and suggest ways to control biases.

II SHORT NOTES

a) Non-response problem

b) Histogram

c) Pie charts

BLOCK - IV: REPORT WRITING AND EVALUATION STUDIES

In this last block of Course-06: Research Methodology, there are four units. They deal with post-research activity - writing and evaluation of research. Report writing is a crucial step in the entire process of research. The purpose of the report is to communicate the entire process of research, starting with aims and objectives, methodology, data collection, analysis and interpretation of results and conclusions/suggestions. Bibliometrics deals with qualitative and quantitative aspects of writing and publications. In a way publications are the outcome of some studies and the qualitative or quantitative measurements help us to evaluate them. The LIS research in India and abroad as provided in the last two units (15 and 16) gives us an evaluative picture to help those interested in taking up research projects in LIS field.

Unit-13 deals with Writing the Research Report. Composition and style of presentation are discussed in this unit.

Unit-14 deals with Bibliometric studies. Bibliometry involves measurement of several interrelated aspects of writing and publication. Bibliometric laws and Citation studies have come to prominence in library and information science field in the last two decades. These aspects were briefly explained in the unit.

Unit-15 provides an over-view of the current trends in LIS research at the international scene. It also gives a brief introduction to the history of LIS research in USA and UK. Under the present status of LIS research it gives institutional support, thrust areas, research methods, funding problems and publications. The unit also touches upon the future trends.

Unit-16 deals with LIS research in India. It starts with the early studies on libraries and then gives the present status under five sections, each of them touching briefly about the research programmes, research areas, funding problems and publications. The future of LIS research in India is also discussed.

UNIT - 13: RESEARCH REPORT WRITING

Structure

13.0 AIMS AN OBJECTIVES

The unit aims at introducing you to writing a research report. It explains the composition as well as style and presentation of research report.

After studying this unit, you should be in a position to

- describe the composition of a research report
- explain the style of presentation
- write a report for a research work done by you.

13.1 INTRODUCTION

Researchers have an obligation to effectively communiate the information about the investigations they have completed. This is done through reporting in professional meetings, conferences, or through learned periodicals. Research work done as a part of the academic work for award of degrees are reported as theses or dissertations. Reporting also occupies an important position in research process. Research reports are detailed and accurate accounts of the systematic studies conducted to solve problems in a specific/interdisciplinary subject field.

Writing the Research report needs a specialized skill, but it is not difficult to learn. The format and the details required for the preparation of a report may vary according to academic discipline. The basic elements and the structure is more or less the same for many of the disciplines in social sciences, however slight variations occur in their framework. The structure of the research reports has been standardized into some models. Such models help us shape the presentation into a consistent and logical framework. If we follow a model reporting and writing easier. The function of the report is not just depict the research findings, but to link those research findings directly to theoretical aspects and hypothesis proposed and tested.

13.2 RESEARCH REPORT WRITING

Report writing is a crucial step in the entire process of research. A carelessly prepared report affects the degree of respect the research commands and at the same time a carefully prepared report will bring recognition to the writer and the work into the limelight. Therefore, writing of the report should be taken as an equally important task in the entire research process and sufficient time should be devoted to it.

The purpose of the research report is to communicate the entire process of research i.e. objectives, methodology, data collection, analysis of data and their interpretation, results and conclusions drawn. Keeping this in view, the author should make attempts to clarify his own thoughts in order to make it meaningful to the readers. The level of knowledge and understanding of the readers should be kept in mind.

There are three stages in report writing - planning, writing and revising, and editing.

13.2.1 Planning

Planning is thinking about the content of the text and its organisation. During this stage you should map out the broad issues involved and sequence them in order to put them in the writing form.

13.2.2 Writing and Revising

Writing is putting down one's thoughts on paper (or on a computer screen). You can start writing your report even at the early stages of your work. Write up description of the procedures and findings more or less you go along. These writings may help you as summaries and turn out to be as a first draft. The initial writings/first draft gives an opportunity to check your original ideas and develop new ideas as you proceed with your work.

After completing the first draft you can think of improving it. So you have to revise it. Novices generally focuss more on changing the wording of the sentences they have written. While expert writers sequence the content and rewrite the longer chunks. When you first read the text you may not see any problem, so read it several times. If you find difficulty in revising you may consult your research supervisor and his rich experience help you a lot.

James Hartley (1997) has given some helpful strategies for beginning thesis writers, which may be summarised as below:

(1) Try to be well organized. Plan well ahead. Try to keep to the plan.

(2) Keep notes, files, xerox copies of materials that may be relevant to your thesis, even if they will eventually be discarded. Some people suggest keeping separate folders for the separate chapters of the thesis.

(3) Write from the beginning. Do not leave 'writing up' until the end - you will forget what you did, and why you did it. So keep a written record. With word-processors one can easily add, change, move around or delete words, sentences and paragraphs. If the thesis reports experiments, then early experiments and pilot studies should be written up in full at the time of doing them, even if this detail is not needed in the end.

(4) Think of how one chapter or other parts of the thesis might be published separately after the thesis is completed. Write them in such a way that it will be relatively easy to do this. After all, few people read theses, so preparing to publish in another format may save you time later on. But don't get distracted. The thesis comes first.

(5) Master at the outset the appropriate procedures for presenting text in your discipline, particularly the presentation of footnotes and references. References should be stored - preferably on a database - from the outset in full detail. There is nothing worse than trying to find something that you read several months before just to put the reference in your thesis.

(6) Find out the requirements of your institution for the presentation of the thesis. (This refers to margins, line spacing, tables, type face size, etc.).

(7) Regularly submit drafts of sub-sections of your thesis to your supervisor, and ask for guidance on your writing - particularly if you are an overseas student. Also ask your supervisor if there are any books or articles on thesis writing in your discipline or examples of previous thesis that can be recommended.

13.2.3 Editing

During the writing stage you write as quickly as you can without paying much attention to punctuation and spelling. Editing means making sure that only details most relevant to your method and your argument appear in the report. This stage involves rethinking and replanning as well as correcting spelling errors and the like. You can divide this stage into different sub-stages: you might edit the content first, then for grammar and style and finally check the format of the references. Word-processors are computer based writing packages that help you in editing the text. Word processors like Wordstar, Softword, MS-Word are designed for text processing and they have built-in thesaurus, which help in checking the spellings. MS-Word provides provision for checking spelling and grammar also.

Grammatik-5 is a computer programme which highlights possible grammatical, mechanical and stylistic errors and also suggests possible solutions. Though some of these programmes are tedious, they are found useful by most researchers. A sample of computer programmes available in Grammatik-5 from Apple Macintosh is shown below:

Programs which indicate grammatical errors:

- adjuctive errors
- adverb errors
- article errors
- clause errors
- comparative/superlative use
- double negatives
- incomplete sentences
- noun phrase errors
- object of verb errors
- possessive misuse
- preposition errors
- pronoun errors
- sequence of tense errors
- subject-verb errors
- tense changes, etc.

Programs which indicate mechanical errors:

- spelling errors
- capitalization errors
- double word
- ellipsis misuse
- end of sentence punctuation errors
- punctuation errors
- number style errors
- question mark errors
- quotation mark misuse
- similar words repeated, etc.

Programs which indicate grammatical errors:

- long sentences
- wordy sentences
- passive tenses
- end of sentence preposition
- split infinitives
- cliches
- colloquialisms
- Americanisms
- archaic language
- gender-specific words
- jargon
- abbreviation errors
- paragraph problems
- questionable word usage, etc.

13.3.4 Books on Academic Writing

There a number of publications on academic writing of research papers, theses and dissertations, scientific reports, etc which may help the researchers writing their reports. The following are some of the major works on this topic:

Brown, S., L. McDowell and P. Race. *500 Tips for Research Students*. London: Kogan Page, 1996.

Cryer, P. *The research student's guide to success*. Buckingham: Open University press, 1996.

Day, R.A. *How to write and publish a scientific paper*, 4th ed. Cambridge: Cambridge University press, 1995.

Fabb, N. and A Durant. *How to write essays, dissertations and theses in literary studies*. London: Longman, 1993.

Gibaldi, Joseph. *MLA Handbook for writers of research papers*, 4th ed. New York: The Modern Language Association of America, 1995.

Meyer, Michael. *The little, brown guide to writing research papers*. New York: Harper-Collins, 1991.

Philips, E.M. and D.S. Pugh. *How to get a Ph.D.* 2nd ed. Buckingham: Open University press, 1994.

Swales, J.M. and C.B. Feak. *Academic writing for graduate students*. Ann Arbor: University of Michigan Press, 1994.

Watson, G. *Writing a thesis: a guide to long essays and dissertations*. London: Longman, 1987.

13.3 COMPOSITION OF RESEARCH REPORT

A research report is a fully documented research work based on the identification of the problem and the context in which it was selected. It covers the procedure followed, the findings, the results derived from the analysis of the findings and a discussion of the results.

A research report consists of:

I. Introduction

It covers the background of the study area, the problem and the objectives.

II. The Body

It covers the Contents between introduction and conclusions.

III. The Terminal Part

The conclusions, generalizations, discussion and summary are included in the terminal part.

13.3.1 Structure of Chapters

Each chapter should contain a title and it should begin on a new page. Each chapter may contain data and the arguments emerging from the data. The chapters may include:

Chapter	I -	Introduction
Chapter	II -	Review of existing Literature on the chosen topic
Chapter	III -	Methodology : Sampling Techniques and Data Collection Methods
Chapter	IV -	Analysis and Interpretation of Data
Chapter	V -	Conclusions, Generalizations, Implications and Discussion
Chapter	VI -	Summary

Introduction: It should clearly and concisely state the significance of the research problem and the purose of the report. It should also avoid lengthy explanation of your purpose of the report. Introduction presents an overview of the content. The opening sentence is another device after the title to get the attention of the evaluators and/or readers. You know that writing involves thinking and hence you have to use all your imagination in writing the opening paras. You should also state the scope of the study, and also methodological and theoretical limitations of your report in the introductory part. You have to define unusual and unfamiliar terms used in the report.

Many beginning writers have a misconception that introduction should be written first since it opens the report. This may not be true always. Generally the first few sentences for many amateur writers is a difficult job and they may not be in proper order. "If you have organised your work properly and outlined the text, and if you are fairly sure of what you are going to say (although there may be interesting surprises in the store—writing always involves thinking, and thinking can lead you anywhere), try writing until you have drafted the rest of the report to write the introduction". (Frankfort-Nachmias and Nachmias, p.358)

Literature Review: Any research work, how innovative the topic or procedure may be it surely rests on the findings and insights of other researchers. A literature review is, therefore, necessary to set the stage or background for your study. It summarises the past research in a logical sequence and then relates your work to it. Literature review should be se.ective, hence you have to include only the sources most pertinent to your work. Another important thing to remember is always use only scholarly works, since they are directly involving the theories or that test hypothesis. It is proper to interpret the events and relating your work to them. Researcher comments briefly on state-of-the-art of the research in this section. Newsmagazine articles are inappropriate since their purpose is to relay information to general public.

Methodology: This part deals with the methods used in the research study. It covers all the steps taken to conduct the study and collect the data (e.g., select of subjects, the location of the study, instruments of measurement used, and special problems pertaining to the study design, etc.), besides the sampling techniques used.

Analysis, Interpretation and Discussion: This section of the report deals with the methods of analysis, interpretation of the results/findings and their integration with the research questions. You have to use your analytical skills and imagination here. You have to state whether your findings support your original hypotheses or validate your line of research. The implications of your work for further research may also be emphasized.

Conclusions: The main purpose of writing the conclusions part is to close the circle of research, moving from past research to the place of your work in the subject field. This requires to link your theoretical or methodological conclusions with the current state of research. This helps to expand the existing knowledge. Without conclusions your work lacks closure. It the conclusions are very brief, you can integrate them with the discussions section.

13.3.2 Pagination

A research report is divided into two main parts - preliminary pages and the main text. For the preliminary pages, Roman numerals (i,ii,iii etc.) are used and for the main text, Arabic numerals (1,2,3 etc.) are used. In general, Roman numerals are placed in parenthesis at the bottom of the page while Arabic numerals are placed in the upper right-hand corner.

Some of the important points to be remembered are: There is no page number given to the title page but it should be taken into counting as page number one. The first page of each chapter is counted, but the Arabic numeral is placed at the bottom of the page. There should not be a full-stop after the page number.

13.3.3 Preliminary Sections

(1) **Blank Page**

(2) **Title Page**

It contains a complete title of the research topic, in the centre of the page in block letters. Below the title, the name of the author is given and at the bottom of the page, the name of the institution or university and the year in which the research was carried on.

Sample Title Page

Title

By
(Name of the Author)

A Dissertation submitted in partial fulfillment
of the requirement for the degree of _______
in the University of _______

Date and
Place

(3) **Certificate of the Guide**

Certificate of the Guide about the originality and genuineness of the research work is given if it is a thesis submitted to a university as a part of fulfillment of the formalities for obtaining a degree.

(4) Acknowledgement Sheet

This contains names of people whom the author wishes to express his/her gratitude for the help received from them. Assistance can be from persons or from institutions.

(5) Preface

This reflects researcher's understanding of the problem and his perspective of the problem. He may also bring out the difficulties encountered, if any, in carrying out the research and the limitations of the study.

A preface is often considered as synonymous with 'foreword'. It includes the reasons why the particular topic was selected by the author.

Preface is also a right place for acknowledgements. If the author decides to discuss the significance and nature of research in the introductory chapter, then he may not require to write a preface, instead he may use the page for 'acknowledgements'.

(6) Table of Contents

It is placed before the preface. This list is necessary to indicate what the report contains - parts, chapters, sections, sub-sections, bibliography, appendix, tables, etc.

Table of contents should indicate page numbers too. To facilitate this process, this table is prepared at the end, after typing the whole report including page numbers. Care needs to be taken in giving page numbers in the contents list, taking into consideration the illustrations which too required to be numbered.

Whether sub-headings of the chapters should be included in the contents page is a debatable point. If there are too many sub-headings, instead of having the table running into several pages, it is better that they are omitted.Instead, the list of sub-headings may find a place, below the title and number of the particular chapter.

The heading 'Table of Contents' should be given in the centre of the top of the page in capital letters. On the right hand should be typed the heading 'Page', at the right margin below which page numbers appear. The sequence of headings should be same as it appears in the text. Main headings should be in capital letters without a full-stop and are numbered consecutively. Sub-headings should be given under the main heading after leaving two spaces. Sub-headings may be upper-lower letters i.e., capital letters for the initial letter and lower case letters subsequently.

(7) List of Tables and Illustrations
(Maps, Diagram and Figures etc., if any)

Regarding listing out tables, illustrations, etc. title of each table and the page on which it appears should be given clearly in Arabic numerals. Figures and illustrations are synonymous terms and refer to maps, drawings and photographs. These too have to be presented numbered in capitalised Roman numerals and with page numbers. They should be placed as close as possible to their first references in the text.

Tables and figurs are the summary of the research findings used to simply the presentation of a theoretical argument. The figures use only words (concepts) to summarise the ideas discussed in the text. Some guidelines for constructing tables are given below:

(1) Tables should be clear, logically planneds and accurate. They should be self-explanatory. They should enable the readers of the report to understand them without reference to the text.

(2) Each table or graphs should be referred to by numbers so that their exact placement is known. (e.g., Table-1; Figure-2). The number and title given to the table or graph should be placed at the top.

(3) The title should clearly to refer to the contents of the table (what, where and when). The title and related information should be centred.

(4) Column headings should refer to independent variables; row headings to dependent variables.

(5) Column and row headings should be clearly distinguishable from each other and may be numbered for clarity.

(6) The column and row figures centered, aligned on decimal points, and aligned with the row headings.

(7) The only abbreviations that are generally permitted are n.a. (not available) and N (frequency). The N may appear in parenthesis directly below calculated percentages if they are small but should always appear as part of the column summary below the total of 100 percent.

(8) Asterisks or other symbols may be used as indicators of footnotes or of unique data entry sources. Keep these to a minimum to avoid clutter and confusion on the part of the reader.

(9) Explanatory footnotes appear directly under the table, in full citation form. Sources of data, also in complete citation form, appear underneath the footnotes.

Report writing differs from subject to subject in social sciences depending on the data, but the basic pattern remains the same.

13.3.4 Appendix

As the name suggests it includes all useful material that could not find a place in the main text. Appendix is a portion of the report which has a direct bearing on the report. It is included where needed, to provide supporting evidence or additional matter related to the text. Elaborate statistical and other tables that could not be presented in the text, schedules or questionnaires used in the collection of data, explanatory notes, etc. where necessary, are given in the appendix. This matter, if given in the main text, would spoil the flow of argument and also would occupy a considerable space.

The appendices may be arranged in a logical order. Each appendix is lettered or numbered as appendix A,B,C or I,II,III and so on. These should be listed accordingly in the table of contents.

13.3.5 Bibliography

The last portion of the research report is the bibliography. It is arranged in an alphabetical order according to the surnames of the authors. Bibliography contains those books which are referred for the research work and those have direct bearing upon the research report.

