NETAJI SUBHAS OPEN UNIVERSITY

STUDY MATERIAL

MLIS

Paper VII

Research Methodology

Library and Information Science

PREFACE

In the curricular structure introduced by this University for students of Post-Graduate degree programme, the opportunity to pursue Post-Graduate course in any subject introduced by this University is equally available to all learners. Instead of being guided by any presumption about ability level, it would perhaps stand to reason if receptivity of a learner is judged in the course of the learning process. That would be entirely in keeping with the objectives of open education which does not believe in artificial differentiation.

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

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

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

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

Professor (Dr.) Subha Sankar Sarkar
Vice-Chancellor

Second Reprint : November, 2017

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

Post Graduate : Library and Information Science (MLIS)

Paper — VII
Research Methodology

Course Writing

Unit 1 & 6 : Dr. C. R. Sain
Unit 2 & 12 : Barnali Roy Choudhury
Unit 3 : Dr. Chaitali Dutta
Unit 4 : Dr. Swapan Khan
Unit 5 : Dr. Ranjan Samanta
Unit 7 : Anindya Basu
Unit 8 & 14 : Madhusri Ghosh Upadhyay
Unit 9-11 : Anamika Das
Unit 13 : Dr. Susanta Banerjee

Notification

Netaji Subhas Open University

Research Methodology
MLIS – VII

Unit—1 ❑ Basic Concepts of Research

Structure

1.0 Objectives

This unit discusses the meaning and definitions of research. Various kinds of researches and their distinction are described in this unit. The purposes of researches are also explained. Research and its difference with scientific methods are explained.

1.1 Introduction

Man has been curious about the world around him. His earliest attempts to explain the phenomenon of the universe resulted in primitive religious concepts. Gradually, man began to see that the operations of the forces of nature were not as capricious as he had believed. He bagan to observe a certain orderliness in the universe, certain cause-effect relationships, and discovered that under certain conditions, events could be predicted with reasonable accuracy.

This reliance on empirical evidence or personal experience represented a step in the direction of scientific inquiry. It was largely unsystematic observation and only when man began to think systematically about thinking itself that the era of science began.

The first systematic approach to reasoning was through the use of the syllogism, attributed to Aristotle and the Greeks. Syllogistic reasoning established a logical relationship between a major premise, a minor premise and a conclusion. This can be explained in the following example,

Major Premise — Man is a rational being.

Minor Premise — Edward Jones is a man.

Conclusion — Edward Jones is a rational being.

This deductive method of logical analysis made an important contribution to the development of the scientific method.

1.2 Definitions of Research

The term 'Research' is associated with the growth and advancement of knowledge. It is defined as a scientific and systematic search for pertinent information on a specific topic. The Oxford English Dictionary defined research as a "search or investigation directed to the discovery of some fact by careful consideration or study of a subject", and that it is "a course of critical or scientific enquiry". The Advanced Learner's Dictionary of Current English defined research as "a careful investigation or inquiry specially through search for new facts in any branch of knowledge". The Encyclopaedia of Social Sciences (Vol. 13-15) defines research as "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 the practice of an art".

Clifford Woody defines research that comprises defining and redefining problems, formulating hypotheses or suggusted solutions ; collecting, organising and evaluating data ; making deductions and reaching conclusions ; and at last carefully testing the conclusions to determine whether they fit the formulating hypotheses. Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the persuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem in research. As such the term 'research' refers to the systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analysing the facts and reaching certain conclusions either in the form of solutions towards the concerned problem or in certain generalisations for some theoretical formulation.

1.3 Objectives of Research

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

(a) To gain familiarity with a phenomenon or to achieve new insights into it. Studies with this kind of objective are termed as exploratory or formulative research.

(b) To portray accurately the characteristics of a particular individual, situation or a group. Such studies with the objectives in mind are known as descriptive research studies.

(c) To determine the frequency with which something occurs or with which it is associated with something else. Such studies with this kind of objectives are known as diagnostic research studies.

(d) To test a hypothesis of a causal relationship between variables. Such studies are known as hypothesis–testing research studies.

1.4 Types of Research

Basically, research is of two types viz. Basic or Fundamental and Action or Applied. The distinction between the two lies with the purpose or objective of research.

When a research is undertaken without any immediate utilization objective, it is called Fundamental or Basic Research. It is also termed as Pure or Theoretical Research. And when a research is undertaken for some practical purpose, that is, with an immediate utilization objective, it is known as Applied or Action Research.

The National Science Foundation Act (USA) 1950 has defined fundamental or basic research as one in which the primary aim of the investigator is to gain "a fuller knowledge or understanding of the subject under study, rather than a practical application there of". While applied research has been defined as one "that is directed towards practical applications of science". Fundamental research also called theoretical research because of its objectives to discover and enunciate broad generalisations which lead to the development of theories, principles and laws. It reveals the basic relationships among things, try to understand why things happen, seeks to uncover the secrets of nature, tries to give us the fundamental knowledge of things and therfore, it is fundamental or basic in character ; it is pure because it has no ulterior motive, it seeks knowledge for the sake of knowledge, just for the satisfaction of curiosity, the fulfilment of the desire to know.

Applied research is wholly utility oriented. It deals with the problems which pressurise for immediate solutions so that some practical and desired benefit may be obtained. Its objective is not to discover theories, principles or laws, but it tries to utilise the findings of fundamental research, i.e. to develop new materials, new machinery, product or goods. According to the New Encyclopaedia Britanica, Vol. 15, 1974. "Applied research carries the findings of basic research to a point at which they can be exploited to meet a specific need." Fundamental research is totally unconcerned about the practical effect of its result or findings, while applied research is wholly or very much concerned about the effects of its results, which must serve a practical purpose in view.

If the data or findings of fundamental research are utilised by applied research, the data or findings of applied research also provide hypotheses for fundamental research. In this way both the kinds of research have gone to make the universe of knowledge in continuous spiral.