Please remember, bibliography is not a parameter to test your report's weight or your knowledge of the subject but it refers to works cited or sources collected. Every work given in the footnote must be included in the bibliography, because the quotation, citation or summary given in the text will have footnote reference.

The citations should follow uniform style with enough details so that the reader can easily locate original source. There are ISI standards for rendering a citation of a book, periodical article, collected works etc.
For example:

The following items of nformation need to be included for books:

Author's full name/Editor, compiler, or translator, if any complete title of the book

Series, if any, and volume or number in the series Edition, if not the original

Number of volumes, Facts of publication - place of publication, publisher, date of publication

Volume number, if any

Page number(s) of the particular citation

Example-1:

Strunk, W., Jr., and White, E.B. *The elements of style.*3rd ed. New York: Macmillan, 1979.

Example-2:

Hartley, J.T., Harker, J.O., & Walsh, D.A. "Contemporary issues and new directions in adult development of learning and memory". In L.W. Poon (Ed.). *aging in the 1980s: Psychological issues* (pp.239-252). Washington, D.C: American Psychological Association, 1980.

For Periodicals the items of information to be included is generally as follows:

Author's name
Title of the article
Name of the periodical
Series number
Volume number
Date of publication
Page numbers.

Gupta, D.N. "Public Distribution System - Involvement of Institutions". *Journal of Rural Development*, 15(2), April-June, 1996, 201-213.

13.4 STYLE AND PRESENTATION

13.4.1 General Style

Style depends on purpose. In the case of a research report, you need to vary the style and presentation according to whether the report is intended to convey information for general use, specific information for a focussed clientele in an event, as in a discussion paper, or as a format required for a training course. Since a research report is a format document, its language and style should, therefore, be appropriate.

You should try to develop a style of your own that is smooth and clear. Well chosen words, short sentences, variety in expression, paragraph unity, proper introduction and conclusion, will add up to your style.

Some of the specific attributes of good reports are legibility, absence of grammatical and spelling mistakes, good structure of sentences, right length of paragraphs with proper headings. A report should be readable following the conventions of the written language for effective communication. A good style does not mean flowery language or use of high sounding words. It should not be ornamental or superficial. It should be appropriate to the subject matter.

You should check to see if your writing is 'even' in its style. Patches of unsatisfactory expression, if any, have to be brought on level with the rest, leveling upward, not downward.

While checking your style, i.e., expression, you have to begin by studying how different sentences have been linked to form paragraphs, paragraphs to form sub-sections and sub-sections to form chapters of the right size and so on.

Please remember that research reports deal with 'ideas-in-relation' and not with isolated ideas. Accordingly, all your sentences in paragraphs and paragraphs in chapters should be logically presented. Another matter which needs to be looked at is the quality of 'individuality' of expression. In everything that you present, your very own style, different from that of others, should emerge. Without affecting the scientific objectivity, you should create an individuality in style for report writing. This will enhance the overall impact of the writing.

13.4.2 Scientific Style of Report Writing

(1) *Objectivity*: Objectivity is the main characteristic of any research. Subjective and emotional attitudes have no place in research or report presentation. The author's background like his caste or religion may unconsciously reflect in his writings, which should be avoided to maintain objectivity.

(2) *Clarity*: Unless your writing is clear to the reader, the value of your research will go down. A confused report bores the reader. Straight-forward communication with the reader in active voice is always advisable.

(3) *Conciseness*: It would contribute to readability and save the time of the reader.

(4) *Diction*: It means choice of words in report writing. The author should have a clear understanding of the words he is using. Many errors in diction arise from mistaking one word for a related word. For example, effect (bring about) for affect (influence). Sometimes, there is a confusion between the meanings. The use of metaphor should be avoided as far as possible in report writing. Ex. " a glaring error".

(5) *Appropriate Means of Expression*: There should be a limitation in the use of figures of speech, particularly those which reflect emotion. Slang or jargon (Ex: Womenfolk/ Hardware which used in the computer related article gives a meaning specific to that profession whereas in ordinary language it has a different meaning) should be avoided in scientific writing. The writer should avoid expressions like 'unfortunately' or 'surprisingly'. When comparing others' data cynicism, sarcasm should be avoided. It is not becoming of a scholar to point a finger at another researcher.

13.4.3 Presentation of Statistical Data

While data are collected in detail, they usually cannot be reported or presented as they are. In order to summarise and present data, tables, graphs or charts are constructed, averages and percentages are calculated. Presentation is for making clear the relationships that exist among the variables under study.

When two variables are related in such a way that the more there is of one (Ex: spending time in library), the more there is of the other (better marks), since the data are relatively straight forward as it stands. Sometimes it is better to present the tabular results as percentages.

As a general rule, if you are presenting your data in tables using percentages, it is best to percentage the independent variable across the dependent variable.

Tabular presentation of data is very basic and useful. There are other ways of presenting data as well, which include several kinds of graphs and the use of charts. All these need to be numbered and given a title. (These are discussed in detail in Unit-12).

13.4.4 Headings and Margins

Presentation of headings of chapters and sections are important. Three kinds of headings, namely, 'centred', 'side' and 'paragraph' headings are used. The central headings are useful in indicating major division in the matter. The side heading and paragraph headings serve to demarcate sub-divisions.

Every chapter must begin on a new page. Chapter number should be in capital numerals. The chapter title is also typed in capitals. These should be a gap of three spaces between the title and the beginning of the text.

Margins should be 5 cms from the left of the page which will facilitate binding. On the right side, two cms of space should be left which will facilitate readability. Similarly, some margin should be left above and below the text which will give a good look to the page.

13.5 SUMMARY / ABSTRCT OF THE REPORTS

Researchers are sometimes required to prepare the summary of their research report. These summaries are stringent condensations of the research reports and usually presented in the form of an abstract. An abstract usually includes a statement of the research problem, a brief description of the methodology used, and the findings. When a researcher prepares an abstract of his own work, they are called 'Author abstracts'. These abstracts are included in the secondary sources of information such as indexing and abstracting periodicals. Abstract help to locate works of interest.

13.6 EVALUTION OF RESEARCH REPORTS

Research work aimed at securing an academic degree (such as Ph.D. or D.Sc.) is compulsorily and thoroughly evaluated prior to the award of a research degree. After working on a topic a period of three or four years (under normal system) for a doctoral degree, the researcher completes his researach work and prepares his/her thesis or dissertation and submits it to the university where he/she registered. Then the evaluation process starts.

Adequacy of the thesis as a doctoral dissertation depends on the characteristics like - originality, the importance of the research questions being investigated and the extent of to which argument is sufficiently backed up by evidence. With his experience the research supervisor is able to judge easily the strength of the work and only when he satisfies himself that the thesis is acceptable for the award of the research degree.

The universities follow certain prescribed procedures for evaluating the dissertations. A doctoral research committee is constituted faculty-wise and it is duly approved by the vice-chancellor to nominate the examiners or adjudicators. Since the researcher and his supervisor probably know who are likely to be experts in the field, many universities allow them submit a panel of examiners. From the panel univerisity selects three examiners - one internal and two external examiners. Some universities still adopt a system of one foreign examiner and two Indian examiners. The research supervisor may be or may not be one of the examiners. The thesis is sent to these examiners for valuation. After the evaluation reports are received, depending upon the recommendations made by the examiners for acceptance, the university set a date for oral examination, often known as the *viva-voce.*

The *viva-voce* is conducted in an intellectually firm, but friendly environment. In many universities the viva-voce is conducted in the concerned faculty, but in some universities it an open viva-voce. The purpose of an oral examination is for the student is able to articulate his/her main findings and to justify these in the light of the evidence gathered. He/she should defend his/her thesis against the questions raised by the examiners about the theoretical framework underlying the thesis, the nature of the research problem, methodology used, treatment of data statistically, interpretation of the findings, possible impact of the findings on the existing knowledge etc. In addition, the examiners also ask questions on literature review, well-known books and journal articles which have not been included in the thesis. After satisfying themselves the viva-voce committe submits its evaluation report. Based on the recommendations of the examiners, the university awards the research degree.

Research work is carried out not only for academic degrees, but also for various purposes. Many such research works are of smaller problems usually involving a shorter period and end up with publication in learned periodicals. The publishers of the reputed journals follow their own evaluation procedures before accepting the articles for publication. They adopt certain standards and conventions regarding the content, format, figures, units of measurement, references, etc. The editorial committee after examining th manuscript of the research article sends it atleast two independent subject experts for their scrutiny and comments. Acceptance or rejection of the research article for publication depends on the views of experts.

13.7 LET US SUM UP

Let us recapitulate briefly what has been discussed so far in this unit.

* Report writing is a crucial step in the entire process of research. It function is not really to depict the research findings, but to link the findings with the theoretical aspects and the hypotheses proposed and tested.

* A research report consists of Introduction, the body and the terminal part.

* The text of a research report is divided into several chapters, which include introduction, literature review, methodology followed, data analysis and interpretation, conclusions/generations and summary.

* The style and presentation generally depends on purpose. Language and style should be appropriate to the level of readers.

13.8 REFERENCES AND RECOMMENDED BOOKS

BEDEKAR, V.H. *How to Write Assignments, Research Papers, Dissertations and Thesis.* New Delhi: Kanak Publications, 1982.

BUSHA, Charles H and Stephen P. Harter. *Research methods in librarianship: Techniques and Interpretation.* New York: Academic Press, 1980.

DHONDYAL, S.P. *Social Science Research and Thesis Writing*. 3A/7, Azadnagar, Kanpur.

DIXON, Beverly R., Gary D. Bouma and G.B.J. Alkinson. *A Handbook of Social Science Research: A Comprehensive and Practical Guide for Students*. Oxford University Press, 1987.

FRANKFORT-NACHMIAS, Chava and David Nachmias. *Research methods in social sciences*, 5th ed. London: Arnold, 1996.

GIBALDI, Joseph and Walter S. Achtert. *Handbook for Writers of Research Papers*. New Delhi: Woley Eastern Ltd., 1987.

HARTLEY, James. "Writing the thesis". (IN *Working for a doctorate: a guide for the humanities and social sciences*/ edited by Norman Graves and Ved Varma. London: Routledge, 1997).pp.96-112.

HASTINGS, Eells. *Writing a Thesis: A Guide to Writing Term Papers, Theses and Long Reports*. Ohio, Yellow Springs: The Antioch Press, 1947.

TURABIAN, Kate L. *A Manual for Writers of Research Papers, Theses and Dissertations*. William Heinemann Ltd., 1983.

13.9 ASSIGNMENT

Compare and contrast the structure of theses in social sciences with that of science and technology.

13.10 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) What is the Composition of a Research Report?

2) Describe about Various Styles of presenting the data in a research report

II. SHORT NOTES

a) Title page of Report
b) Purpose of Research report
c) Importance of graphics in a Report
d) Chapterisation

UNIT - 14 : BIBLIOMETRICS

Structure

14.0 AIMS AND OBJECTIVES

The unit aims at introducing you to Bibliometrics as a method of research in library and information science.

After studying the unit, you should be in a position to

- explain the quantitative method, viz., bibliometrics, in relation to research methods in library science
- describe the growth pattern of bibliometrics, as an emerging major area of research and study in the field of library and information science
- trace the advances in the field from a mere frequency counting technique to an advanced method of mapping scientific and sociological disciplines.

14.1 INTRODUCTION

Research is defined as a systematic enquiry. All research must follow some method, for successfully performing the task and getting the right results. Thus method here refers to the technique that is followed in the process. There are different techniques that are adopted, such as, survey, questionnaire, case study, observation, etc. Survey can be of two types, field survey or literature survey. Literature survey can be done in two ways, by qualitative (e.g. content analysis) or by quantitative approach (e.g. citation analysis). In our context, bibliometrics is a quantitative method that can be followed in research. This is one way to understand the meaning of the term bibliometrics. But some qualitative studies can also come under the bibliometric technique, and so there is flexibility in the use of the term.

What is the relationship of research, bibliometrics and library science? This connection can be understood if we understand what library science is. Library science, with its teachers, practitioners, allied specialists, etc, has become prolific, whether in publishing, or developments, or reporting research, and has come to a stage when its literature furnishes the basis for a multifaceted discipline. In this same sense of inter- and multi-disciplinary nature, bibliometric studies, have further widened the bases, with the coming of statistics, mathematics, and operations research. The existing, list of related areas, include, computer and information science, cybernetics, cognitive psychology, artificial intelligence, semiotics, and linguistics, computer engineering, decision science, education, etc. This trans-disciplinary connection makes library science, bibliometrics and research all compatible with one another.

Among the different research methods, bibliometrics stands as a part of the quantitative techniques adopted in library and information science. The other research methods, are experimental method, survey method, historical research, operations research, etc. There is no hard and fast line between these, and each researcher has the flexibility in adopting one or more than one, and conduct the study in any combination.

Similarly, there are no lines between quantitative and qualitative methods. Bibliometric analysis depends heavily on numbers (statistical data, figures, charts, bar diagrams, etc). But it also looks for studying the underlying behavior of the documents, such as, why do people cite, why do author's show preference in using certain journals, why are some author's most productive in few years, and so on. And only through research and study the field of bibliometrics has gained universal acceptance, leading to publications, regular conferences, and further research.

14.2 BIBLIOMETRICS-DEFINITION, SCOPE AND DEVELOPMENT

What is the meaning of the term bibliometrics? Biblio= 'book'; metric= a scale or measure. Alan Pritchard was the first to use this term in 1969. He suggested that the word 'bibliometrics' connotes "the application of mathematics and statistical methods to books and other media of communication." The term, according to Pritchard, included, all quantitative studies dealing with written communication.

14.2.1 Definitions

Sengupta defined bibliometrics as "Organization, classification and quantitative evaluation of publication patterns of all macro micro communication along with their authorships by mathematical and statistical calculus" (1988:40).

Hertzel finds a practical approach in the bibliometric analyses and states: "Bibliometrics is the analysis of the structure of literature using various tools, counting, rank-frequency distributions, and citation analysis; and although the structure of literature is basic to all disciplines, it is particularly important in the area of information retrieval." (p.198).

Bibliometrics then refers to applying various methods of analysis of recorded communication, aimed at identifying different characteristics features and behavior of the documentary sources. The features include both: a) the bibliographical details (or variables), b) factors that are connected with document generation, (i.e., author's interests, publisher's affiliation, subject scatter, publication date, geographic concentration, etc.) Behavior of the documents studied, includes, analyzing co-authorship pattern (to know the level of team spirit); role of institutions (political, economic, technological exploitation, etc) in publishing and conducting research, etc. This analysis is arrived at by using techniques like, grouping, measurement, observation, and mapping. Bibliometrics then can be defined as a branch of library and information science, concerned with the techniques of measurement and evaluation of communication media.

14.2.2 Utility

Why do we need this method of study, what advantages does it have, and what functions does it fulfil? The advantages in using this technique, interestingly, are seen in both qualitative and quantitative approaches.

Bibliometrics and scholarly communication are two areas relevant to communication research in the context of content analysis and organizational communication. Bibliometric technique is now applied to topics from political communication to the new media. This in turn has provided a chance for information scientists and communication researchers to collaborate in an area of common interest. This then initiates a long desired integrated approach, or the systems approach, in which there is a linking of all related activities, such as communication research, library research, and information seeking research. Content analysis of different forms of media, printed, or non-printed, is useful for communications, libraries, and information scientists. This is only one instance, and the method of bibliometrics provides an answer for all the three different areas.

The qualitative utility of bibliometrics are seen in the following functions, as stated by Ferrante:

a) Tracing the spread of ideas as a study of epidemics

b) Establishing a transition point between zones of higher and lower utility

c) Classifying segments of a literature through inter-connections of citations

d) Determining the impact value of a given document

f) Locating criticisms of published results of research and experiments.

Bibliometrics is, however, more concerned with quantitative utilities and approaches. This quantification has led to formulation of laws, which are derived on the basis of some characteristics of literature. The characteristics applied are as given below, and have led to the emergence of respective laws:

i) measuring the productivity of an author; based on the number of published articles (Lotka)

ii) ranking of words in a text, based on frequency of occurrence of words (Zipf)

iii) scattering of articles on a subject in various periodicals (Bradford)

This quantification utility has the following functions, as stated by Brookes:

a) Design of economical information systems and networks

b) Improve efficiency of information handling processes

c) Identification and measurement of deficiencies in existing bibliographical services

d) Prediction of publishing and distributing trends

e) Discovery and elucidation of empirical laws that can provide a basis for developing a theory of information science.

Interestingly, the last mentioned function results in a combination of the quantitative and qualitative dimensions, in this it meets research needs in both the areas and forms the basis for more qualitative research.

The bibliometric technique deals with what is called a variable. A variable in this sense can be of three types, viz., producers, artefacts and concepts. The category of producers includes, authors, publishers, corporate entities, institutions, countries, etc. Artefacts refers to information products, such as, books, articles in journals, news items, written speeches, conference papers, etc. Concepts include all that the written documents contain, that is, words, chapters, quotations, citations, etc.