1.5 Research Methods and Methodology

Research methods may be understood as all those methods or techniques that are used for conduction of research. Research methods refer to the techniques that researchers use in performing research operations. It otherwise means that the methods which are used by the researcher during the course of studying his problem are termed as research methods. Since the objective of research, particularly in applied research, is to arrive at a solution for a given problem, the available data and the unknown aspects of the problem have to be related to each other to a make a solution possible. Keeping this in view, research methods can be put into the following three groups :

(i) Those methods which are concerned with the collection of data ;

(ii) Those statistical techniques which are used for establishing relationships between the data and the unknown factors ;

(iii) Those methods which are used to evaluate the accuracy of the results obtained.

In brief, research methods are but tools of research for data collection and for establishing relationship between facts.

Methodology means a systematic procedures and techniques required to be followed for accomplishing an activity. It is a way to systematically solve the research problem. Research methodology is a body of methods, that is, procedures and techniques of collection, organisation, analysis and evaluation of data or facts which are appropriate for a specific problem. Researchers not only need to know how to develop certain indices or tests, how to calculate the measures of central tendency or how to apply particular research techniques, but also they need to know which of these methods are relevant and which are not in a particular problem. All this means that it is necessary for the researcher to design his methodology for his problem as the methodology may differ from problem to problem. In research, the scientist has to expose the research decisions to evaluation before they are implemented. He has to specify very clearly and precisely what decisions he selects and why he selects them so that they can be evaluated by others.

From above, we can conclude that research methodology has many dimensions and research methods do constitute a part of the research methodology. In brief, research methodology consists of a series of tasks undertaken from the selection of the problem to the analysis of data in a research study.

1.6 Research and Scientific Method

The terms research and scientific method are closely related. Research, as we have discussed, can be stated as "an inquiry into the nature of, the reasons for, and consequences of any particular set of circumstances, whether these circumstanes are experimentally controlled or recorded just as they occur. A scientific method is the way in which one can test opinions, impressions or guesses by examining the available evidence both for and against them. It is simply the pursuit of truth which is determined by logical considerations. Scientific method is the most assured technique for controlling a host of things and establishing stable belief. The ideal of science is to achieve a systematic interrelation of facts.

The scientific methods include logical reasoning, and depends on both deductive and inductive kind of reasoning. Deductive reasoning proceeds from general to particular while inductive reasoning proceeds from the analysis of particular facts or events to a general conclusion. The deductive and inductive process is the essence of scientific method.

Logic and Scientific Method :

Logic involves reasoned knowledge and all sciences are applied logic. The universal feature of science is its general method which consists in the persistent search for truth. But the search for truth depends on evidence, the determination of which we call logic.

Scientific method is an enquiry to ascertain the validity of beliefs, hypotheses and propositions through factual evidence. In scientific method, logic aids in formulating propositions explicitly and accurately so that their possible alternatives become clear. Further, logic develops the consequences of such alternatives and when these are compared with observable phenomena, it becomes possibe for the researcher or the scientist to state which alternatives is most in harmony with the observed facts. All this is done through experimentation and survey investigations which constitute the integral parts of the scientific method.

The scientific method is, thus, based on certain basic postulates which can be stated as follows :

(i) It relies on empirical evidence ;

(ii) It utilizes relevant concepts ;

(iii) It is committed to only objective considerations ;

(iv) It presupposes ethical neutrality, i.e. it aims at nothing but making only adequate and correct statements about population objects ;

(v) It results into probabilistic predictions ;

(vi) Its methodology is made known to all concerned for critical scrutiny and for use in testing the conclusions through replication ;

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

Thus scientific method implies an objective, logical and systematic method, i.e. a method free from personal bias or prejudice, a method to ascertain demonstratable qualities of a phenomenon capable of being verified, a method where in a researcher is guided by the rules of logical reasoning and the method where in the investigation proceeds in an orderly manner.

1.7 Some other types of researches

(i) Descriptive vs. Analytical

Descriptive research includes surveys and fact-finding enquiries of different kinds. The major purpose of this descriptive research is to describe the state of affairs as it exists at present. The main characteristics of this method is that the researcher has no control over the variables. Researcher can only report what has happened or what is happening. In analytical research, the researcher has to use facts or information already available and analyse these to make a critical evaluation of the material.

(ii) Conceptual vs. Empirical

Conceptual research is related to some abstract ideas or theories. It is concerned with the works of philosophers or thinkers to develop new ideas or to reinterpret an old ones. On the other hand, empirical research is concerned with experience and observation alone. It is data-based research where conclusions are drawn from the data collected and is capable of being verified by observation or experiment. It can also be called as experimental type of research.

(iii) Quantitative vs. Qualitative

Quantitative research is based on measurement of quantity. It is applicable to phenomena that can be expressed in terms of quantity. Qualitative studies, on the other hand, is concerned with qualitative phenomenon. It aims at discovering the underlying motives and desires through depth interviews. Attitude or opinion research, i.e. research designed to find out how people feel or what they think about a particular subject or institution is also qualitative research. We can analyse various factors which motivate people to behave in a particular manner or which make people like or dislike a particular thing.

(iv) Few others depend on time that carried on over time, for many years. This type of research is known as longitudinal research where the characteristics of a subject or phenomenon and the way in which these characteristics change with growth and development. Research may be a laboratory type where research is carried out in an controlled environment. Sometimes, research may be clinical or diagnostic research. In this type, case study method or indepth approaches to reach the basic causal relations. The exploratory type aims at the development of hypothesis rather than their testing. Historical type of research utilises historical sources like documents, remains, etc. to study events or ideas of the past at any remote point of time. In management studies where decision making is an important aspect in the field. Such a decision making process is achieved through operations research.

1.8 Exercise

1. What do you mean by research and what are the objectives of research ?
2. Discuss the various types of resarch.
3. Make a distinction between Basic Research and Applied Research.
4. How is research differs from scientific method ?
5. Short notes :
   (a) Research Methodology.
   (b) Benefits of Research.
   (c) Empirical Research.

1.9 Summary

This unit introduces us with the various definitions of research and various types of researches. It also introduces the distinction between research and scientific method. Basic concepts of research are discussed in this unit.

1.10 References

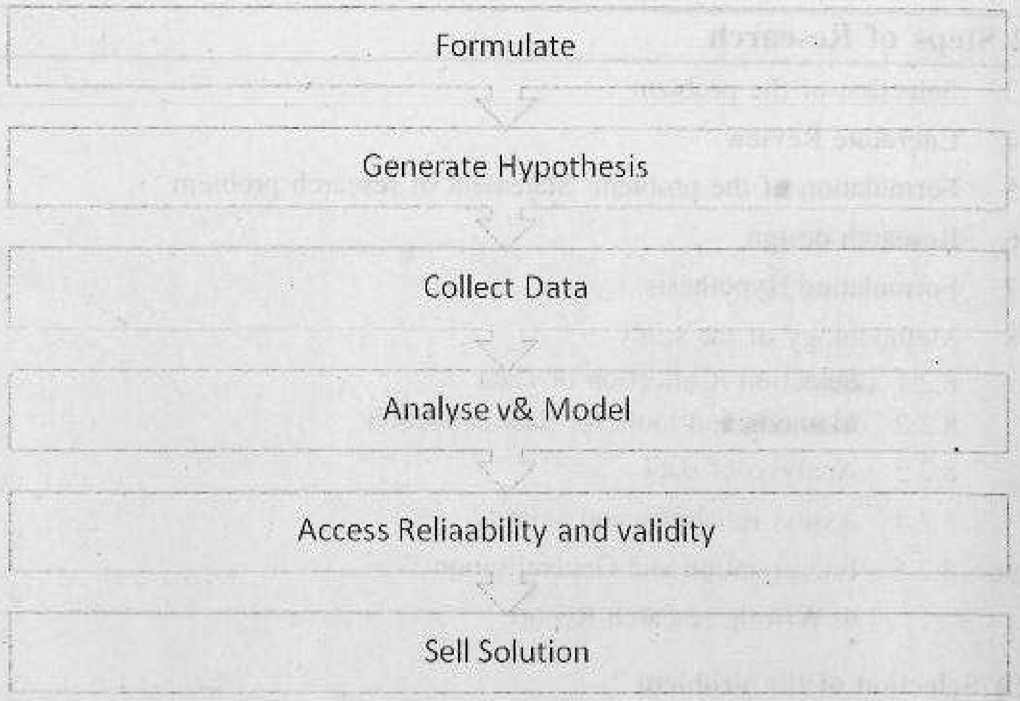
1. Ghosh, B. N. Scientific Method and Social Research. New Delhi : Sterling Publishers, 1987.
2. Goode, W. J. and Hatt, P. K. Methods in Scocial Research. London : Mc Graw Hill Book, 1981.
3. Kothari, C. R. Research Methodology : Methods and Techniques. 2nd. ed. New Delhi : Wishwa Prakashan, 1990.
4. Ohdedar, A. K. Research Methodology. Calcutta : Bengal Library Association, 1993.

Unit : 2 ❑ Steps of Research

Structure

2.1 Introduction

This unit is dealt with a brief discussion about the steps of research methods in special reference to library and information science. Generally, there were no such preseted dejure or defacto rules and regulations to conduct a research work. Therefore, researcher may follow few steps for their further reference/guidance. Research is a scientific process of organising, planning conducting, analysing and reporting. Researchers have to achieve their goal oriented research outcome within a stipulated time frame. Karl Raimund Popper 1979 expressed research process as following:



(Source: Adams. J et al. (2014), Research Methods for Business and Social Science Students, Sage Publication.)

Research is a purpose of search and a newly observed part of previously established knowledge. It is a method of searching for new particular aspect of any particular field of study. It applies a scientific method in the study of problem. It has following salient feature.

- It is an original work.
- It focusses a particular field of study
- It is unbiased and it is not influenced by any opinion or school of thoughts
- Its quality depends on the attitude of the researcher(s)
- As the research is based on scientific method, then its main basis is logical appearance.
- It is based upon observer-able experiences or empirical data.
- Therefore, it demands accurate observation and description. It also inculcates scientific and inductive thinking.

2.2 Steps of Research

3 Selection of the problem

4 Literature Review

5 Formulation of the problem/ Statement of research problem

6 Research design

7 Formulating Hypothesis

8 Methodology of the study

8.2.1 Selection /Collection of Data

8.2.2 Methods and tools for data collection

8.2.3 Analysis of data

8.2.4 Assess reliability and validity

8.2.5 Interpretation and Generalisation

6. Writing research Report

2.2.1 Selection of the problem

First and most important stage of a research work is to find out/ formulate the actual problem by understanding the subject property of his/her own interest initially. In this stage researcher should be aware of his/her ability/command over the topic, available literature on the topic and also availability of research guide who may direct him/her on a right way within stipulated time frame. Selection of suitable research problem in other way may lead research work smoother and efficient. It is thus selection of research problem which has high volume to the society and to identify those problems that need to be shorted out after choosing a right research gap for a study. It is a difficult task as it depends on the time, effort and commitment on the part of a researcher. However, three points are to be kept in mind as below:

- What are the contemporary interest?
- What is your own interest?
- What is the gap in the field of the study?

2.2.2 Literature Review:

Literature Review is a major component of a research. What the review should be conducted/focused on, that important part of the research work is to be communicated here. This part basically includes spaces for critical reading in order to improve domain specific knowledge and skill development. Literature review is important in each and every steps of research work to understand as follows:

Which part(s) of the undertaken domain has already been explored?

What are the main and in-depth theoretical perspective?

Who are the experts in this field?

What are the gaps still remains to explore the particular field?