We may list one or two of these to show the variety of variables, that encompass a single term called monograph. The term covers all the following: atlas, manual, thesis, proceedings, statistics, treatise/laws, dictionary, encyclopedia, pamphlet, collection/anthology, festschrift, directories, bibliographies, report, etc. This clarity of the terms is basic to the scope of bibliometric studies.

Thus there are many ways in which bibliometrics is useful and we can add on to this list some more, like

a) results of bibliometric studies help in evaluating an organization, in terms of its research programmes, government policies, image in the academic community;

b) useful for evaluating an individual journal, by analyzing its citations, and the results are useful in ranking core journals—which helps in selection process and acquisitions;

c) evaluation of countries production, helps in identifying policies, indicating level of resourcefulness, etc;

d) useful in studying the impact factor of individuals, publications, countries in different subjects, and thereby enabling the mapping of the respective fields.

14.2.3 Scope

The scope of the term bibliometrics covers literature in documents, published and unpublished, printed or in other media, as follows:

a) form: book, journal articles, theses, multimedia.

b) body responsible for the production, as well as transmission, distribution of the document, like authors, institutions, their affiliation, subject interests, geographical origin, preferred forms of publications.

c) subject covered: single discipline, multi-discipline.

d) language of the document, single or multiple language.

e) date of publication, frequency of publication.

f) size of the document, amount of information.

g) place of publication, distribution limits, if any.

h) text of the document, citations in the document, citing behavior, citing pattern, citing needs.

These factors, or precursors, are also called as variables, which provide the basis for study and research in the field. By using these variables, bibliometricians, provide a support directly or indirectly in development of information base. Thus bibliometric results assist in improving bibliographic control, identifying the interrelationships for improving secondary services, identifying most cited sources, preferred sources, and possible bias in the research. Similarly, citation use and citing pattern give useful information about subject, language, country relationships.

14.2.4 History

The practice of using quantitative method to measure information sources existed since long. For instance, F. J. Cole and Nellie Eales, in 1917, graphically mapped the literature and called this as 'Statistical analysis'; E. Wyndham Hulme in 1922, studied the literature and called it 'statistical bibliography'; S.R. Ranganathan in 1948 called it librametrics, while in fact he had been applying the elements of statistical calculus to the library problems as early as 1925. This implies that quantitative approach to study the literature or the library activities is also included in the term bibliometrics.

Why was this called statistical bibliography or librametry? The answer lies in two different perspectives: one, is the that anything relating to library or literature was traditionally connoted by the term bibliography. The other is, the method that was applied referred to measurement and also application of mathematics and statistics. That is it was a quantitative or rather technically a statistical analysis of the growth of biblio (or books) by Coles and Eales, Hulme, and others; S.R. Ranganathan, being a mathematician, made use of his mathematical knowledge in measuring the work done in the library, and he coined a direct name.

Statistical bibliography as a term did convey some sense. It had two definitions, "to shed light on the processes of written communication and of the nature and course of development of a discipline (in so far as this is displayed through written communication), by means of counting and analyzing the various facets of written communication"; and "the assembling and interpretation of statistics relating to books and periodicals...to demonstrate historical movements, to determine the national or universal research use of books and journals, and to ascertain in many local situations the general use of books and journals."

Bibliometrics has today gained a place and is understood to deal with quantification of written communication—books and other media. This implies that counting or measuring the frequency of the literature are the basic two ways. By an extension of the term, we are still with Ranganathan's coined term, librametry, because bibliometric studies do include statistical approaches to the study of library and its services.

14.3 BIBLIOMETRIC STUDIES

14.3.1 Relationship with Scientometrics and Informetrics

The study of growth patterns of literature started formally by the publication of the work of Cole and Eales in 1917. According to one estimate today there are thousands of publications that deal with the subject. The field has gradually diversified in different areas. Some major sub-fields are econometrics, informetrics, scientometrics, etc.

Scientometrics, deals with evaluation of factors, like scientific progress, levels of scientific development, social relevance, science of science, and science policy. Econometrics, focuses on, study of economic factors, economic indicators, economic theories, and all aspects relating to economics. Informetrics, refers to, analysis of information productivity patterns. It is a field which brings into focus the concept of 'organized complexity' of the Info Society. To promote this activity FID's committee for Informetrics was set up in 1980.

Study of the literature, or written communication, then has advanced considerably. As we see in the following sections the field has its own laws, and has become a major area of research interest, next only to automation. While all studies use numbers and counts, bibliometrics are distinct in using social science methods in quantifying the literature, using citation data to determine scatter phenomena, using citation analysis to identify the research fronts of a field, and in understanding the structure of subjects.

14.3.2 Sources of Bibliometric data

Counting the literature has many advantages and hence it depends on some sources for getting data. Then what are the sources of getting the data? The sources which provide data are: journal articles, books, patents, theses, etc. These are either in printed or computerized databases. Among these, the articles in journals and books provide much of the data (variables) required for bibliometric analysis.

In this sense, bibliographies and secondary services are a better source for this data. Bibliographies are the most preferred source for data, as these are easily available, cost less and are also comparatively comprehensive. Bibliographies include most of the bibliographical details, and also provide additional material, like number of references, abstract, etc. Indexing and abstracting sources also are comprehensive and cover the required variables for bibliometrics.

Among other sources of data, theses are valuable. But these too are mostly not comprehensive in providing bibliographical details of each cited document. Library catalogs are another useful of data. Apart from journal articles, journals are also a source for data collection, in the sense, that one can find measurable data about age of the journal, subject, geographic and institutional affiliation.

For this study there is also need of other information sources, like directories, biographical tools, encyclopedias, bibliographies, conference/seminar proceedings, news items, review essays, list of forthcoming conferences, letters and correspondence, illustrations, advertisements, etc. To explain one of these is appropriate. We need a biographical directory for various reasons, like author's address, biographical details, list of publications, author's past and present affiliation, achievements, awards and honors, etc. This information helps in finding the links of the author with the text, with the cited sources, with the subject of the text, language of the text, country and date of publication of the text, and so on.

An understanding of the sources also calls for a basic awareness of all the above tools, in their physical form as well as other allied details. In the physical form of a source, a bibliometric study deals with either text, citation or bibliographical details. In the context of text, one has to be aware of the subject, language, etc. To know the subject, one must be aware of content analysis, indexing and vocabulary control mechanisms, in the respective subject.

For instance, in a bibliometric study of word frequency one must know what words belong to the subject, what words are from related areas and what words are from unrelated areas. This knowledge helps in grouping and classifying the text under study. Each form has a specific place in the dissemination of information, and the use of these as variables, necessitate the understanding of physical form too. For example, one must know the difference between a book and a pamphlet. According to the guidelines given by UNESCO a book, is a non-periodical printed source, must at least have 49 pages. Whereas a pamphlet is a small work, which is again according to UNESCO must be between 5 and 48 pages. Such is the wide area of data sources.

14.3.3 Applications

It is appropriate to say that basic bibliometric techniques are of two types: productivity count and literature usage count; dealing respectively with a) books, articles, words in a text, place of publication, subject matter, time and date of publication, publishing institution, authors, author's institution, etc.; and, b) citations in published works, circulation, frequency of borrowing or browsing different library materials, failure and success in search strategies, search options, search strategies, etc.

This categorization suits the definition of bibliometrics given by Daniel O'Connor and Henry Voos, that the 'scope of bibliometrics includes studying the relationship within a literature (e.g. citation studies) or describing a literature." Similarly the authors of most often cited source, Nicholas and Ritchie in the book 'Literature and bibliometrics,' divided these two categories as those dealing with productivity count (or descriptive) and literature use (evaluative).

Frequency is the catch word in bibliometrics. Depending on the frequency of a particular result, the utility of the data can be decided. For example, the frequency of the use of a textbook is most in an academic library. This data may be of use in a number of ways, such as, collection promotion, collection development, collection maintenance and in withdrawing the other less used items. Then frequency approach to data collection is of value. But there are critics of this approach who feel that the frequency counting method itself is not always foolproof. They feel that any data drawn from one source is likely to be imbalanced, and hence all such data must be counter checked for validity. Similarly, there are critics who question the utility of quantification itself.

Quantification as a tool for research did exist since long, but its independent use, and use without the support of the long known technique of qualitative approach has led to much debate. The research of today finds it feasible to combine these two methods. Bibliometrics has also found one means of using the qualitative field, in its citation analysis, and within this, the impact factor studies. We may come to these later.

The field of bibliometrics grew due to many significant factors, like, literature explosion, emerging information society, ever increasing user needs, growth of disciplines and sub-disciplines, collection developmental pressures, continuously shrinking budgets, a growing sense of professional responsibility among the information managers, space shortage, advances in storage and retrieval systems, etc.

Most of these are interrelated. For instance, literature explosion is directly linked with growth of new subjects, advances both in the production, and the means of storage and retrieval of information, etc. Literature explosion, or explosion of information and explosion of the published material (which is not in any way explosion of knowledge) has occurred due to many reasons. Significantly, the twentieth century is called the information age. Ability to publish decides the ability to survive, and by implication to the theory of survival of the fittest.

Literature explosion is an other catch word for many studies. In an interesting study, by Attia on explosion of literature aimed at addressing three facets of this modern phenomena: a) the concept of information dimensionality, b) determination of common parameters to measure the amount of information within each dimension; and c) a working definition of 'explosiveness." The preliminary findings of the study reveal that there has not been an explosive increase in the information production area, whereas there is an explosive growth in the information distribution. Regarding the information flow the findings reveal that there are explosive and unexplosive increases as well as unexplosive decreases at different points for most of the variables.

This literature explosion has led to a major problem of book selection. The post World War II era has totally changed the scenario, from the previous era of no selection at all, to detailed book selection policies and guidelines today. This means much of experimentation has gone in and today bibliometric is getting acceptance as means of providing support in developing guidelines for selection of materials, as well as for other more important activities, like storage, processing and weeding of materials in a library.

14.4 BIBLIOMETRIC LAWS

The three laws that have emerged in this field are as follows:

Bradford's Law of Scatter

Zipf's Law of word frequency

Lotka's Law of author productivity

14.4.1 Bradford's Law

Of the three laws, Bradford's law has received greater attention in literature of library science. Bradford's Law of Scattering was promulgated by the British bibliographer, S.C. Bradford, interpreter of UDC. His theory dealt with the scattering of the literature of a discipline and its consequences, in his now most cited work, entitled: Documentation. His concern was with the twin problem of overlap and omission in the coverage of primary journal articles by indexing and abstracting services. Bradford's main concern was to identify the levels of scatter or dispersion of publications of one field into other fields. So much was this scatter and so large was the spread over of the literature that this work of Bradford opened the eyes of many, who had not really taken serious note of this matter.

Earlier, in another study, Bradford analyzed this pattern of dispersion of literature on a subject in various periodicals, in a paper on applied geophysics and lubrication. He found in this study that 9 journals covered 429 articles, the next 59 journals covered 499 articles, and the next 258 journals accounted for the 404 articles. This implied that the first nine journals contribute one-third (about 33%) of the articles in a subject, the next 589 journals accounted for another one-third (33%), and the remaining 5 X 9 X 9 journals for the last one-third (33%).

That is to say, periodicals are usefully categorized in three groups:

i) those periodicals which carry four references in one year, in a particular subject;

ii) those which carry between two and four references in a year, and

iii) those which carry one or fewer references in a year.

The first category forms the nucleus of periodicals in a particular subject; this becomes therefore the most useful group with most productive and relevant titles of periodicals in the concerned subject. This nucleus is visible in the law which Bradford stated as 1:n:n2

Then according to Bradford, the first nucleus of the one-third (or around 33%) of the titles, in a subject are the most highly productive title, and these are the core journals. Core journals make up the most of the required material, and hence, if identified using proper bibliometric methods, proper data sources, and counter-checking the bibliometric results with other sources will provide enough basis for decision making in acquisition of library materials.

A library can decide to subscribe these only, and for the other required titles, inter-library loan, xerox supply service, etc., can be tried, as there is a shortage of budgets, there is always an increase in the prices of journals, increasing user needs, increasing of new subjects, etc.

Since that date of the Bradford's prediction of the scientific publications many bibliometricians have improved on this statement and given different perspectives. The refinement of the law was attempted by B.C. Vickery. He found discrepancy between the verbal and graphical representations, and restated the mathematical expression. He was one of those pioneers who pointed out that the application of the law should not only be to the three zones which Bradford had identified, but with suitable modifications of the values of the ratio n, to any number of zones. In terms of practical applications of the law, it was F.F. Leimkuhler who initiated the process, and B.C. Brookes and other bibliometricians took up this task further.

Of all the areas in which this law is applied, the most prominent is the study of journal productivity, most specifically in science. Core titles are derived on the basis of either numerical or verbal analysis.

14.4.2 Lotka's Law

While Bradford was concerned with the scatter of publications, another scientist, Alfred Lotka studied, in 1926, the frequency of publications of authors. He was concerned with the number of papers published by scientists, and desired to determine the level of contributions by scientists of different caliber in the progress of science.

In his study of the literature Lotka found that the productivity of scientists confirmed to inverse square law such that for every 100 authors contributing one article, 25 will contribute 2 articles, 11 will contribute 3 articles, and 6 will contribute 4 articles, etc. The observations for author's with single articles were 57.9% for Chemical Abstracts (6,891 contributors) and 59.2% for the Physik Data (1,325 contributors). The law as stated by Lotka is "the number of chemists publishing n papers is proportional to 1/n2 in each," resulting in authorship of a large number of documents by a small number of authors became obvious.

Bibliometricians have further studied this pattern in other areas, and find that this law is applicable to the subjects of history of technology, map librarianship, etc. They suggest that there is a need for alteration in the consonant 1/n3.5 for a field like library science;

This led to the theory of success breeds success. This means that an author of many papers is likely to publish more, as compared to one who is less productive. So also, a journal which is used more frequently is more likely to be used again, than a less frequently used journal; and in this same sense, an article which is highly cited is likely to be cited further, than an article which is rarely cited. Derek de Solla Price developed his own theory that "half of the scientific papers are contributed by the square root of the total number of scientific authors."

Price's general theory of Cumulative Advantage Processes require a brief explanation. Cumulative Advantage Process operates in situations where success breeds success. This theory proposes a statistical model that predicts the distributions they produce. These distributions fit those derived from Lotka, Bradford, Zipf and others like the Pareto law of income distribution. Incidentally, this also fits in the empirical results of citation frequency analysis. It has really a cumulative advantage, and in a sense, it unites all the laws and citation data which come under the umbrella of bibliometrics.

On the same line of thinking, Ravichandra Rao argues, that those documents which have been frequently borrowed are more likely to be borrowed again, than those least borrowed in a library.

14.4.3 Zipf's Law

As against the above two laws, dealing with scatter of publications and author productivity patterns, Zipf undertook yet another area of counting the frequency of words in a text. He ranked the most frequently appearing words in a descending order. He analyzed the works of fiction and found that the product of a word (r) and its frequency (f) was a constant. For instance, the tenth most frequent word (r=10) occurred 2,653 times (f= 2,653), the hundredth word (r = 100) occurred 265 times (f= 265), the two hundredth word (r= 200) occurred 133 times (f= 133), etc., which means that he listed the ranked words in descending order of frequency, and multiplied the numerical value of each rank (r) with its frequency (f) and arrived at a product (c).

Stated in his own words, the Law is "that if words occurring in a natural language text of sizable length were listed in the order of decreasing frequency, then the rank of any given word in the list would be inversely proportional to the frequency of occurrence of the word."

Zipf's equation is $rf = c$

(where r and f are, rank and frequency of words, respectively and c is a constant).

In this law, it is clear that Zipf made a word frequency analysis in a text. He derived the law from a general principle of "least effort". That is, words whose cost of use is small or whose transmission demanded least effort are more frequently used. The law has been successfully tested in different fields and the results have shown, among other things, the richness of an author in vocabulary.

For explaining this word count technique a search of this text on bibliometrics was conducted. It showed 9912 words in all; the decreasing order of terms and names (excluded are other words — though Zipf would include all occurrences), are listed as a sample of the occurrences:

Citation (variants included)	57
study (variants included)	37
author (variants included)	24
bibliography (variants included)	19
many	18
etc.	18
bibliometric	18
article (S)	17
bibliometrics	16
Bradford (variants included)	15

Before we end this discussion of the laws of bibliometrics, it is appropriate to state that of all these three, Bradford received the greatest attention and has still much appeal to all. Attempts are also made to show the inter-relationships between the laws. For instance, Abraham Bookstein, analyzed the three laws and states: "Both Zipf's and Bradford's laws emphasize the higher-ranking entities. Lotka. on the other hand, begins with the low end." (Hertzel: 190)

14.5 CITATION ANALYSIS

Citation directs our attention to something already existing, that is, it means to cite, or quote or refer. It is a specific reference to a work from which the text is quoted, or showed its source as an authority for a statement or idea. In one way the citing author transfers the authority about an idea to someone else.

One may ask why do authors cite? The answers can be many: as a reference for further study, acknowledge the quote, appreciate a related work, suggest another work for comparison, prove a statement, and so on. In fact all this may indicate that there is a similarity between a text and the references that are cited. And this supposed similarity has led to many studies.

By this method of analyzing the cited list of references one can evaluate the citations received by "articles, authors, institutions and other aggregates of scientific activities" says Ravichandra Rao. Two clarifications are essential at this stage. One relates to evaluating the citations received by an institution. By this it means that we assess the publications of the institution, as well as acceptability of its publications, which is very easily visible in the cited sources.

The other clarification is of the terms citing author, cited author, and citations received. Citing author is the author of the text, cited author is the reference quoted in the text, and citations received refers to the number of times an author is quoted or referred. Nicholas and Ritchie, provide more clarity to other terms, "A distinction must be made between a reference made by an item (to another item) and a citation received by an item (from another item). References are, of course, also citations: the distinction is according to whether the source or citing item is, being considered.

Study of the citations has been the most active area in bibliometrics. Analysis of the relationship between the cited and citing document is one of its main purposes. The first attempt to find this relationship was, in the form of citation counts, by Gross and Gross in 1927, who produced a list of journals that were highly cited by the Journal of American Chemical Society. In 1961 Eugene Garfield started the Science Citation Index, and later fulfilled the other areas: Social Science Citation Index, and Arts and Humanities Citation Index. Garfield took up another study to find the impact factor of citing and cited documents. Kessler initiated the study of bibliographic coupling, Small took up the study of co-citation, and the area is progressing. A study of citing pattern, by the National Enquiry reveals that there is an increase in the number of references in each text. The reason for this increase could be that more secondary sources are providing access to citations, and that has directly resulted in increase in the access to literature.

The same Enquiry also found, that "Citation references in journal articles are preponderantly to other publications within the field. Nearly all citation references made in classical studies are to journals in classics. Corresponding figures for the other disciplines are English, 84 per cent; philosophy, 82 per cent; sociology, 74 per cent; and history, 68 per cent".

A word can be said about the relationships that is in the document, as well as, with other documents. The inner links are many, like the link of the text with the references of a previously existing document. A link with the external source is seen in different ways, such as the influence of citing and cited documents on each other, the author's sources of data in the document; similarly there are links in the place and time of the publication. It is right to quote Nicholas and Ritchie, in this context of relationships: "All documents rely to a greater or lesser extent on information contained within previously published documents. Inside the literature there exists, in fact, a web of relationships. By exploring the full bibliographic ramifications of these relationships (by matching, for instance, the bibliographic description of the host publication with that of the cited publications), the nature and strength of each relationship can be determined (pp.10-11)."

A lot of survey and studies have been done in the area of citation analysis. Basically these are of two types: studies dealing with productivity of science literature; and, studies reflecting the use of the literature. Wherein studies based on counting entries are from the productivity category. The techniques used in analysis of citations mostly are obsolescence, clustering, citation indexes, etc., and these are discussed below.

14.5.1 Obsolescence

Basically stating it as use-less, this term obsolescence also means no life or half-life documents, or no longer in use. The half-life of a document in a subject is that period during which half the literature used today was published. That is, if one is analyzing half-life of the literature in psychology, all documents published in the subject in the year 1996 are its base. A survey of use of the publications of 1996 and the previous years is to be conducted analyzing the annual use of literature in the subject. The results may show that half of the literature used in these libraries during 1996 was published during the past ten years. In such a case half-life of the literature of this subject is ten years.

While all bibliometric studies try to find the utility of publications, this technique of obsolescence has been specifically used in confirming or rejecting the idea that document use or value declines with its age. That is, a document may become of least use after some time of their appearance for reasons, which may include, ephemeral nature (e.g., pamphlets have a short life span), outdated information (like, advances in the field), decline in the general interest (for instance, new subjects grow and old decline regularly), etc. Then obsolescence refers to no longer useful or reliable documents. This is what a group of bibliometricians predict.

All these predictions are based on the idea that decline in use is related to the growth of literature. Nicholas and Ritchie are of the opinion that "Recent research would seem to indicate that the 'decline' and the reasons for it are exaggerated or misinterpreted. Two factors have combined to create this situation. The first is that decline has been exaggerated by intense and multiple purpose use of recent literature (up to two years old)... The second factor is that few studies have taken account of growth of the literature when calculating supposed rates of decline".

There are others who are cautious and do not declare any document as dead or useless, unless it satisfies many more tests. While Maurice Line defined half-life as "half the active life," this was interpreted as "meaning the time during which one-half of the currently active literature was published." This flexibility in interpretation and deriving what actually half-life or obsolete means, has lead another group, which includes Brookes, who refuted the Line's approach, and concluded that it was "both questionable and impractical". He, rather stressed on the point that "the rate of obsolescence is a function of both the subject literature and of the local usage of that literature" (Hertzel, p. 192-193).

14.5.2 Clustering

This technique refers to the use of grouping or creating families of each category of documents. By creating a network, or cluster, of related documents, bibliometricians, like Mark P. Carpenter and Francis Narin, who found that "the process (clustering process) used to divide sets of journals into subject areas has two underlying assumptions: first, that journals which deal with the same subject area will have similar journal referencing patterns; and second, that journals which deal with the same subject area will refer to each other" (Hertzel, p. 196). Nicholas and Ritchie state: "As each document has a number of relationships, some of which are not reciprocated to the same degree, a very complex multi-dimensional network is created. From this a 'cluster' or related documents can be defined, free of prejudicial classification. Once clusters are identified subject descriptors can then be allotted".

Bibliometricians have also worked in other areas, which include co-citation, bibliographic coupling, self citation, etc. Bibliographic coupling, looks similar, but there is a difference. Bibliographic coupling links source documents and is a fixed measure for any number of coupled papers. Two works are said to be coupled bibliographically if they have a common citation. Here it links source documents.

Co-citation indicates the extent to which a citing author perceives a pair or more papers as related. Co-citation is the frequency with which two citations from earlier publications, are cited together. All these, in short, try to build networks to measure the strength of relationships. Co-citation also enables mapping of science. This activity of mapping a field using citations has been greatly facilitated by the citation indexes.

14.5.3 Mapping of Subjects

The structure of any subject is a difficult [illegible], as is the understanding of the universe of knowledge, growth of new subject, etc. Garfield thus states, the same more specifically in the context of the structure of science: "The structure of science is an intriguing subject to those who study science as a system. Beyond the intuitively comfortable hypothesis that science is a mosaic of small units, rather than a structural monolith, there are many more questions than answers. What is the nature of the basic units in the mosaic structure? How do they relate to each other? Are the intellectual and social structures similar? Are they made up of the same basic units? What is the relationship between them? Is there a variety of configurations at the infrastructure level? (Garfield: 1979: 98).

Co-citation and clustering have resulted in what is called mapping the structure of subjects—macro and micro. Specialists like Small, Griffith, Price, Narin, Goffman, Jahn, Garfield, etc., tried to study the relationships in citations through these networks, and have attempted to map the structure of specialities, to start with.

The mapping of the structure of disciplines is attempted both in natural and social sciences, and sciences include physics, biomedicine, cancer, etc. and the social sciences, includes, psychoanalysis, sociology, etc. To quote from Garfield, again, these "represent the most sophisticated attempts made... to use citation analysis to define the structure of science on the scale and at the level of detail needed for science policy purposes. On the one hand, their method seems capable of depicting all the major scientific specialities in a single, coherent structure. On the other, it offers a range of resolution broad enough to permit the detailed examination of any substructure level that is appropriate to the questions an investigator may choose to pose. This combination of scale and variable resolution produced a functional capability, for analytical purposes, that is analogous to an automated design system in which a computer is used to display a product at all production levels, from final configuration down through the entire hierarchy of various sub-assemblies.

14.5.4 Citation Indexes

What are these indexes, and how are these different from other indexes? The answer lies in the point that "A citation index is an ordered list of cited articles each of which is accomplished by a list of citing articles. The citing article is identified by a source citation, the cited article by a reference citation. The index is arranged by reference citations. Any source citation may subsequently become a reference citation." (Hertzel, p. 194).

The traditional indexes required highly trained indexers; whereas in citation indexes the test of an article provides the indexing through its citations. A clarification is necessary to distinguish between source citation and reference citation.

i) Science Citation Index (SCI)

This is published as a five part index. The two key parts are the Citation Index and the Source Index. The other parts are the Patent Index, the Permuterm Subject Index, and the Corporate Index. Citation Index links items published during the year with past items cited in references. The Source Index is an author index to the items published during the year, and it is also organized alphabetically by surnames of authors. The Patent Index, specifically, deals with patents. The Permuterm Subject Index is a title-word index. The Corporate Index takes care of the journals published during the year listing organizations with which the author's are attached, or affiliated. Each organizational entry displays the articles that the staff published during the year.

Citation indexes led to the development of advances in citation studies, like Journal Citation Reports, impact factor, etc.

ii) Journal Citation Reports (JCR)

Journal Citation Reports are part of the Science Citation Indexes(SCI), and deal with citation counts and links of the literature. JCR provides the following data on the source journals covered by SCI:

a) how often each journal is cited

b) how many items it published

c) how often (on an average) each item is cited, which is called impact factor

d) how often (on an average) each item is cited during the year of its publication, which is called immediacy index

e) the source journals responsible for the references to each journal, the number of references received from each, and how they were distributed by the publication years of cited source.

f) the number of references each journal published, to what journals, and how the references were distributed by the publication years of the cited issues.

By identifying what the citations are, and who is citing, JCR studies what journals cite which journals, and with what frequency, and defines the core and non-core of the literature on any subject.

14.5.5 Impact factor

This refers to the impact of a journal, not of an article; when a journal is analyzed for its productivity, the total number of items published by the journal influences the number of times it is cited; the more it publishes, greater would be the number of opportunities it has of being cited. This is evident more so in large-sized of journals, than in the small ones, as a large journal has more articles. To get a correct impact factor, the analysis must show the average citation rate per item in a journal. This is derived by dividing the number of times the journal has been cited by the number of items it has been published.

14.6 LIBRARY USE STUDIES

Bibliometric studies have been useful in different ways in libraries. Libraries need feedback and also need constantly information about book selection, evaluation of services and facilities, etc. And in many of these areas bibliometric analyses has been found effective. Three main areas can be identified in which many research studies have been taken up: book and serial acquisitions, library use patterns, and in matters relating to decisions about less used materials, that is weeding of documents.

Many studies, have used journal articles for identification of core journals. These have applied Bradford's law and for micro and macro subject areas, like superconductivity, surgery, biochemistry, biophysics, cancer, religion, etc. While the technique is yet to receive full acceptance, nevertheless, citation indexes, impact factor, and other improved measure have found some acceptability of bibliometrics. Libraries are trying to convince the users of the validity of such results.

Under each aspect it is necessary to mention one or two examples, to illustrate the exact meaning and scope of the bibliometric application. To begin with, we can see the impact of bibliometrics in identifying a list of core titles, using the Bradford law.

Mohammed Taher studied in his doctoral work the subject of Islamic studies, and aimed at identifying, among others, core journals. The study used the verbal suggestion of Bradfordian law, which prescribes the three zones of a literature, and categorized the literature of the field into the respective zones: the most relevant, less relevant and least relevant sources. The study, gathered data from a printed bibliography, having international coverage, for 1980s in English. One thousand and seventy three(1073) journals were indexed and this accounted for 24 titles (in first zone), 118 (in second zone) and 931 (in third zone). This means 24 titles are capable of satisfying the majority of the user needs in the subject. As against this verbal analysis, a statistical analysis of the same data, identified 13 journals as core titles. This difference in results, was probably due to statistical measure, and the like.

I.N. Sengupta analyzed the literature in medicine and allied fields, with an aim of identifying core journals. The data was gathered from the references provided at the end of each article of Annual Review of Medicine, for the years 1965-1969. The results showed that 35 journals, out of 975, accounted for 52.8% citations. These 35 have been identified as core journals in the field.

Bradford's law has been, as said earlier, applied in journal citation analysis only. So also Lotka's law is yet to find other areas of application. Ravichandra Rao and Burhanuddin's case studies are relevant in this aspect, to measure the author productivity as seen from an interpretation of Lotka's law. Ravichandra Rao combines both the Bradford and Lotka laws in one study, and hence its importance. No study, on the application of Zipf's law was found in the literature, the reasons being that the study demands much hard work, if done manually and the labor that is put in this work gives less results to the researcher.

In the area of library use patterns, there is less of criticism, and even lesser number of studies, in indicating the efficiency or otherwise of library services and facilities. Ravichandra Rao in his doctoral study analyzed use behavior in a library. The study focused on six automated Canadian university libraries, aimed at identifying a probability distribution of transactions over documents and also over users. The results were, that most size-frequency data of transactions follow a negative binomial distribution, which by implication is a manifestation of success-breeds-success phenomenon; the rank distribution is likely to follow Bradfordian distribution; frequent users were more likely, than infrequent borrowers, to borrow documents which were borrowed frequently; and users in the area of arts were likely to borrow documents, irrespective of whether the documents were borrowed frequently, or not.

Mohammed Burhanuddin in his doctoral study assessed the author productivity of the Indian writers in American studies. The study analyzed the data for the years, 1969-1987, and identified 856 serials, and among others, studied the availability of these in a leading, also representative library. The study found that 72.2% percent of cited items are available in the library. The same study also checked for the availability of other cited items, such as, 1410 cited books, the availability was 71.9%; 24 reports and unpublished items, and the availability was around 20%. These figures then indicate that published materials, such as books and journals are sufficient, but reports and unpublished materials are highly inadequate in the library surveyed. The same study evaluated the dispersion phenomenon in this field, and found that: out of 325 cited journals, 33.5% or 18 titles were enough to be identified as core journals of American Studies. This represented the first zone; the second zone has 80 journals and third has as many as 227 journals.

These sample studies are illustrative of the type of use one can make of bibliometric method. In case of library use, while many studies try to conduct qualitative experiments, much effort has not been done to measure the efficiency of library services using quantitative measures and bibliometrics. So also weeding and improvement of library require much attention by the bibliometricians.

14.7 INDIAN CONTRIBUTION TO THE DEVELOPMENT OF BIBLIOMETRICS

In terms of growth of this subject, India has made many attempts and can find a place as an important contributor among the world's prominent bibliometrically advanced countries. U.S., U.K, Canada and others.

India, has to begin with a FID Committee on Informetrics, established in Delhi, around 1985, which also promotes bibliometrics. Many books have appeared dealing with bibliometrics, including those by I.N. Sengupta, I.K. Ravichandra Rao, B.M. Gupta, S. Subba Rao, Mohamed Taher, etc. There are at least dozens of Indian scholars who have published their contributions in Indian and international sources, to name a few: B. Maheshwarappa, B. Guha, R. Shalini Urs, B.K. Sen, M. Mahapatra, K. Meghanatha Reddy, K.S. Ragahavan.

Coming to teaching this subject, it is a full paper, in some universities, and is a part of research methods in some universities. Whereas some universities have not yet updated their syllabus, and hence have failed to include quantitative and bibliometric methods in the programme.

Extensive bibliographies have appeared in Indian library science periodicals, like the IASLIC Bulletin, Annals of Library & Information Science, etc. IASLIC had its seminar on this topic in 1985, DRTC had its first All-India bibliometric conference in 1969 (as the theme of its Annual Seminar), again, DRTC conducted Refresher Seminar in 1981, and there have been many more programs conducted all over the country.

Indian bibliometricians do regularly attend the international bibliometric conferences, and the recent meet held at Chicago had many Indians participating in the programme. And as a part of the exchange of ideas program, bibliometricians from other countries have continuously visited India, and those include: Leo Egghe, F.W. Lancaster, and others. Indians who are now in the United States, and other countries have also contributed to the development of the field, prominent among these are Chandra Prabha and K. Subramanyam, who are living in the U.S., and have published extensively.

14.8 DIFFICULTIES IN BIBLIOMETRIC STUDIES

Like all evolving fields Bibliometrics has to find ways of getting acceptability in the universe of knowledge. At present there are scholars who do not feel very easy with this quantification technique, some outright refuse to give any weightage.

Take for instance, a quote from Carlos Candra, relevant in the context of this other side's view: "no amount of quantitative or mechanical manipulation of texts, bibliographies, or other tools will enable one to identify the key contributions to information science (Hertzel: 198)."

There are many problems in using bibliometric method, such as, determining the exact meanings of definitions that are in use (terms like monograph, journal, newsletter, house journal, etc); identifying a source which can provide maximum data for analysis (as many studies rely on indexing and abstracting services); finding a comparable source, as are existing in the developed countries, for analysis of the literature in the developing and underdeveloped areas; combining the results of bibliometric study with other research methods to ensure the validity of the results, etc.

There is much debate on the validity of these studies. On one side the critics of bibliometric studies are not ready or do not care to accept the bibliometric results. And for different reasons, library users do not accept the decisions, taken by librarians to add or drop subscription to journals based on library related studies. A list of core journals will provide maximum required items with minimum cost and effort. This is taken as a convenience, than a strategy to face the problems.