What are the main problems to conduct research work on the said topic?

Is the topic being open to solve identified problems?

The relevant publications are to be carefully studied. The main purpose of this study is to indicate the problems that are already studied and those are yet to be investigated. Review requires a critical understanding of the literature that demonstrates the height order intellectual skills of analysing, evaluating and creating. This would help the researcher to know how the same are conducted, the methodology employed issues covered and the further works has been suggested. Only proper literature review process may relate all questions and may make researcher able to have vivid explicit and pinpointed idea on the said topic. This process is linking/bridging the gap between research questions and research findings with proper evidence in terms of previous scholarly communication. It is also a parallel guide to frame research work which may increase confidence, courage and strength to connect future work with the previous work. In this case, researcher may follow a good number of resources including Journal articles, Monographs, Thesis and Dissertations, Indexes, abstracts, bibliographies, Encyclopaedias, Handbooks etc. to have in-depth knowledge. Researcher may use resources under Creative Commons license materials to avoid academic dishonesty. Try to plan your search in terms of scope, range and timescale. Do critical reading, critical evaluation critical analysis and critical thinking of existing literature and formulate research Problems/questions.

2.2.3 Formulating Research problem:

Besides selection equally important is the formulation of the problem. After having critical and depth study of specified topic researchers are able to frame a range of questions on the base idea like who, what, when, how much, what the reason for, what are the functions, what is the process, what are the cause/result, what is its applicability, what are the conflicts/arguments etc. An effective problem formulation involves the following:

1. Definition of the problem
2. Scope of the problem
3. Justification of the problem
4. Feasibility of the problem

A researcher must have to answer the sources of data or information those are very relevant and needful in research work. Two types of sources are used i) Primary Data and ii) Secondary Data. Published literature and now-a-days web information are being popularly used.

2.2.4 Formulating Hypothesis:

Hypothesis mainly depends on anticipation. It is uncertain. The hypothesis is a tentative solution of the research problems. This provisional idea is literally based on a limited amount of evidence. Once the evidence has been gathered and the hypothesis has been tested it becomes the statement that the research work sets out to defend proposes to develop and attempts to prove (John Dewey, 1993). It is specified and object oriented task to achieve particular goal.

2.2.5 Research design:

Research design is a plan of action to pursue a research work which indicates the actual actions to be followed to achieve research goal. It includes study types (descriptive, experimental, survey), research questions, hypothesis, variables (controlled, uncontrolled), methods of data collection and plan for statistical data analysis. This is the place where researcher will identify and state the exact method for his/her research work for effective outcome of the research undertaken.

2.2.6 Methodology of the study:

2.2.6.1. Selection /Collection of Data: The core and relevant data are to be collected for analysis and interpretation. Primary data and on the basis of it secondary data should be collected for fruitful execution of research work.

2.2.6.2 Methods and tools for data collection: Researcher may follow any of one or more than one method(s) to pursue his/her research work viz. Survey, Interview and focus groups.

2.2.6.3 Analysis of data: Creative attitude and efficiency can help to achieve proper analysis of data. Immaterial data/information are to be omitted/discarded. In this stage the collected data would have to be classified after collecting data, the researcher may have to be utilized the data based on his need and purpose. Sometimes the difference and comparison of data are also being made from the perspective of the study.

The collected data may not express the real truth and thus the conscious observation is required for arranging data and fruitfulness of the collected data are also to be verified before organisation and arrangement of data. Qualitative data analysis and quantitative data analysis and Statistical analysis may take plac here if necessary.

2.2.6.4 Assess reliability and validity: Reliability and Validity of data must be measurable. How far collected datais reliable and validated and authenticated, that should be assessed.

2.2.6.5 Interpretation and Generalisation:

Selection, collection and interpretation of data is very much interrelated with each other and work collectively. Each and every entity effect each and every part of this circle. Researcher have to state his/ her own findings regarding researcher work should properly have communicated here logically and methodically.

2.2.7 Writing research Report :

This section is going to discuss on how to write a research report by giving particular emphasis on different sections of research step by step. Research report is to be written in three parts all over. **First part** includes Title Page, Preface, Acknowledgment, List of Tables, List of Abbreviations, Content page. **Second part** includes Introduction, Literature Review, Objectives of the study, Statement of Research Problem, Data Collection, Data Analysis, Interpretations, Findings, Conclusion & Recommendations by following **Third part** including References (by following referencing style APA 6th edition), Index, Glossary.

Unit 3 ❑ Ethical and Social Aspects of Research

Structure

3.0 Objectives

This unit intends to provide an understanding of the social and ethical aspects of research, which are no less important considerations than any other aspect of research. A researcher should always be aware of the elements of risks involved in neglecting social norms and sentiments, as well as the ethical bindings that a human society follow.

3.1 Introduction

Some researchers work on human subjects. In medical sciences, human and other living creatures are often used. In such cases there may be certain risk factors, which should be taken into cosideration by the researchers. However minor those factors are, there may be certain serious implecations. According to Best and Kahn[1] : "While these risks may be minor and the importance of the research may be high, the risks must be considered. ... Any set of rules or guidelines that attempts to define the ethical limits for all human researches raises controversy among members of the scientific community and other segments of society."

It is true that rigid controls often hamper or restrict the effectiveness and spontaneity of research. But without some restraints, chances of serious injury to and infringement on human rights remain as a possibility.

3.2. Need for Ethical Guidelines

Considering the risk factors and their impact on society and research itself, a few things are to be taken care of. Those are :

3.2.1. Informed Consent

The researcher may involve any human being as his subject of research, depending on a clear consent on the part of the subject. The information, a researcher obtains through research on human subjects is of immense value and potential. On the other hand, the researcher must be fully aware and responsible for the well beings and welfare of the subject. Chances of physical harm to the subjects, during research are remote in the area of social sciences ; but the possibility of such harm in physical and applied sciences can not be ruled out. Hence, informed consent from the subject or participant in any research is a must.