One point deserves clarification to highlight the hidden drawback of this technique. Bibliometric studies are focused on citations, that is use of documents. In this focus however, all use of books without citing it in the text, remains unnoticed. Shelf users do not always cite all that has been browsed, referred at the shelf, or even those books that are read all the time. This is to say that readership survey is also an equally important method, which can supplement and compliment the citation study, but citations as such are not the sole indicators of use.

Regarding the citation studies, specifically about data gathered on highly cited titles, one view is: "citation lists do not indicate the least-used titles in a particular library, which is the information needed by librarians to determine which subscriptions to discontinue. The correlation between journals least cited and journals least used is usually less strong than the correlation between journals most cited and journals most used in a library (Martyn and Lancaster, p. 196)."

Garfield has listed many objections that are raised by critics of citation studies. Such as, " A person's rate could conceivably be inflated by self-citations. A paper may be cited frequently in refutation or as a negative example...there is the problem of sloppy, and even biased, bibliographic citations...Not everyone cites all the obvious, classical antecedents or is conscientious about citing all the sources actually used. Not everyone conducts an exhaustive literature search or uses all sources that should have been used. Not every one limits references to only material that was actually read (Garfield: p. 63)."

While objections raised by one group of critics deals with what citation counts measure, another group deals with what citations do not or cannot measure. And the debates go on to similar other issues. And Garfield, clarifies some of these: "Any fair appraisal of citation analysis as an aid in evaluating scientists must acknowledge that there is much about the meaning of citation rates that we do not know. We are still imprecise about the quality of scientific performance they measure. We still know very little about how sociological factors affect citation rates. There is still much uncertainty about all the possible reasons for low citation rates. And there is still much to learn about the variations in citation patterns from field to field (p. 249)."

It is hoped that the field will find a right place and be as helpful and positive as, Hertzel feels: "At present it is composed of methods and techniques without a coordinating theory; but it seems possible that as more theoretical information is converted to fact, and behavioral patterns are established for the assessment and evaluation of the structural components of literature (i.e., authors, publications, words, laws) and their relationships, the more probable causal explanations will be evident and bibliometrics will be closer to being recognized universally as a science (p. 199)."

14.9 LET US SUM UP

Let us recapitulate briefly what has been discussed so far in this unit.

* Bibliometric technique is an important area of research. In the words of Hertzel, "it seems evident that bibliometrics, which was first used in attempts to evaluate journals for collection development in libraries, has been recognized and expanded to the study of the structure of literature in the larger encompassing field of information science".

* The Bibliometric laws are applied to various use studies in libraries and information centres. They are also helpful in mapping of the structure of different disciplines.

* Citation analysis studies, besides helping us to rank journals in a subject area, have also opened up many issues and problems, by providing measures for analyzing the patterns of growth and productivity of literature .

14.10 REFERENCES

ATTIA, A. H. Mohammed. "Application of information theory concepts in the investigation of the growth pattern of production, distribution, and velocity of information". *Ph.D.thesis*. University of North Texas, 1987. (DAI: 48/08, p.1918, Feb.1988)

BROOKES, B. C., "Numerical methods of bibliographic analysis" *Library Trends*, 22, 1973, 18-43.

EGGHE, Leo, "The discipline 'bibliometrics': Its origin and present coverage" (Preprint)

FERRANTE, B.K., "Bibliometrics: Access in the library literature" *Collection Management*, 2, 1978, 199-204.

GARFIELD, Eugene. *Citation indexing: Its theory and application In science, technology, and humanities*. Philadelphia: ISI Press, 1979.

HERTZEL, Dorothy, H. "Bibliometrics: History of the development of ideas," (IN Encyclopedia of library and information science. Edited by Allen Kent. New York: Marcel Dekker, 1987. vol. 42 (Supplement 7), pp. 144-219.

HJERPPE, Roland. *Bibliography of bibliometrics and citation indexing and analysis*. Stockholm: Royal Institute of Technology, 1980.

An outline of bibliometrics and citation analysis. Stockholm: Royal Institute of Technology, 1980.

MAHAPATRA, M. "Inclusion of bibliometrics in library and information science in India" *IASLIC Bulletin*, 30, 1985, 175-179.

MARTYN, John and Lancaster, F. Wilfred. *Investigative methods in library and information science: An introduction*. Virginia: Information Resources Press, 1981.

NATIONAL Enquiry into Scholarly Communication. Scholarly communi-cation. *Report of the National Enquiry*. Baltimore: Johns Hopkins Univ. Press, 1979.

NICHOLAS, D. and Ritchie, M. *Literature and bibliometrics*.London: Clive Bingley, 1978.

RAVICHANDRA Rao, I. K. *Quantitative methods for library and information science*. N w Delhi: John Wiley, 1983.

———— "Two essential components of information science: quantitative methods and automation." *DRTC Annual Seminar*, 1984, Paper EA.

SEN, B. K. and Narendra Kumar, "Indian contributions in bibliometrics, 1958-1984." *Annals of Library Science and Documentation*, 33, 1986, 85-103.

SENGUPTA, I. N. *Bibliometric research, growth of biomedical literature*. V.1. Calcutta: SBA Publications, 1988.

———— "Bibliometrics and identification of core periodicals (Kaula Endowment Lecture)" *Herald of Library Science*, 29: 3-4, 1990, 226-251.

14.11 RECOMMENDED BOOKS

BORGMAN, Christine L. *Scholarly communication and bibliometrics*. New Delhi: Sage Publications, 1990.

GARFIELD, Eugene. "Current Comments: Scientometrics comes of an age". *Essays of an information scientist*. Vol. 4 (1979-1980). Philadelphia: ISI Press, 1981. pp. 313-318.

HANDBOOK of libraries, archives and information centres in India/edited by B.M. Gupta. New Delhi: Segment Books, 1996. (Volume 13: Bibliometrics, scientometrics and informetrics)

HAWKINS, D. T. "Unconventional uses of on-line information retrieval systems: On-line bibliometric studies". *Journal of American Society of Information Science*, 28(1), 1977, 13-18

MACHLUP, Fritz, and Mansfield, Una (eds). *The study of information: Interdisciplinary messages*. New York: John Wiley, 1983. 4v.

SATIJA, M. P. "On quantifying knowledge" *Herald of Library Science*, 23:1-2, 1984, 23-30.

SMITH, Linda C. "Citation analysis" *Library Trends*, 30:1, Summer 1981, 83-106.

SWEENEY, Francis, (ed.). *The knowledge explosion: Liberation and limitation* (Series: The Centennial Colloquium at Boston College). New York: Farrar, Straus and Girous, 1966.

14.12 GLOSSARY

Bibliographic coupling: the relation of two documents by virtue of their joint descent from the third (or Two works are said to be coupled bibliographically if they have a common citation. Here it links source documents).

Bibliometrics: quantitative analysis of recorded information.

Citation: reference to a previous work (also called as reference)

Citation analysis: analysis of references in a document

Citation Index: A list of articles that, subsequent to the appearance of the original article, cite that article.

Co-citation: Two citations are cited together, to show the similarity of the subject matter. It is a reversal of bibliographic coupling, as Garfield says (Garfield: 1979: 99).

Core Journals: most important journals (highly ranked titles, based on the applications of Bradford's Law)

Econometrics: quantitative study of economic theories, and economic problems.

Frequency: The number of times a reference repeats itself.

Impact Factor: how often (on an average) each item is cited, decides its impact.

Informetrics: quantitative study of Information and trends in information transfer.

Librametry: quantitative study of library services.

Naukometrija: Russian term for scientometrics, *See* Scientometrics.

Ranking: measure of the relative position in a list.

Science Citation Index (SCI): Index of scientific publications, along with citations, keyword index, patent index, author index. It is published by Institute of Scientific Information, Philadelphia.

Scientometrics: quantitative study of scientific developments, and science of science.

Statistical Bibliography: statistical analysis of the literature or contents of a bibliography; this term is now replaced by bibliometrics.

Variable: a source for bibliometric study: such as book, author, publisher, journal, institution, etc.

14.13 ASSIGNMENT

1) Explain the need for bibliometric studies
2) What is the law of scatter, and how is the core list of titles derived according to this law.
3) Explain the key terms in bibliometrics.

14.14 MODEL EXAMINATION QUESTIONS

I ESSAY QUESTIONS

1) What are bibliometric laws. How are these different from one another?
2) What is citation analysis ? Explain the impact factor.
3) Trace the history of bibliometrics, from frequency counting techniques to mapping of natural sciences and social sciences.

II SHORT NOTES

a) Frequency counting
b) Quantitative methods
c) Science citation indexes

UNIT - 15: RESEARCH IN *LIS* - THE INTERNATIONAL SCENE

Structure

15.0 AIMS AND OBJECTIVES

In this unit you will be introduced to the concept of research in library and information science (LIS), its development, present scenario and future perspectives.

After reading this unit, you will be able to

- explain the need and purpose of LIS research

- list out the developments in LIS research
- enumerate the developments in LIS research in USA, UK and India
- describe the areas or aspects of LIS research
- reveal the financial support available for LIS research by various funding agencies.
- enumerate the sources of LIS research information
- explain the methods of LIS research and the problems involved
- discuss the future perspectives of LIS research.

15.1 INTRODUCTION

Research is an indispensable component for any discipline for the generation of new knowledge, which coupled with skill leads to effective action. Research is a means of continuously developing discipline with the capability to utilise the knowledge generated in other disciplines and it makes use of scientific method. However, defining research in professional context is difficult as it involves academic as well as professional activities. Hence, LIS research is known first by its areas of study, methods employed and the end it seeks. It is governed by the principle of objectivity and rejects all authority except that of valid evidence. Let us discuss these aspects in detail in the following sections of this unit.

15.2 LIS RESEARCH - NEED, PURPOSE AND FUNCTIONS

The acceptance of research programmes in the field of library developed gradually. Initially, many library leaders felt that advanced study had no place in librarianship, hence there was no organised research activity for a long time. However, the recognition for LIS as a profession and the contributions of professional organisations like American Library Association (USA), Library Association (UK) and International Federation of Documentation (FID) in initiating, encouraging and organising research work in library science during the last quarter of 19th and early years of 20th century have changed the situation. The constant efforts of eminent librarians around the world like Melvil Dewey, C.A. Cutter, Panizzi, J.D. Brown, B.W. Sayers, S.R. Ranganathan <u>et al</u> established librarianship as an important profession. The generation of truly scientific knowledge about LIS, theoretical principles and practice skills had made it possible for a scholarly recognition of LIS as a profession. To expand the scope of the profession, research has become inevitable.

With the proliferation of library schools and the acceptance of doctorate degree as legitimate course of study within the library field, research became firmly rooted in library education process. According to Shera, "a profession that would know itself, that would anticipate or to use Dennis Gabor's phrase `invent the future' must support and engage in productive research". Vickery has stated that research in LIS is legitimately concerned with the systematic investigation of problems regarding the nature of library and information work, by ascertaining principles and procedures. Though initially the LIS research used to arise out of individual's personal motivation, over the years particularly in the last two decades the pursuit of research has assumed aspects of team work. This change has broadened the base of LIS research and stimulated for creation of research teams/projects with financial assistance from various organisations. The tremendous growth of research activities may be attributed to increasing social pressures, financial support from government and other organisations and the realisation of librarians to raise the status of library science on par with other disciplines.

15.2.1 Need for Research in LIS

There is an ever increasing need for research activities in LIS. Let us study them here.

i) *Research as a Means of Development:* Research is the fundamental requrement for the growth and development of any discipline. LIS research is essential to extend the boundaries of its own body of knowledge and to improve the practices of the profession. It is a tool for progress and it broadens the outlook and scope of the subject.

ii) *As a Tool to Develop New Methods:* Research helps to develop tools for the organisation of knowledge and to get access to information. For example, many functional schemes of classification, cataloguing codes, indexing techniques, new techniques of library operation and its services etc have been developed through research.

iii) *To Sharpen the Existing Methods:* LIS research makes it possible for new applications of existing knowledge and research is the gateway to achieve it.

iv) *Scientific Inquiry:* To inquire into, or to make a scientific examination of the many underlying assumptions and adaptation of methods of information handling research is essential. It encourages an orderly process of investigation, analytical thinking and critical approach to the study.

v) *Advancement of Scholarship:* Working with hypothetical assumptions or opinions of the theoretical or practical concepts will surely leads to continuing advancement of scholarship of library professionals.

vi) *Increases the status of the Profession:* LIS research brings prestige to the individual undertaking research and also to investigate where it is being conducted and also to the profession as it can be compared on par with other disciplines. It enables the profession to achieve greater academic and professional respectability.

vii) *Interaction with other Subjects:* Research sets no boundaries to the subject of investigation. The process leads to the mutual enrichment of the other subjects, such as statistical applications, computer applications, etc. This enhances the scope of research work.

15.2.2 Purpose of LIS Research

The purpose of LIS research is clearly to support the functioning and development of libraries as social and educational institutions. It aims to find the best methods of acquiring, organising, storing, accessing, retrieving and disseminating information or data through research. Hence, it has become an important aspect of library professional and educational activity.

15.2.3 Functions of LIS Research

Krishan Kumar has listed out the following functions of research:

a) It produces knowledge

b) It enables discovering of new application of old knowledge

c) It helps in the development of the researcher intellectually and professionally. It is a continuous process.

d) It leads to better teaching as new knowledge is integrated into teaching programmes

e) It brings prestige to the person and the institution

f) It can prove to be a source of income. Funded research is a source of finance.

g) It enables an institution to attract better faculty, researchers and graduate students

h) It enables finding solutions to problems, to resolve conflicts in society

i) It enables man to relate move effectively to his environment and achieve his purposes.

j) It promotes progress of the society.

The functions of LIS research are also the same with emphasis on the acquisition of more accurate knowledge and understanding of libraries and librarianship. In simple, "its focus is upon the techniques whereby information is created, used and communicated as well as upon the discovery of quicker and more efficient methods to present the right recorded facts and figures to persons who need them, once the information has been stored". (Busha, Charles H., 1980)

15.3 HISTORY OF *LIS* RESEARCH

Though libraries existed since ancient times, the occupation of the librarian was considered as storekeeping. The librarian did not require any skills except those of calligraphy to be effective. However, the growth of library collections, their increased use and the social, political and educational changes altered dramatically the duties of a librarian and necessitated for the formulation of a body of techniques appropriate to the tasks to be performed. This has paved the way for education and training of library professionals in the 1880s. The period of formal education and training was continued without any change till the publication of Williamson Report in 1923, which has stated the librarianship as a discipline. This change has put its impact on the process of library functions. In spite of great controversies over considering the librarianship as a separate branch of science, it has overcome them and the society has recognised it as a profession.

Now, it is "recognised as a complex, multi-disciplinary subject ranging from computer and telecommunications through cybernetics to psychology, logic and techniques of classification and indexing and as such the information work and information science may be defined briefly as professional discipline concerned with the accumulation, storage and transfer of recorded knowledge". This professional character necessitated formal research.

Research in LIS existed since earlier times. History of library science reveals that the librarians in earlier times have undertaken a number of experiments to solve the problems encountered by them. Research was, then, focussed on solving immediate and practical problems.

Research in LIS has made a formal beginning after the publication of Williamson Report (USA) in 1923. The report pointed out, *inter alia*, the poverty of research in library science. The librarians were not interested in theoretical principles of their practices and the lack of curiosity among them distanced them from the academic world. The seeds of LIS research were first sown in USA in the 1930s. Williamson's article in *Library Quarterly* (1931) titled as "the place of research in library science" can be said as the starting point to direct the profession towards research. He has also chosen the topic "Research in Librarianship" for the Founders Day Celebrations in the School of Western Reserve University (USA). The contributions by Butler and Wilson, Cornovsky et al tried to emphasise research in library science by focussing on various aspects of research. Berelson in his paper read at Chicago Conference in 1948 initiated the library schools to conduct research. The research in LIS, thus, got its foundation.

15.3.1 Early Library Studies

Historically speaking, as has already been mentioned, LIS do not have a long tradition of research, scholarship or organised research activity in the modern sense. However, individual librarians in the past have experimented, though not following any scientific method. They have made a number of innovations in finding solutions to the library problems, new methods of organising knowledge, new techniques for information retrieval and new equipment required for libraries. The last quarter of 19th century saw the emergence of innovative and enthusiastic librarians in USA and UK. For example, Dewey's innovation of Decimal Classification, Cutter's contribution to various library practices, especially the rules for subject indexing; the development of cataloguing rules by Panizzi, Burgoyne's little book entitled, *Library Construction: architecture, fittings and furniture*; Berwick Sayer's contribution to classification. Here in our country S.R. Ranganathan's seminal contribution to almost all branches of LIS since the 1920s have influenced the library science throughout the world, irrespective of their country of association. Their experiments, along with those of other practioners have a prominent place in the history of LIS research.

15.3.2 Period of Organised Research

The organised research activity in the field has started since the 1930s in USA, 1960s in UK and late 1950s in India. These research activities can be broadly categorised into the following two types:

i) Research conducted in Library Schools leading to Ph.D. or M.Phil. degree; or Fellowship by research institutions or professional associations; and

ii) Research carried out by individual researchers or research teams in library schools, libraries, documentation centres, etc on certain problems without leading to the award of any formal degree.

With the proliferation of library schools and acceptance of doctorate as legitimate course of study within library field, increasing opportunities for higher positions, research became firmly rooted within the library schools.

The recent decades witnessed a tremendous growth in research activities supported by organisations and associations. The following sections will provide you the detailed information in this regard.

15.3.3 LIS Research Activities in USA

The USA may be said to be the first country to have started organised research programmes in LIS. It was at the initiative of Melvil Dewey that the degree course in Library Science at Columbia University was established in 1891. However, the advancement in this direction started with the establishment of the Graduate Library School at the University of Chicago in 1928 with a million dollar grant from the Carnegie Corporation. The school, from its inception, included provisions of Ph.D. programmes and the first Ph.D. was awarded in 1930. Between 1930 and 1959, 129 dissertations were produced in USA. Out of which 80 were awarded at the University of Chicago. During the 1950s, six other library schools, namely, Illinois, Michigan, Columbia, Western Reserve, California and Rutgers also started offering Ph.D. programmes. At present 27 universities offer doctoral programmes in the US alone.

The establishment of US Social Science Research Council Public Library Inquiry in 1950 was a land mark. It involved social scientists with a variety of backgrounds, who used different methods and techniques of research. In 1956, the Council on Library Resources at Washington was established with funds from Ford Foundation to support research and development in LIS. It was responsible for major projects in USA and abroad.

In the late 1960s, the research activities accelerated with the active participation of American Library Association (ALA). It established (i) a Committee on Research to look after the research needs and coordinating the research activities of the association; (ii) an Office of Research and Development to advance the theory, methods, principles of library science; and (iii) a Library Research Round Table to contribute towards the extension and improvement of library research.

Elaborate experiments have been carried out during the 1970s on information retrieval techniques especially at Case Western Reserve University, which were reported by Saracevic in 1971. The first large scale impact study on online search service was carried out by Wanger and her team in 1976.

The National Science Foundation provides extensive grants for basic and applied research in information science under the Division of Information Science and Technology. The areas included standards and measures, structure of information, behavioral aspects of information transfer and informetrics.

During the 1980s the LIS research in USA has attained the top position with surveys carried out with funds from organisations. The SMART system on experimental automated retrieval system; the MAPPER expert system by Borko (1987) at the University of California at Los Angeles for cataloguing of maps are a few examples.

Thus research in USA is carried out at several levels. These include research (i) by individual researchers, (ii) by library schools as part of doctoral programmes, (iii) by professional associations, (iv) by foundations, (v) by Federal and State Governments, and (vi) by research organisations. The country's research contribution to the world library and information science is highly significant.

15.3.4 LIS Research in UK

In the United Kingdom, Library Association used to award till 1946, Fellowship to Library Association (FLA) to those who produced satisfactory thesis on an approved subject. Later it has been made as part of final year course on LIS.

The Classification Research Group (CRG) was formed in 1952 and its members like D.J.Foskett, B.C.Vickery and J.E.L.Farradane whose contributions played a key role in influencing the research progress in library classification. Farradanes's (1961) work on relational operators in the context of analysis of relationship between different concepts is a landmark.

Aslib is also promoting research in LIS since the 1950s, when its research department was opened. Out of its several projects, the Aslib-Cranefield Project is outstanding. Two major experiments on evaluation and testing effort, popularised as Aslib-Cranefield studies, were carried out by Cyril Cleverdon. In Cranefield-I, started in 1958, post-coordinated indexing system was compared with more traditional indexing languages, using 18000 documents and 1200 search topics. The results of Cranefield-I led to formulation of hypotheses which were tested in Cranefield-II.

A breakthrough in LIS research in UK has occurred with the establishment of the Office of Scientific and Technical Information (OSTI) in 1965. The major functions of OSTI included research and development as an important area of activity. (Though OSTI is defunct now, its various research functions have been transferred to the British Library Research and Development Department, BLRD). In addition to OSTI, research projects are being carried out in libraries and information centres by individuals independently or under sponsorship of professional associations or institutions, such as LA, Aslib, BLRDD etc. The BLRDD supports wide range of research projects covering different fields of LIS activities. At present, it plays a pivotal role in the promotion of LIS research in UK. It has set up a number of research centres such as the Centre on User Studies, Centre for Library and Information Management, Primary Communications Research Centre, Information Technology Centre, etc.

Research studies in LIS are also carried out at Schools of Library and Information Science of several universities and polytechnics. Though education for librarianship in UK began with the establishment of the first school of librarianship at the University College of London in 19[illegible]9, the facilities for research programmes leading to Ph.D. degree started sometime late in the mid 1960s. The library schools are extending facilities both for Ph.D. work as well as for carrying out special research projects. The establishment of Postgraduate School of Librarianship and Information Science at the University of Sheffield in 1963 can rightly be considered as a turning point in the development of formal research programmes leading to Ph.D. During the 1970s two noteworthy research contributions were made at Bath University. (i) Maurice B Line's (1971) investigation to information requirements of the social scientists, and (ii) a controlled experiment on the comparative catalogue study reported during 1975.

The 1990s witnessed LIS research emphasis on computerized information retrieval systems. For example, Alice Vickery (1960) report about EUROPE, which will serve as a multipurpose intelligent (i) referral system, (ii) online system, and (iii) computer assisted instructor for business managers in Europe.

Thus, LIS research has a long history of development which has contributed to the present status of the profession. The revisions of classification schemes such as UDC, DDC, CC, etc and cataloguing codes like AACR have made the organisation of library and information resources more scientific and suitable to manual as well as computerised information retrieval processes. The innovations in the methods of LIS operations and the adoption of information technology have minimised the three Ms (material, manpower and man hours) and encourages the professionals to perform their duties with ease. Thus during the past 25-30 years much development in LIS research activity has taken place in USA and UK.

15.4 PRESENT STATUS OF LIS RESEARCH

The LIS research, after going through decades of ups and downs now has reached a state of relative maturity. It has been well recognised as a source of new knowledge, both theoretical and applied and the responsibility has been bestowed on LIS research to plan, budget, organisaional design, information retrieval programmes to suit the contemporary information society. P.B.Mangla has summed up the contributory factors to this advancement of LIS research as follows:

1) vast expansion of library collections, number of users and number of librarians and information scientists;

2) need to provide specialised services to the users, such as CAS, SDI, etc.

3) availability of large financial support from various agencies such as foundations, national and international organisations and other agencies, besides governments;

4) development of research activities in library schools, libraries and information centres;

5) rapid and innovative developments in information technology, microreproduction, communications technology, etc.

6) active role played by professional associations;

7) vast expansion in research programmes in science & technology and social sciences, which made it essential for library and information science schools to develope newer ways and means for information storage, retrieval and dissemination; and

8) the emergence of research journals for exchange of information.

Krishan Kumar has also emphasized similar reasons for the present status of research activity in LIS. According to him, the reasons are:

i) Library collections have grown steadily;

ii) Large increase in the number of librarians and information specialists;

iii) Increase in the scale of library and information activity;

iv) Greater availability of financial support for research in LIS;

v) Funding and growth of doctoral programmes in library schools;

vi) Rapid and innovative changes in technology have forced the profession to provide more emphasis to research in LIS, whereby LIS can keep pace with complexities of the modern world.

Hence, in order to meet the challenges of ever changing information environment, the LIS professionals are producing substantial research findings, carried out in universities or in libraries/information centres.

15.4.1 Institutional Support to LIS Research Programmes

As has already been discussed under Section 15.3, there are a number of institutions that supports research activities in LIS. They may be categorised as under:

i) International Organisations

LIS research is also carried out at international level by professional organisations such as IFLA, FID, Unesco and International Atomic Energy Agency. These are actively involved in conducting and supporting research through publications, seminars / conferences and the development of information systems.

IFLA: The Education and Research Division of IFLA strives for the promotion of LIS research. Further IFLA programmes, such as UBC, UAP, International MARC programme, Preservation and Conservation (PAC), Universal Dataflow and Telecommunications (UDT) and Advancement of Librarianship in the Third World (ALP) work for the advancement of librarianship. It communicates the current professional information through its publications, namely, IFLA Journal, IATUL Quarterly, IFLA Directory, World Guide to Libraries, Archives and Information Science Associations. It organises international seminars/conferences at different places and facilitates exchange of professional thoughts.

FID: The division of FID/CR involved, from its inception in the specialised research activity. The development and revisions of Universal Decimal Classification (UDC) is an example of its research programmes. FID/CR organises conferences regularly which forms a forum for the discussion of research work in classification. The *FID News Bulletin* is truly international in its character. The FID is also developed FID-ISORID programme, the International Information System On Research In Documentation which reports research projects, completed or ongoing throughout the world.

UNESCO: Unesco has been particularly prominent internationally in conducting statistical and other research activities and in supporting the projects of other international organisations. For example, Unesco collaborated with FID in the establishment of ISORID.

Its UNISIST project aims to support improved information transfer in the field of science and technology. The General Information Programme (PGI) running as part of UNISIST is striving for the advancement of information handling facilities with the application of information technologies. Eg. CDS/ISIS. PGI provides many activities in the field of information, documentation, archives and libraries for member states. It plays an important role in projects in many Third World countries. Its publication, the *PGI/UNISIST Newsletter* is a source of details about the programmes carried out by PGI.

ii) Professional Associations

In addition to the international organisations, various library associations have been supporting research projects in specific areas of LIS.

The American Library Association (ALA) and Special Library Association (SLA) are playing important roles in the promotion of LIS research. ALA, having 11 divisions, is supporting research programmes under school, public, college and research libraries, information technology, information services, etc. All the 11 divisions have their publications to communicate current developments. The American Society of Information through its periodical publication, namely, Journal of American Society for Information Science (JASIS) publishes current research programmes. The *Special Libraries* and *JASIS* are the two important, dependable primary journals for research information. The Council on Library Resources / The Coalition for Networked Information provides funds to libraries in the use of new technology to improve operating programmes and to improve services.

In UK, the Library Association has 23 Special Interest Groups (SIG) that support professional advancement. Its publication *LA Record* carries current research information. The ASLIB-Cranefield studies have already been discussed. Aslib has its own publication, the *Aslib Proceedings* for exchange of professional information.

In addition to professional associations the national libraries also encourage research activities. The Library of Congress (LC) has undertaken a number of research projects. For example, Cataloguing In Publication, MARC projects, etc. The British Library (BL) through its BLRD department is the major source of research activity in UK.

iii) Universities

Indeed, the progress of LIS research is closely associated with the doctoral programmes in universities. From the starting of the doctoral studies at University of Chicago in 1928, the research competence in university departments has increased throughout the world. The very existence of library and information studies in university context implies a belief that the course has been recognised as an academic discipline. As a consequence of this academic acceptance, there is a steady growth of the number of people studying for higher degrees leading to Ph.D. In USA about 27 universities offer Ph.D. programmes and in UK 13 while in India about 38 universities offer the same. That is how the number of doctoral programmes in basic, applied and action research has been ever increasing. According to ALISE, the purpose of these programmes is to "prepare a person to advance knowledge in the field of LIS. The degree is a research degree and as such has three broad purposes. They are - to heighten powers of critical thinking; to develop skill in addressing research problems; and to generate through dissertation work an advancement in research or theory.

15.4.2 Areas of LIS Research

Research in the past focussed its attention mostly on the historical aspects of the profession since the scope of LIS at that time was limited and it was in its emerging stage. It was only from the 1950s, that the library research paid increased attention into the problems and functions of librarianship when the craft of the profession began to crystalise. The shift came because of the impact of changing information society, where in (i) information has been recognised as an important national resource for all walks of human activity; (ii) production and availability of large quantities of information has to be disseminated by one way or the other.

During the 1930s and 1950s, the research programmes were essentially descriptive and historical studies, which have included the areas such as history of books, printing and publishing, history of individual libraries, their governance and administration; studies of technical service; studies of reading interests, bibliographic compilations, etc.

Later the situation has considerably changed. The areas of research included management of information sources, personnel administration, technical services, use pattern of different types of communication media, user studies to seek their opinion on existing facilities and to examine their actual requirements, classification and cataloguing with emphasis on thesaurus and subject indexing, etc.

However, the use of computer and concern for the user oriented services and the rapid developments in LIS curriculum have thrown open many new areas. The traditional areas of classification, thesaurus, content analysis and consolidation, cataloguing, indexing, abstracting, current awareness, etc still exist but these are superseded to some extent by new areas such as resource sharing, networking, characteristics of users, end-user surveys, etc. which have opened new vistas. Innovative ideas for work norms, standards, staff pattern, etc are being worked out. The present thrust area is application of information technologies for various library operations and information retrieval mechanisms which save labour, improve speed and accuracy, maximise convenience and dependability and reduce costs.

In general the areas of LIS research can be broadly listed out as follows:

i) LIS Activities, such as buildings, planning and management; preservation, professional, application of IT.

ii) Information Storage and Retrieval, including classification cataloguing and other methods of organising knowledge.

iii) Dissemination of Information - information services, systems, user studies, cooperation, etc.

iv) Bibliometric Studies

v) Education for LIS

vi) Historical Studies on Library Associations, Library legislation, Library history and Library Movement.

vii) Impact of IT, Comparative studies on Online information retrieval, CD-ROM databases, etc.

The current trend is towards qualitative research, interdisciplinary research, technology-based research and information use studies. Some of the thrust areas identified for current and future research include

i) Testing the comparative efficiency of the retrieval system

ii) Computational linguistics, e.g., automatic indexing, automatic abstracting and machine translation, etc.

iii) Communication technology;

iv) Information Systems and Computer-based specialised services

v) Database design and management services

vi) Cost-benefit and cost-effective studies of operations, services, etc.

vii) Use of IT for various purposes in libraries, information systems and services

viii) User Studies - end-user surveys.

Current Trends in LIS Research

There is a significant change in the selection and approach to research topics over the years. The topics of the 1970s reflect the conventional case studies of libraries, use studies of information sources and user studies, LIS education, profession, etc. The 1980s developed studies on computer applications and bibliometric analysis in addition to the traditional areas. The 1990s have witnessed the predominant emphasis of LIS research on information technology and its application to information handling. This progress is more evident in developed countries like USA and UK. Although, the basic principles and functions of library and information science profession remained the same, the approach to achieve them has changed dramatically with the advent of IT. This has greatly influenced the aspects of LIS research. A peep into the research areas reported in *Research in Librarianship* (England), *Dissertation Abstracts International - (A) Social Sciences* (Michigan, UMI) and *Current Research in LIbrary and Information Science* shows the trends in LIS research. The topics studied for research during the 1970s include:

1) Professionalisation - structure, limits, sociological concepts, etc.
2) Public library system
3) Academic library system
4) Special library system
5) Reading habits/interest
6) Information analysis and consolidation
7) Citation analysis of specific literature
8) Use patterns of primary and secondary sources
9) Library use statistics
10) User studies or attitude studies.

The 1980s research areas encompass studies with more emphasis on users and computer applications besides the traditional studies of provision and use of different library systems. Use, user studies and bibliometric studies were also given much prominence. The areas may be categorised as

1) User studies: Perception of users, empirical studies, instructional support, behavioural analysis, etc.

2) Indexing system: Controlled vocabulary, subject and cross reference indexing and automated classification systems for information storage and retrieval purpose.

3) Cooperation and Resource sharing: Depository library system; Design for cooperation, etc.

4) Application of technological investigations -Software development, computer approach to managerial tasks; Hypermedia workstation; Development of databases; Search theory of information; Visual literacy on computer culture; Evaluation of optical storage media; End-user computing, etc.

A review of current issues of *Dissertation Abstracts International* presents the latest trends in LIS research. The topics included are -

1) Marketing analysis and strategies
2) Public access to microcomputers and the software
3) Computer-based large group learning system
4) Library CD-ROM LAN performance in patron use
5) Library networks development
6) Multi-disciplinary implication; Citation ranking and Bibliometric studies
7) Impact of IT and assessment of different library system

Indeed the CRLIS has listed the research topics under the following broad areas, which reveals the influence of IT on LIS research areas.

1) Profession
2) Education and training - Implementation of Computer- Assisted Learning (CAL)
3) Public library system
4) School library system
5) University/college library system
6) Special library system
7) Library use and users
8) Computer applications
9) Electronic media/ Multimedia - Standards, Evaluation, Global inventory projects, Online catalogues, Cooperative archives, serial management, etc.
10) Library management
11) Networks - Internet, Visual library museum pages, Community tele service centres, evaluation, WWW indexing tools, etc.
12) Preservation - digital archiving and long-term access to digital data
13) Security of digital information and communication
14) Information work - Information kiosks, information strategies and export market, online catalogues, cooperative activities, Virtual Reality Technology Library.
15) Searching - Databases
16) Software - Development and use studies
17) Non-bibliographical information: fulltext databases and databanks;
18) Imaging technology
19) Electronic publishing, etc.

From the above description, you might have taken note of the changing trends in LIS research areas shifting from traditional areas to technological applications. However, a further analysis reveals that the thrust of research on application of IT is from developed countries and the developing countries are still working out on conventional topics with little variation. This is because in developing countries the technological application in libraries is in its infancy whereas in developed countries it has attained a state of relative maturity.