3.2.2. Confidentiality

The researcher should maintain strict confidentiality of the information he/she obtains on or about the human subject. Respect for the privacy of the subject should be maintained. No information on the subjects should be disclosed without their permission.

Any filming or recording of the subjects, done in research may cause identification and / or harm to the subjects. The reputation, social image or position of the subject in the society may be damaged for these recordings or filming. Therefore, as the best practice, identity marks of the huamn subjects should either be obfuscated or psuedonyms be used.

3.2.3. Protection from Physical and Mental Stress

The researcher should try his/her best to keep the subject of his / her research free from any sort. of physical or mental stress. In fact, the researcher should not normally deceive the subjects about significant aspects of the research. All possible precautions should be taken to protect the physical well being as well as emotions of the subjects.

3.2.4. Honouring One's Self Respect

The dignity and self respect of every individual should be protected with great care. Other wise it may cause irritation and humiliation to a subject which is not at all desirable and congenial for the research and researcher.

3.2.5. Knowledge of Outcome

Everybody has a right to know any fact or information about the nature, results and conclusions of the research and it is the duty of the researcher to provide the same, as promptly as practicable.

3.2.6. Fabrication of Data

THe researcher should be very careful in presenting the data that have been collected by him/her during research. Data should never be manipulated or fabricated. Transparency and truthfulness are the two important parameters of a good research.

3.2.7. Plagiarism

Plagiarism is an ethical issue. It is the practice of using some other person's or expressions by the researcher in his/her writing without mentioning or acknowledging the source. The researcher should refrain from following such practices.

3.2.8. Honouring Comkitments

Researchers should make sincere efforts to honour all commitments they have made to subjects/participants.

Considering above points, one should admit the need of a code of ethics for research.

3.4. A Brief History of Research Ethics

The concept 'research eithics' first came to light after world war II. During the Nuremberg Trials, people first came to know about the dangerous and harmful 'research' that had been conduced in the concentration camps by the German physicians. In the so-called 'research' millions of innocent people were put to death due to the inhuman attitude and unethical mindset of those physicians[1]. Best and Kahn[1] said : "In 1932, a study began on long-term effects of syphilis. For this study by the United State Public Health Service, 399 African-American men with syphilis were denied treatment and told they had 'bad blood'. This study continued unitl 1972 during which time the physicians in charge of the study made sure their patients / subjects did not find out the truth and did not receive any appropriate treatment for syphilis. In the 25 years since the study become public, it has been considered a powerful statement regarding racism and ethical misconduct."

These type of researchers on becoming public shook the conscience of the society and naturally voices were raised against such unethical practices. As a result, in 1953, the first code of ethics for psychologists was issued by the American Psychological Association. It was a milestone in the history of research ethics. It was revised in 1963 and in 1970, the Board of Directors appointed an Ad Hoc Committee on Ethical Standards in Psychological Research forupdating the 1963 code in the light of Changing social contexts and the changes taking place in science as well as in professions. A series of debates, conversations and correspondence was going on involving almost all cross-sections of the society resulting in the final draft, adopted and published in 1973[1].

In December 1992, the ethical standards were published by the American Psychological Association and in the same year, the American Educational Research Association (AERA) adopted the Ethical Standards of the AERA. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research was established by the United States Congress in 1974.

Some of the significant events and studies are discussed here for a clear understanding.

3.4. Conclusion

Researchers should value their research but at the same time should not forget their responsibilities towards their subjects, co-researchers, and to the public as well. They should not adopt unfair practices and should not discard unfavourable data. Invasion of privacy, lack of confidentiality or harm to the subjects in the name of research should be avoided by any means. Appropriate credits should also be given to those who have provided help and support to their research.

3.5. Exercise

1. What do you understand by 'Plagiarism'?
2. Why ethical guidelines are needed in conducting a research fairly?—Discuss.

3.6. References

1. Best, J. W. & Kahn, J. V. (2003). Research in education (9th ed.). New Delhi : Prentice-Hall of India.

2. umar, P. S. G. (2004). Research methods and statistical techniques. New Delhi : B. R. Publications.

Stop '*gendercide*' efforts

[illegible]	[illegible]	[illegible]
Women	·Husband ·In Laws ·Parents	·Doctors ·USG Clinic owners

In 1980s, 90s & 20s; respectively 2 , 4 & 6 million female foeticides conducted in India. It is no exaggeration to call this gendercide. Women are missing in their millions-aborted, killed, neglected to death .

If female foeticide is done , who are involved?

Stakeholders can make a difference. Result is enormous through implementation of the programme effectively.

Unit 4 ❑ Research Methods : Quantitative

Structure

4.0 Objectives

After reading this Unit, reader will be able to:

- define quantitative research;
- understand the nature of this type of research;
- recognize its advantages and disadvantages;
- identify the research situations appropriate for quantitative study;
- explain the basic research design for quantitative study;
- specify the aspects of quantitative research;
- identify the problems faced in conducting quantitative research;
- find the scope of its application in library and information science;

4.1 Introduction

Quantitative methods highlight objective measurements and the statistical, mathematical, or numerical analysis of data collected through polls, questionnaires, and surveys, or by influencing pre-existing statistical data using computational techniques. Quantitative research focuses on gathering numerical data and generalizing it across groups of people or to explain a particular observable fact.

The collection of information in quantitative research is what sets it apart from other types. Quantitative research is focused specifically on numerical information, also known as 'data.' Because the research requires its conductor to use mathematical analysis to investigate what is being observed, the information collected must be in numbers.

For quantitative data analysis in this type of research, issues of validity and reliability are significant. Quantitative researchers endeavour to demonstrate that their selected methods succeed in measuring what they purport to measure. They want to make sure that their measurements are stable and consistent and that there are no errors or bias present, either from the respondents or from the researcher.