Thus, computerisation is the buzz word in the present day information environment. Increasing research activities within library and information science began to interface with other subject areas, such as audio-visual technology, microform technology, computer science, communication and instructional technologies. These developments enhanced the research potentialities of LIS and gave scope to develop new theories, principles and practices.

15.4.3 Methods of LIS Research

The term 'Research Methods' may be understood as all those techniques that are used for conducting a systematic study. It refers to the methods an investigator uses to collect, analyse, assess and present findings on a problem. Methods in LIS research are borrowed from other disciplines. The social science research methods are often applied to LIS research. The research methods used in social sciences are mostly based on scientific methods and they provide a framework within which hypothesis, testing and inferential thought may be carried out in a disciplined manner. The scientific method consists of the following steps:

i) Identifying a problem

ii) Formulating a tentative hypothesis

iii) Collecting data or truths basing the problem

iv) Classifying the data

v) Analysing the data

vi) Deducing inferences and verifying the hypothesis.

The research in LIS is conducted following the above steps systematically or using the framework of a scientific method. The historical, survey or case study methods most often used in LIS research comply with the framework of scientific method.

The initial research in LIS followed historical and descriptive method. Later, case study method is adopted for a diagnostic study of existing library facilities. The current trend is for survey, experimental and operations research methods. However, while selecting a particular method, it is imperative to make sure that the adopted method must be based on scientific approach.

15.4.4 Funding of LIS Research

Traditionally, the funding of LIS research is primarily by institutions, learned societies, and professional associations.

In USA, the funds come from Ford Foundation, the Carnegie Corporation, the ALA, etc. The Federal and State governments also support the advancement of research. According to a recent compilation, the Carnegie Corporation financed 26 library science research projects at a cost of \$350,653. The Ford Foundation awarded five grants of \$5,200,000 in the field of librarianship. Between 1956-61, the Cooperative Research Branch awarded four library science related projects. The Educational Media Branch awarded seven grants for library science research costing over \$158,000. Further, the Library Services Act also assisted in the projects on rural public services. The National Science Foundation (NSF) provides extensive grants for basic and applied research in information science under the Division of Information Science and Technology. The USA also offers fulbright fellowship to sponsor research in the Third World countries.

In UK, the funding is by several organisations such as Government Departments, Research Councils, Learned Societies and Foundations. OSTI, since 1965, coordinated research activities with funds of millions of dollars. For example, Funding during the period 1973-74 was \$80,000. The Library Research Committee of LA was provided with grants running to thousands of pounds from the LA besides several adhoc grants from other organisations. The BLRDD is a primary funding agency of LIS research in Great Britain. It supports research with huge financial grants. The doctoral programmes in the universities are funded by the individual universities through fellowships, research assistantship, etc. Traditionally, the main purpose of financial support for research has been to advance the subject base, but now-a-days more emphasis is on improvement of practices such as design of new systems, application of new technologies, provision for wider range and quality of services, etc.

15.4.5 Problems of LIS Research

LIS research reached a state of relative maturity in exploring various facts and finding solutions is beset with certain limitations and often faces criticism. The problems and criticism on LIS research are -

i) The quality and quantity of research has risen dramatically over the decades. However, the topics chosen are not inspiring and relevant.

ii) The research output is episodic rather than sustained and cumulative. This has blocked the progress of the profession.

iii) The research techniques applied for the purpose are far from satisfactory. For example, the surveys are not considered to be a foolproof method. The limitations of questionnaire method create problems as the data gathered does not achieve a higher degree of specification and it is difficult to draw generations.

iv) Statistics, descriptive or inferential are the language of scientific research. It is this component which has been singularly lacking in library science. Very little has been written on how to apply statistical methods to analyse data.

v) A major impediment to LIS research is paucity of funds. Even in USA, very few LIS research projects are being funded compared to other disciplines. The plight of developing countries is still more discouraging.

vi) There is an apathy on the part of the profession towards research. The profession has certainly not provided enough recognition to research achievements.

vii) Communication of research results is not satisfactory. Hence, the findings are not being implemented.

viii) There is lack of research tradition. In many countries research in LIS is a recent phenomenon.

15.4.6 LIS Research Publications

The major categories of information sources which reveal the research information can be categorised as primary, secondary and tertiary.

i) Primary Sources

A majority of professional journals in the field of library and information science and the newsletter/bulletins from the professional associations report about the research activity. Some of the important journals which carries research articles are

Library Quarterly (1930 -)
Journal of Librarianship
Journal of American Society for Information Science (JASIS)
RQ (Reference Quarterly)
Radials Bulletin
BLRD & D Reports
Newsletters of Professional Associations
Library and Information Research News (LIRG of LA)
Library and Information Science Research
JELIS

These primary sources are the vehicles for publication of important articles on theory and practice of librarianship, as well as accounts on a wide range of research activity.

ii) Secondary Sources

The secondary sources serve as principal tools for knowing the primary sources of information in a subject area. Some of the important secondary sources of information in the field library and information science are

Library Literature
Library and Information Science Abstracts (LISA)
Current Research in Library and Information Science (CRLIS)
Information Science Abstracts
Current Research in Information Profession

iii) Tertiary Sources

These sources provide preliminary background information. These can set the researcher to start his work. They may provide specific points of terminology, programmes or factual data. Encyclopedias, glossaries, bibliography of bibliographies, statistical sources, directories, etc are few examples .

Encyclopedias:

World Information Technology Manual (1991)
Encyclopedia of Library and Information Science
ALA World Encyclopedia of Library and Information Science

Glossaries:

Harrod's Librarians Glossary and Reference Book (1990)

Besides, a few monographs like *British Librarianship and Information* (LA), *Annual Review of Information Science and Technology, The UMI List of Doctoral Publications*, etc are of much use to know about the research information in LIS.

15.5 FUTURE OF *LIS* RESEARCH

The combination of technology, economics and the rising flood of information forces the profession to undertake new research programmes to acquire the skills and knowledge to navigate the computer networks. The technological developments and other innovations have made due impact on the organisation and functioning of libraries and thus are promoting LIS research. Hence, the thrust areas of future research would be application of new technologies to library operations and information retrieval. The libraries of tomorrow will become more like information service centres. As such they would require more powerful tools for storage and retrieval of information.

According to Shobas, "we are not about to seek the skills of indexers, abstracters, classifiers who became obsolete because of computers. On the contrary, there is a crying need for more of these skills in computer-based information retrieval. Too many information retrieval systems today are flawed, some seriously because of the absence of such information experts during the design stage". Further, the changes occurred in communication and storage media in the form of electronic publications and compact discs necessitates exploration in these areas. You might have read or heard about Internet. It has become an essential part of our library work. It helps to search web sites and to locate information, down-load files or transfer data wherever we want. The networks which provide global access to information, would be a potential area for further research.

Thus, the future of LIS research would like to concentrate more on application of IT. Already the new technologies like Expert Systems, Artificial Intelligence, Hypertext and Hypermedia paved their way into LIS profession which demands for further research.

15.6 LET US SUM UP

LIS research has a long history of development. Though its development in initial stages was not on scientific basis, the research programmes from the 1950s have contributed to the development of the profession. The advancement of LIS research has been made possible with the financial and moral support of foundations, professional associations, national libraries and universities, etc. There is a shift in the areas of research work from traditional topics to current demands of information society. Nevertheless, the thrust of current LIS research has been influenced by information technologies and the future of LIS research has to concentrate to explore the potential IT applications to library and information activities.

15.7 REFERENCES AND RECOMMENDED BOOKS

ALISC. *The Ph.D. in library and information science: A position paper of the Association for Library and Information Science Education*. October 14th, 1990.

Busha, Charles H. *Research methods in librarianship*. New York: Academic Press, 1980.

Krishan Kumar. *Research methods in library and information science*. Delhi: Har-Anand; Vikas, 1992. p.9

Mangla, P.B. "Research in library and information science in USA, UK and India: An overview". IN *Perspectives in library and information science* (C.G. Viswanathan Festschrift). Lucknow: Print House, 1982.

Shera, J.H. *The foundations of education for librarianship*. New York: Becker and Hayes, 1972.

Vickery, B.C. "Academic research in library and information studies". *Journal of Librarianship*. 7(3) July 1975; p.153-159.

15.8 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) Discuss the need and functions of LIS research.
2) Describe the development of LIS research.
3) Discuss various sources of information on LIS research.
4) Trace out various funding bodies for financing the research projects in USA and UK.
5) Discuss the role of international organisations in promoting LIS research.

II. SHORT NOTES

a) Organised Research
b) Research project
c) Professional Associations and LIS Research
d) Future of LIS Research

UNIT-16: RESEARCH IN *LIS* - IN INDIA

Structure

16.0 AIMS AND OBJECTIVES

The aim of this unit is to give an overview of the developments in LIS research in India.

After studying this unit, you should be able to

- get an idea about the professional concept of LIS
- explain the genesis of LIS research in India
- describe the contributions of S.R.Ranganathan to library research
- enumerate the thrust areas of LIS research in India
- discuss the financial support available to LIS research in the country
- trace the future perspectives of LIS research in India.

16.1 INTRODUCTION

Library and information science (LIS) like any other applied discipline, is evergrowing, which ensures gradual making of librarianship as a profession. However, the growth of LIS is possible only through research process and discovering answers to questions through application of scientific procedure. The LIS research programmes provide new ideas and solutions to meet the information needs of society in a better way at the national level. The research activity in LIS has increased remarkably in India during the last two decades, though there is an evidence for its existence even earlier times. The following sections trace out the development of LIS research in India, describes the areas of research and its future perspectives etc.

16.2 *LIS* RESEARCH IN INDIA

The concept of LIS research in its present day meaning is in fact a phenomenon of the post-Independence period. The credit for initiating and stimulating scientific investigation in librarianship goes to the doyen of Indian librarianship, Dr S.R. Ranganathan. Though a developing country, India's contribution to LIS during 1950 and 1970 is particularly pioneering and significant voluminous writings of S.R.Ranganathan in almost all aspects of librarianship contributed original ideas which had an impact on LIS research. These contributions are well recognised all over the world. These contributions also helped the profession to achieve a status in India.

16.2.1 Recognition of LIS as a Profession in India

A profession comes into being when a group of individuals recognising the existence of a genuine social need, join together to find means of meeting that need through the elaboration certain techniques. Librarianship in India has exhibited an evolution to rise to the status of a profession and now it parallels with the professions having comparable characteristic such as a body of relatively abstract knowledge, susceptible of practical application and service-orientedness which is determined by dominant socio, political, economic and ideological environment. It has been made possible because of the sole contribution by S.R. Ranganathan who with his scholarly, authentic and voluminous writings comprising of 60 books and above 1200 papers put India on the world map of professional librarianship.

16.2.2 Need for LIS Research in India

P.B. Mangla has identified the following factors which necessitated for vigorous research activity in the country.

1) Increasing emphasis on research qualifications of different senior level positions in libraries, information centres and documentation centres;

2) Gradual recognition of the role of libraries and information in programmes of socio-economic development, industrial growth, R&D activity, etc.;

3) Increasing need to develop and provide computer-based information sources to meet the varied and complex information requirements of the users;

4) Need to develop well organised information systems at the national, state and regional levels;

5) Need to develop and utilise information technology; and

6) Need to develop inter-disciplinary research with areas such as linguistics, mathematics, sociology, psychology, computer science, etc.

16.3 HISTORY OF LIS RESEARCH IN INDIA

The history of LIS field does not have a long tradition as that of libraries in India. However, the development of research is associated with the development of libraries. Let us have a brief overview of the development of libraries and library education in the country which has a direct bearing on the developments in LIS education and research.

16.3.1 Early Libraries

The archaeological evidence and scriptural description of Nalanda and Taxila universities indicate that the libraries at these two universities as the ancient Indian libraries. However, most of the Indian libraries belong to the 19th and 20th centuries, with smaller number of books and learning was limited to the privileged few. Indian library system has grown well along with the national movement for independence. The spread of public libraries and their organisations was the topic of interest. Though attempts were made, the public libraries have not assumed their role in the rural sector as there was no legislative support.

Among the academic libraries, only university libraries were prominent. The Wheat Loan Scheme of USA under PL-480 have given a facelift to the profession and the professional community got its recognition among the academic community.

Development of special libraries was a post Independence phenomenon. They have grown with research in science and technology, industry and institutional research in different subject fields.

16.3.2 LIS Education and Research

The LIS education has its roots in 1910, with Borden and Dickinson organised LIS schools at Baroda and Lahore. The facility for library science education in the country has increased to a great extent during the past 40 years. Now, more than 80 departments of library science in the universities are functioning in the country offering this course at graduate, postgraduate and research levels.

However, the concept of research was never considered during the early period, except for few solutions to practical problems in libraries. The era of LIS research started with S.R. Ranganathan. During the early period, i.e., upto 1970s if any worthwhile research has been done in the field of library science and documentation has been largely due to the contributions of a single individual and a few of his associates.

16.4 CONTRIBUTION OF S.R. RANGANATHAN FOR LIS RESEARCH IN INDIA

Comprehensive data is not available about the various research projects carried out during early development of LIS research. However, LIS research on a scientific basis had its beginning with the entry of S.R. Ranganathan in the later part of 1920s. Prior to that, librarianship had been considered as a craft involving clerical operations and house keeping for a collection of books and other reading materials. His works have dominated the research activities for five decades and he is responsible for developing some of the fundamental and theoretical principles. The development of basic laws, normative principles, cannons for classification and cataloguing, subject indexing and chain procedure, theoretical and practical base of book selection, reference service, documentation, library administration and management, professional education and development are few examples.

Ranganathan was the driving force behind the FID/CR. He acted as chairman of FID/CR from 1951 to 1960. His ideas on classification and cataloguing became the core material and areas of research in different library schools all over the world. Ranganathan thought that public libraries could be developed only through state legislation and saw that the four Southern state enacted library legislation. His concept of reference service and facets of documentation work, was developed with special terminology combined with exposition. He was noted for his research among the librarians and the academic community of those days respected him as a pioneer researcher. The contributions of Ranganathan can be listed out briefly as follows:

1) Normative Principles

The Five Laws of Library Science were enunciated first in 1928 and published in 1931. The Laws have made a great impact on the development of library thought in the world. These are used as guiding principles for expounding different aspects of library science. Through these Laws the prevailing practices in libraries could be explained and evaluated.

2) Classification

When the Colon Classification was published in 1933, it was the first fulfledged faceted classification. It made a marked departure from the then known schemes for classification of subjects. The theory forming the basis of classification was embodied in the *Prolegomena to Library Classification* (first published in 1937). It was a profound work that has influenced classificatory thought throughout the world. By 1956, considerable research has been made to develop depth schedules of specific subjects. In 1963, a breakthrough fundamental research yielded a helpful methodology on the basis of stated postulates and principles. This has opened up avenues for large scale, continued, developmental team research in the subject.

3) Cataloguing

In 1934 Ranganathan designed and published the first complete code for the preparation of *Classified Catalogue Code*. In 1938 the chain procedure was developed. This was applied on large scale in British National Bibliography till Precis was invented in 1971. The canons of cataloguing, the Classified Catalogue Code (Ed.5 published in 1964) and the chain procedure technique have influenced considerably the whole world in the field of cataloguing.

4) Document Usage and Reference Service

Ranganathan has developed a theory for the provision of assistance to readers in the 1930s which as published and popularised in 1940 as the Reference Service. This is a typical example of applied research, where psychology and classification meet. It describes the psychology of readers, reactions of reference librarians to different kinds of questions, methods of reference service, etc.

5) Documentation

Ranganathan has made clear statement on the objectives and functions of documentation. This has helped in organisation for documentation and drawing up a suitable curriculum for the training of documentalists. The book entitled *Documentation and its facets* was published in 1963 and it was regarded as a unique contribution of its kind.

6) Library Management

The *Library Administration* (1959) by Ranganathan is a remarkable piece of applied research. It analyses all types of routines in library. The work analysis and job analysis of each library routine in Madras University library was based on this theory. The staff formula, the three cards system and other procedures were familiarised in India and abroad.

7) Library Legislation

Ranganathan has worked out a number of organisational plans on the basis of Five Laws. The first Model Library Act was framed in 1930, which was revised later to suit the independent status of India. Draft bills have been drawn up for several of the constituent states and union territories.

8) Library Buildings, Fittings and Furniture

Considering the economy, standards and Five Laws, the primary elements in the design of library buildings, furniture and fittings have been established. The Indian Standards Institute has published these standards as two documents.

9) Standardisation

Several standard in the field of documentation, book production, technical terminology, etc were formulated by the Sectoral Committee on Documentation of Indian Standards Institution have become the basis for international standards.

10) Professional Education

Ranganathan was a pioneer in the establishment of full-time graduate courses in LIS in the universities. He also started postgraduate courses and doctoral degrees at Delhi for the first time in the country.

11) Contribution to LIS Research

Ranganathan has done himself individual research for several years. However, he encouraged team research also. He initiated team research and inspired and guided many of his students to participate in the meetings of the 'Library Research Circle' founded by him at Delhi. Further, he has sown seeds for fulfledged level of LIS research when he set up the Documentation Research and Training Centre (DRTC) at Bangalore in 1962. The research activities at DRTC aimed at developing some of the fundamental theories. The important areas of the research work were classification and related fields.

16.5 PRESENT STATUS OF *LIS* RESEARCH IN INDIA

The LIS research in India has remained as 'solo research' till the late 1940s. Only in the late 1940s and 1950s, research in librarianship had started moving on scientific lines. The formal university research programme started in 1948 at the University of Delhi. It was at a very slow pace till 1980s. In the recent times there is a spurt in the research activity. The reasons for this situation may be attributed to

(1) increase in the number of schools offering LIS education;

(2) recognition and prescription of doctoral qualifications for appointment, promotion and career development of faculty and college/university librarians by the University Grants Commission (UGC)

(3) Emergence of funding bodies like UGC, ICSSR, CSIR, etc.

The present research activities in India can be broadly categoried into the following two types:

i) Research conducted in LIS departments in universities leading to the award of M.Phil. and Ph.D. degrees; and

ii) Research carried out by individual researchers on research teams in library schools, documentation training centres, etc on certain problems without leading to the award of any formal degree.

The LIS research in the country is mostly taking place in universities and in the training centres, viz., INSDOC and DRTC. In a way institutional research is very minimal.

16.5.1 Research Programmes

The research programmes in the country are being carried out by different organisations.

(1) Universities

In universities, the research programmes are at two levels - M.Phil. and Ph.D.

M.Phil. Programme:

With the recommendations of the Sen Committee of UGC, the M.Phil. programmes in various disciplines were initiated in different universities. University of Delhi was the first university to introduce M.Phil. programme in LIS in 1980. Today, many universities are offering this course leading to the award of the M.Phil. degrees.

The major objectives of M.Phil. programs are

i) to make the students proficient in the methods and techniques of research and their application to the problems of LIS;

ii) to give the students specialised knowledge in selected areas of LIS;

iii) to train the students for doing further research work leading to Ph.D. degree; and

iv) to expose the candidates to teaching methods relevant to LIS.

Ph.D. Programme:

Research studies leading to the award of doctoral degree offered by the university departments. Research at Ph.D. level was initiated by Dr S R Ranganathan in 1948 at the University of Delhi with the support of the then Vice-Chancellor, Sir Maurice Gwyer and the first doctorate degree was awarded in 1957 to D.B. Krishna Rao. It was only after 20 years that the second Ph.D. degree was awarded by the Punjab University and the recipient was J.S. Sharma. Today 38 universities in the country provide facilities for doctoral research.

The major objectives of Ph.D. are

i) to make a significant contribution to existing knowledge;

ii) to give evidence of carrying out independent research and ability to advance knowledge in the field of LIS;

iii) to be able to attain mastery of a specialised area of LIS;

iv) to promote cross fertilization of ideas through interdisciplinary studies; and

v) to provide a selected number of candidates for teaching positions in LIS

Dev Raj Sharma in his document on 'Doctoral Research in Library and Information Science' (1997) has made an attempt to compile the bibliographical details about the theses/ dissertations submitted to the Indian universities on which Ph.D./D.Phil/ D.Litt. degrees have been awarded from 1950 to March 1997. Approximately 350 theses have been awarded during the period under various Indian universities. As has already been mentioned, the LIS research was at very slow pace upto 1970s. The sudden upsurge is visible from 1980s onwards. This is evident from the year-wise break up of the number of theses as cited by Dev Raj Sharma.

During the 1950s and 1960s only two theses each have been contributed while in 1970s nine PhDs have been awarded. The following table reveals the year-wise break up of PhDs awarded during the 1980s and 1990s:

S.No.	Year	No. of PhDs
1)	1980	04
2)	1981	04
3)	1982	03
4)	1983	11
5)	1984	13
6)	1985	20
7)	1986	10
8)	1987	09
9)	1988	09
10)	1989	15

S.No.	Year	No. of PhDs
11)	1990	16
12)	1991	26
13)	1992	50
14)	1993	27
15)	1994	46
16)	1995	34
17)	1996	40
18)	1997 (March)	08
19)		
20)		

Source: Dev Raj Sharma. *Doctoral Research in Library and Information Science in Indian Universities*. Palampur: [Author], 1997.

Further, some of the universities offer post-doctoral programmes with an intention to further advance the research studies. Dr Sen Gupta was awarded D.Sc. for his studies on Bibliometrics.

(2) Institutional Research

In addition to universities, research work is also carried at by training centres like Documentation Research and Training Centre (DRTC) and Indian National Scientific Documentation Centre (INSDOC). The DRTC was set up in 1962 by S.R. Ranganathan. Since its inception, it has been carrying out research studies on documentation and related areas. The results of the research are communicated through its official organ, *Journal of Library Science with a Slant to Documentation and Information Studies* and annual seminars. It is the only institute which is engaged in individual and team research so far as the aspects of LIS are concerned. The revision of Colon Classification (Seventh edition) is its major contribution in recent times.

INSDOC, especially with the advent of information technologies, has undertaken a number of programmes which contribute to the advancement of knowledge in library and information science. It communicates LIS knowledge through its official organ, namely, *Annals of Library Science and Documentation*, besides other publications.

The Defence Science Information and Documentation Centre (DESIDOC) has also taken an active part in the development of library application software packages based on CDS/ISIS. Its publication, *DESIDOC Bulletin*, endeavors to bring the recent developments in the field of Information Technology, with a slant to LIS.

(3) Library Associations

Compared with ALA and SLA of USA and LA and Aslib of UK, the contribution of library associations in India from ILA and IASLIC in research activities is negligible. However, they are conducting annual seminars regularly to encourage the professionals to share their ideas, innovations and experiences. The official organs of these associations, namely ILA Bulletin, IASLIC Bulletin are communicating their innovations in the field. There is a need for our professional associations to take greater interest in research activities like those of western countries.

16.5.2 Areas of LIS Research

Sewa Singh has identified the following topics of current research in which either research has already been done or is in progress at present.

(1) Scientific Management of Libraries

A number of works have already been done in this field, starting with S.R.Ranganathan. However, very few studies have been undertaken on the application of IT in library operations.

(2) Classification and Indexing

Colon classification, Classified Catalogue Code and POPSI are the remarkable contributions in this area. Other researchers have also made studies in the related areas like relational indexing.

(3) Thesaurus Construction

The DRTC and individual researchers have made attempts to construct thesaurus in various minor subject fields. This concept requires further search.

(4) Bibliometrics

A number of bibliometric studies have been conducted to do the literature analysis in various subject fields. This mainly consisted of ranking of journals, citation analysis, country-wide contributions, etc are the main focus areas.

(5) User Studies

The present trend is to conduct user studies to assess the opinion of and to seek suggestion from the user on the existing library facilities and/or impact of IT etc.

(6) Computer Applications

A few, but valid studies have been carried out in this area and software development for various library operations is the most favourable aspect under this field.

(7) Reference and Information Services

A majority of the user surveys have tried to find out the actual need for information services. Assessment of existing services and comparison of manual and automated services also of common interest for research.

(8) Library Systems

Evaluation of public libraries, community information, academic libraries and different types of special library systems have been carried out using case study techniques. Some consider that these studies have become abundant and redundant also.

(9) Library Legislation and Movement

Quite a few studies have given emphasis on public library movement and legislation in different states of the country, following the historical method of research.

(10) LIS Education

LIS education and research with emphasis on curriculum manpower requirements, teaching faculty, teaching methods and material, accreditation of LIS schools, etc are the major topics of interest. These studies have brought forth facts on these areas and made several suggestions for improvement.

Interdisciplinary studies have also been conducted by LIS researchers. Eg., Children's literature and libraries; Management of computerised libraries; Artificial intelligence and library functions, etc.

P.S.G. Kumar has identified the following as the areas of study for LIS research in India:

Automation, Bibliometric studies, Cataloguing, Classification, Communication, Cooperation, Documentation, Information Systems, Legislation, Library Movement, Management Studies, Manuscripts, Reference Service, Technology Transfer, Universe of Subjects, Use and User Studies, etc.

A recent study by A.A.N. Raju helps us to identify the trends in LIS research in India during the 1990s. His study reveals that the topics of doctoral research from 1993 to 1997 cover the broad topics such as -

User studies

University libraries

Bibliometrics

Information storage and retrieval

Management

Public libraries

Information centres/services

Information systems and networks.

Among the above list of topics user studies occupies the prime position. A further analysis indicates that user studies includes attitudes, information seeking behaviour of users, etc.

Even among the topics registered for Ph.D. degrees also, 'User studies' seems to be a **hot** favourite for research scholars followed by Bibliometrics/Citation analysis, information centres/ services/networks and university libraries.

A.A.N. Raju has also observed that much of the LIS research is undertaken in traditional areas of LIS, the reason may be the traditional areas are easy for pursuance. The other reasons, according to him, include -

a) non-availability of expertise in the LIS department to guide the scholars who are coming forward to select topics on the various aspects of computer applications in LICs and Information Technology;

b) lack of confidence and enthusiasm on the part of the research scholars to select topics under emerging information and communication technologies and computer applications;

c) lack of infrastructure, expertise, resources and computerised LICs, which can be used as laboratories for pursuing research.

Suggested Areas of Research

P.B. Mangla has listed the following areas for further research work in LIS in India:

(1) User Studies: Information requirements of the users in different subject areas, literature use, library use pattern, user training, etc.

(2) Subject Indexing: Terminology, thesauri construction and maintenance, etc.

(3) Clarification and Cataloguing: Special classification schemes, cataloguing of manuscripts, books in different languages, non-book materials, etc.

(4) Information Systems and Services: Developing the information systems, computer-based information services, etc.

(5) Management of Libraries and Information Systems:

(6) Automation in Libraries and Design of Information Systems

(7) Library and Information Science Education and Research

(8) Interdisciplinary Studies in Library and Information Systems and Services

(9) Library Buildings and Equipment

16.5.3 Funding of LIS Research in India

The financial support for LIS research programmes in India is mainly through grants from the University Grants Commission and institutions of higher learning and research.

(1) **University Grants Commission** (UGC)

The UGC has a scheme of Junior Research Fellowships, which are awarded on the basis of national level test. The test is being conducted twice a year. The fellowship offers an amount of Rs.2300/- (as JRF) and Rs.2800/- per month (as SRF after completing JRF for two years) and in addition an amount of Rs.5000/- as contingency grant per annum. These facilities are for conducting full-time and regular research. In addition, the UGC is also funding the publication of doctoral theses. It also supports minor and major research projects for post-doctoral research.

(2) **Indian Council of Social Science Research** (ICSSR)

The ICSSR considers financial assistance for any research proposal on any theme that has a 'significant social sciences orientation'.

It offers three types of grants:

i) Awards to teachers, with a maximum value of Rs.7,500/- to pursue their research interest, to promote marginal support to scholars and to write a book/monograph on their past research;

ii) Research grants are sanctioned to an individual scholar or team of scholars for research programme on a theme extending to 3-5 years and costing over Rs.1 lakh; and

iii) Assistance to eminent social scientists of repute, even after retirement, is available in the form of a grant of Rs.10,000/- per annum and would be given on year to year basis at the maximum of three years.

(3) National Information System in Science & Technology (NISSAT)

The NISSAT also encourages individual or team research projects that have relevance to scientific information. It provides assistance to research and development to have greater efficiency in service functions of a library and information system. The schemes will cover the subjects of

- Information science;
- Computer applications in information handling in management;
- Information technologies;
- Scientific and technical translation.

Besides promoting and supporting the conduct of basic and applied research, the NISSAT schemes also support surveys, compilations like union catalogues and directories, literature outlining methodologies/tools/techniques of information handling, information centre development, etc

(4) Financial Assistance from Other Institutions

The Tata Educational and Research Institute (TERI) and Baba Atomic and Research Centre (BARC) also encourage institutional research.

16.5.4 Problems of LIS Research in India

The LIS research in India encounters certain problems which are as follows;

(1) Lack of encouragement for theoretical research: After the death of S.R. Ranganathan, no one is interested in conducting or promoting theoretical research in library and information science in India. Almost all the research works conducted at present are of applied in nature.

(2) Concentration only on limited areas: There is a shortage of investigations in certain subject areas, such as information storage and retrieval, scientific communication and its characterisation, the suitable software for library operations, etc.

(3) The profession at large has not taken any advantage of the research findings. Hence, the innovations and creativity made by researchers are becoming a waste. There is a need to integrate the research results with the needs of the specific libraries.

(4) A majority of research scholars in our field are apparently part-time scholars. The financial support available to these part-time scholars is not encouraging.

Krishan Kumar has identified the following weaknesses of LIS research:

i) Difficulty in dealing with social problems;
ii) Complexity of human behaviour;
iii) Lack of operational situations as a source of hypotheses;
iv) Lack of objectively measurable output;
v) Weaknesses of methods, techniques and tools of research;
vi) Lack of proper academic background;
vii) Lack of resources;
viii) Lack of tradition.

16.5.5 LIS Research Publications

A researcher starts his research with tertiary and secondary sources of information and lands in primary sources to get the required background data.

The following are the primary publications which provide research or scholarly information:

(1) Annals of Library Science and Documentation

(2) Library Science with a slant to Information Studies

(3) ILA Bulletin

(4) IASLIC Bulletin

(5) Library Herald

(6) Herald of Library Science

(7) DESIDOC Bulletin

These sources communicate primary research information.

The newsletters of the professional associations and educational organisation also provide information about the completed research projects as well as the progress of the ongoing research programmes. For example, the University News carries the news about the Ph.D. degrees awarded by the universities in India.

However, there is an inadequacy of secondary sources of research information in India. The indexing journals like *Index India, Guide to Indian Periodical Literature* are the general sources, which also list the articles in library and information science.

The bibliographies on dissertations, treatises and statistical reports and indexing services also guide the researches in locating the relevant articles.

16.6 FUTURE OF *LIS* RESEARCH IN INDIA

The contemporary libraries have become more like information centres. The tools of new technology are introduced from time to time to increase the performance of the libraries to meet the increasing demands of the users. Hence the thrust of future research will be on the application of new technologies to increase the speed and accuracy of information dissemination and to reduce the cost. Hence the LIS research has to concentrate on forecasting the nature of libraries and information centres and their services. In India, both the systems of access, conventional as well electronic, are likely to operate side by side for at least several years to come. However, over a period, the balance will shift towards electronic document delivery system, computer-aided searching devices, etc.

16.7 LET US SUM UP

In this unit, we have discussed about the developments in LIS research in India. The contributions of early libraries to the development of research is minimal. The primary works of Dr.S.R. Ranganathan have paved the way for the promotion of research activity. The last two decades have witnessed a spurt in LIS research in India. A number of universities have started offering courses leading to doctoral programmes. However, institutional research is limited when compared to the Western countries. The contribution of library associations is also minimal.

The UGC and ICSSR are providing financial support for LIS research, but it needs to be further encouraged. The primary publications that disseminate research information are very inadequate. The thrust areas of research need to be critically examined keeping in view the changing information environment and the latest developments in IT and its adoption in libraries and information centres in the country.

16.8 REFERENCES AND RECOMMENDED BOOKS

Dev Raj Sharma. *Doctoral Research in Library and Information Science in Indian Universities.* Palampur: [Author], 1997.

Krishan Kumar. *Research methods in library and information science.* Delhi: Har-Anand & Vikas, 1992. p.305.

Kumar, P.S.G. "Doctoral research in librarianship in India". IN *Handbook of libraries, archives and information centres in India.* Vol.I: Libraries and archives, edited by B.M.Gupta, et al. New Delhi: Information Industry Publications, 1985.

Mangla, P.B. "Research in library and information science in USA, UK and India: An overview". IN *Perspectives in library and information science.* (C.G. Viswanathan Festschrift; Vol.2). Lucknow: Print House, 1982. p.44-61

Raju, A.A.N. "A look into LIS research in India during the past five years". IN *Fifty years of library and information science education in India: Seminar papers*/ edited by A.A.N. Raju, N.Laxman Rao and S.Sudarshan Rao. Hyderabad: IATLIS, 1977. p.186-198.

Sewa Singh. "Research in library and information science: Status and trends" IN *New vistas in library and information science*, edited by A.A.N. Raju et al. (Prof GVSL Narasimha Raju Festschrift). Delhi: Vikas, 1995. p.382-395.

16.9 ASSIGNMENT

(1) Prepare a list of items appeared under the section 'Theses of the month' in University News and identify the major areas of research conducted in Indian universities.

16.10 MODEL EXAMINATION QUESTIONS

I. ESSAY QUESTIONS

1) Write an essay on the contribution of Dr.S.R.Ranganathan for the promotion of LIS research in India.

2) Describe the present status of LIS research in India.

3) Discuss the role of UGC in promoting LIS research in India.

4) Briefly describe the various sources of information used in LIS research in India.

II. SHORT NOTES

a) Library Research Circle

b) Research fellowships