The analysis can be left until the end of the data collection process, and if it is a large survey, statistical software is the easiest and most efficient method to use. For this type of analysis time has to be put aside for the data input process which can be long and laborious. However, once this has been done the analysis is quick and efficient, with most software packages producing well presented graphs, pie charts and tables which can be used for the final report.

To better understand this style of research it is needed to break down its major doctrine. There are three: observing and explaining something that happens, collecting information, and analyzing the information. The combination of these three parts is at work when presenting clear and well-researched findings.

Observing and explaining occurrences is the first step. The search for this explanation can be presented in the form of a question. It can also be expressed as a hypothesis. In the case of a hypothesis the search for an explanation is made as a statement to be proved or not – depending on the goals of the researcher.

Self Check Exercise :

1. Define quantitative research. Why it is called 'quantitative'?

Answer..
....
....
....
....
....
....
....
....

4.2 Types of Quantitative Research

Quantitative research primarily can be categorized in two types – experimental and non-experimental. Experimental research method consists of three studies i.e. True experimental, quasi experimental and pre-experimental research. Non-experimental mainly leads to descriptive study which includes survey, correlational, comparative study, evaluation study, and ex-post facto study. The difference among them primarily relates to the degree the researcher designs for control of the variables in the experiment.

Quantitative Reaearch Methods

Experimental
- True experimental
- Quasi experimental
- Pre-experimental

Non-experimental
- Correlational
- Survey
- Comparative
- Evaluation
- Descriptive (Ex-post facto)

4.2.1 Experimental

4.2.1.1 True Experimental:

This research method often called experimental research; use the scientific method to establish cause-effect relationship among a group of variables in a research study. Researchers make an effort to control for all variables except the one being manipulated (the independent variable). The effects of the independent variable on the dependent variable are collected and analyzed for a relationship. The true experiment is often thought of as a laboratory study, but this is not always the case; a laboratory setting has nothing to do with it.

Characteristics of Experimental Research:

There are several characteristics involving experimental research which distinguish experimental research from other research methods. These key ideas, as presented by Creswell (2008) include:

- Random assignment
- Control of extraneous variables
- Manipulation of treatment
- Measurement of outcomes
- Comparison of participant groups
- Possible threats to validity

Moreover, this research design provides unbiased approximation of the variable effects and associated doubts; enables the experimenter to detect important differences. It permits conclusions that have wide validity and shows the direction of better results.

Examples :

- The effect of positive reinforcement on attitude toward school
- The effect of teaching with a traditional lecture approach on students' achievement
- A comparison of the effect of personalized instruction vs. traditional instruction on computational skill
- Law Of Segregation - The Mendel Pea Plant Experiment
- Transforming Principle - Griffith's Experiment about Genetics

Experimental research is guided specifically by a hypothesis. Sometimes experimental research can have several hypotheses. A hypothesis is a statement to be proven or disproved. "It is an assumption based statement of a proposition or a reasonable guess which the researcher seeks to prove through his study". Once that statement is made experiment is started on to find out whether the statement is true or not. This type of research is the bedrock of most sciences, in particular the natural sciences.

4.2.1.2 Quasi Experimental

Quasi-experimental research may look very much like true experimental research in that it does involve the manipulation of an independent variable. Quasi-experimental research studies lack one or both of the essential properties of randomization and a control group. It has a weakness in that is not possible to deliver 'cause and effect' results. In other words, researcher cannot conclude from this research that, for example, doing one thing causes a particular phenomenon (e.g. smoking cigarette causes cancer).

When performing a study, a researcher is attempting to show that variable 'A' influences variable 'B' to do something. They want to demonstrate cause and effect. Random assignment helps ensure that there is no pre-existing condition that will influence the variables and mess up the results.

Quasi-experimental designs include:

- time series design
- multiple time-series design
- regression/discontinuity analysis
- equivalent time samples design
- separate sample pre-test/post-test design
- separate sample pre-test/post-test control group design
- nonequivalent control group

Types of Quasi-Experimental Research

Quasi-Experimental Research can be categorized into three types of quasi-experimental research: cross-sectional, longitudinal, and cross-sequential.

1. **Cross-sectional** research studies make a comparison of different groups at the same time. In human growth and development, cross-sectional research is most often used to look at different groups of people at different ages.
2. **Longitudinal** research studies look at one individual or one group over a period of time. An example of longitudinal research would be a study where a group of two-year-old children are separated into those who use processed food and those who do not.
3. **Cross-sequential** research studies compare two separate, but similar, longitudinal research studies that are done at different times. Cross-sequential research studies are much more complicated, expensive, and time consuming than other research designs. Because of this, it is seldom used, even though it is a very effective way of collecting useful data.

4.2.1.3 Pre-Experimental

This method follows basic experimental steps but fail to include a control group. It is also called a weak experiment because the design lacks essential components of a true experiment: random assignment of participants to groups and manipulation (Suter, 2006). Pre-experimental design is usually undertaken for exploratory purposes. Typical of pre-experimental design is the elimination of a control group, thus it is often called a single-group experiment. A single group is often studied but no comparison between equivalent non-treatment groups is made. Pre-experimental research is needed because there are many independent variables that the researcher cannot manipulate, either ethically, or practically. For example, some quantitative variables that cannot be manipulated are age, intelligence, personality traits etc. Pre-experimental research is three types:

(i) **One-shot case study design:** In this research a single group is studied at a single point in time after some treatment that is presumed to have caused change. It focuses on to evaluate the influence of a variable and attempts to explain a consequence by an antecedent. No control or comparison group is employed in this study. This design is diagrammed as:

    Group 1 → Treatment Implemented'! Posttest Administered

(ii) **One group pre-test/post-test design:** Generally a single case is observed at two time points, one before the treatment and one after the treatment. Changes in the outcome of interest are presumed to be the result of the treatment. This method does not include control or comparison group. A benefit of this design is the inclusion of a pretest to determine baseline scores. This design can be diagrammed as:

    Students Assigned to Group → Pretest Administered → Treatment Implemented → Posttest Administered

(iii) **Static group comparison design (cross-sectional study):** A group that has experienced some treatment is compared with one that has not. Observed differences between the two groups are assumed to be a result of the treatment. This design can be diagrammed as:

    Intact Groups → Group 1 → No Treatment Implemented → Posttest Administered → Group 2 → Treatment Implemented → Posttest Administered

4.2.2 Non-Experimental

4.2.2.1 Correlational

Correlational research examines the relationships between two variables using statistical analyses. It is done to establish what the affect of one on the other might be and how that affects the relationship. In correlational research the survey is conducted on a minimum of two groups. In most correlational research there is a level of manipulation involved with the specific variables being researched. However, it does not look for cause and effect and therefore, is also mostly observational in terms of data collection. For example, just because two data points synchronization doesn't mean that there is a direct cause and effect relationship.

This type of research recognizes trends and patterns in data, but it does not analyse or prove these observed patterns. Cause and effect is not the basis of this research. The data and their relationships, and distributions of variables are studied only. In this study variables are not manipulated; they are only identified and are studied as they occur in a natural setting.

Examples:

- The relationship between intelligence and self-esteem
- The covariance of smoking and lung disease
- The relationship between diet and anxiety
- The relationship between an aptitude test and success in an algebra course

4.2.2.2 Survey

Survey research is one of the most important areas of measurement in applied social research. The broad area of survey research encompasses any measurement procedures that involve asking questions of respondents. This research uses questionnaires, interviews, and sampling polls to get a sense of behavior with strong precision. It allows researchers to judge behavior and then present the findings in an accurate way. Survey research can be conducted around one group specifically or used to compare several groups. There are many different types of surveys, several ways to administer them, and many methods of sampling. Some data collections methods are:

Questionnaires: a series of predefined questions for data collection from individuals. It is two types in nature:

Closed-ended questions - predetermined answers are given generally 'yes' 'no' type from which respondents choose their answer.

Open-ended questions - respondents are asked to answer each question in their own words i.e. answers are descriptive in nature.

Sampling: It is a technique of selecting units from a population of interest so that by studying the sample researcher may fairly generalize the results back to the entire population of interest. Many types of sampling methods are used in survey research. Types of sampling are listed below:

Probability Sampling: It utilizes some form of random selection. The method sets up some procedures that assure that the different units in the population have equal probabilities of being chosen. Various types of probability sampling are:

- Simple Random Sampling
- Stratified Random Sampling
- Systematic Random Sampling
- Cluster (Area) Random Sampling
- Multi-Stage Sampling

Non-Probability Sampling: In this method some elements of the population have no chance of selection or where the probability of selection can't be accurately determined. It involves the selection of elements based on assumptions regarding the population of interest, which forms the criteria for selection. The selection of elements is nonrandom. Types of non-probability sampling are as follows:

- Accidental or Convenience Sampling
- Purposive Sampling
- Quota Sampling
- Expert Sampling
- Snowball Sampling
- Heterogeneity Sampling
- Modal Instance Sampling

Interviews: In quantitative research (survey research), interviews are more structured than other researches. In an interview method, the researcher asks a standard set of questions for data collection.

- Face-to-face interview – The researcher establishes rapport with potential participants and therefore gains their cooperation. This method yields highest response rates in survey research.
- Telephone interview – This method is less time consuming and less expensive and the researcher has ready access to anyone having a telephone. The response rate in this method is not as high as the face-to- face interview.
- Computer Assisted Personal Interviewing (CAPI) – It is a form of personal

interviewing where the interviewer brings along a laptop or hand-held computer to enter the information directly into the database. Though this method is expensive to set up it saves time involved in processing the data.

4.2.2.3 Comparative

Comparative research aims to make comparisons across different countries or cultures. This method has long been used in cross-cultural studies to identify, analyse and explain similarities and differences across societies. A major problem in this research is that the data sets in different countries or across societies may not use the same categories, or define categories differently (e.g. definitions of poverty). There are many names this research: comparative public opinion, cross-national public opinion, or even comparative political behavior.

This research is a broad expression that includes both quantitative and qualitative comparison of social entities. Social entities may be based on many shapes, such as geographical or political ones in the form of cross-national or regional comparisons. The ultimate goal of comparative research is to search for similarity and variance of the entities. Those searching for similarity often apply a more general theory and search for universals or underlying general processes across different contexts. To conduct comparative research some rules must be followed carefully. Some of the basic rules of comparative research have been provided by Roger Jowell (1998) as follows.

- **Knowledge about country:** Researchers should not to interpret survey data relating to a country with little knowledge.
- **Limit the number of countries:** They should oppose the appeal to compare too many countries at once.
- **Contextual variables matter as well:** Cross-national surveys should pay as much attention to the choice and compilation of aggregate-level contextual variables.
- **Aware of limitations:** Researchers should be as open about their limitations as they are enthusiastic about their explanatory powers.
- **Rules for methods:** To transform cross-national surveys from parallel exercises into joint ones, collective development work, experimentation, scale construction, and piloting should be undertaken in all participating nations.
- **Be critical of findings:** Analysts of cross-national data should try to suspend initial belief in any major inter-country differences they discover.

In this study researcher emphasizes diversity of factors and the focus is on the similarities within a category of cases with the same outcome that (i) distinguish that category from other categories (countries with other forms of austerity protest) and (ii) explain the outcome manifested by that category. In other words, the study of diversity is the study of patterns of similarities and differences within a given set of cases.

4.2.2.4 Evaluation

This is an "applied form of research that involved finding out how well a programme, practice, procedure or policy is working" (Polit & Hungler, 1999). On other words "assessment study with value judgement is the evaluation study of research. This type of research is directed to evaluate the performance of a project, or programme, process or product that has already been started or implemented" (Ohdedar, 1993).

Assessment + Value = Evaluation

Three types of evaluation research are:

(1) Concurrent evaluation – It evaluates ongoing programme simultaneously;
(2) Phase evaluation – This research evaluates programme at its different phases;
(3) Terminal evaluation – At the end of a programme this research is carried out.

Examples of evaluation research:

- Process or implementation analysis;
- Outcome analysis and the analysis of the outcome of changes in processes;
- Impact analysis or the impact of a new treatment;
- Cost-benefit analysis and so on;

As with the survey research methods, this type of research may best be carried out as a qualitative piece or research, depending upon the original research question.

4.2.2.5 Descriptive (Ex-post facto)

Descriptive research does not fit neatly into the definition of either quantitative or qualitative research methodologies, but instead it can utilize elements of both, often within the same study. This type of research includes surveys and fact-finding enquiries of different programmes. The main purpose of descriptive research is to describe the state of affairs as it exists at present. Sometimes descriptive research is used as synonymous term of survey research. This research describes and interprets *what is.* This research is called ex-post facto research as it uses variables that are already happened. It is also called exploratory observational or casual comparative method. "It deals with conditions or relationships that exist, opinions that prevail, processes that are going on, effects that are manifest, or trends that are developing. It is mainly concerned with the present though it often considers past events and influences as relating to current status" (Ohdedar, 1993).

Most quantitative research covers two areas: studies that describe events and studies aimed at discovering deductions or causal relationships. Descriptive studies are aimed at finding out "what is," so observational and survey methods are normally used to collect descriptive data.

The purpose of descriptive research is to

- observe;
- describe and
- document

This research involves the collection of data that will provide an account of individuals, groups or situations. Instruments involved to obtain data for this research include

- questionnaires;
- interviews (personal interviews with the aid of study guide, closed questions);
- observations (participatory or not, checklists, etc.)

Characteristics of descriptive research:

- It does not reply questions about how/when/why the characteristics occurred;
- It gives meaning to the quality and standing of facts that are going on;
- It involves collection of data in order to test the hypothesis or to answer questions;
- The variables in descriptive research are not usually controlled.
- It is criticized for its inability to control variables;
- Study compares the characteristics of two groups or cases to determine their similarities or/and differences.
- It reveals problems or abnormal conditions to provide basis for decision making;
- Study may or can be repeated for purposes of verification and comparison.

Types of descriptive research:

On the basis of the observation of variables descriptive research is categorized into three types.

(1) **Cohort method** – This method looks forward in time and observes variables for long period of time. It is called observational study. So, observational studies are all about observing people, and they occur into two settings.

(i) **Naturalistic:** It is called field observation where a researcher observes the subject in its natural environment, collect data . and drew conclusions from this. This makes the observations more true to what happens in the chaotic, natural world.

(ii) **Laboratory observation:** In this type of study a researcher observes the subject in a laboratory setting. This gives the researcher a little more control over what happens. An example of a laboratory observation in psychology would be done to understand something about children at a certain age.

(2) **Cross-sectional method** – It measures to see if a substance or activity is related to other. Generally it looks at current or present activity. In this study researchers record information about their subjects without manipulating the study environment. A researcher can measure the cholesterol levels of daily walkers and non-walkers along with any other characteristics that might be of interest to the researcher. Here researcher never influence non-walkers to take up that activity, or advise daily walkers to modify their behaviour.

However, cross-sectional studies may not supply specific information about cause-and-effect relationships. This is because such studies offer a picture of a single moment in time; they do not judge what happens before or after that event.

(3) **Case-control method** – It compares one group to another and finding their differences if any using backward data analysis. This study was historically borne out of curiosity in disease etiology. The study is designed on the basis of the case history of a patient. The diseased patient is questioned and examined, and elements from this history taking are joined together to disclose factors that are liable for disease.

The goal of this research is to retrospectively resolve the experience to the risk factor of interest from each of the two groups: cases and controls. These studies are carried out to determine the abnormal from the group. Sometimes this study is called "retrospective studies" or "case-referent studies." This study gains some advantages of short time span, look at multiple risk factors, quick answering system etc. But it suffers from a problem of data quality which mainly depends on memory.

Examples of Descriptive Research:

Generally the study questions in descriptive study start with "What is...". For example, descriptive researches want to resolve the following questions:

- What is the impact of social networks on changing attitude of college students?
- What is the intangible motivational factor in increasing output of a factory?
- What is the average age at which children learn to speak?
- Why does school environment control student behaviour?

This research includes both quantitative and qualitative data to describe the population being observed. For example, it can be observed that why certain groups of workers are suffering on eye diseases, while others of the same type and in the same working conditions are good in health.

This research can be done about the effects of the social network sites on changing human behaviour, study about mass media effectiveness, study of the relationship between a particular TV programme and it's appeal to people and so on.

The above examples are fit into three types of descriptive studies:

Simple descriptive: The study wants to know a single answer from a study such as percentage of student having autistic problem.

Comparative descriptive: It describe and compare multiple groups. Generally the research question is like what is the percentage of boy vs. girl students who are depressed?

Correlational: The study describes the statistical relationship between two or more variables. For example research question wants to find the student-teacher ratio in each classroom and calculates the average student achievement on the state assessment.

Self Check Exercise :

1. How will you categorize quantitative research?
2. Why true experimental research is sometime called 'experimental research'?
3. In which research 'cause-effect' relation does not arise and why?
4. Why pre-experimental group is called single-group experiment? Describe three types of pre-experimental research.

5. Mention the tools for data collections in survey research.
6. Why comparative research is carried out very carefully? Describe with examples.
7. Why descriptive research is called 'Ex-post facto' research? Describe various methods of descriptive research.

Answer...

4.3 Characteristics

Data collection: This method follows structured research instruments for collecting data. Data gathering instruments contain items that seek measurable characteristics of the population (e.g. age, sex, the number of family members, income, educational status etc.). To avoid researcher's bias, this requires adherence to the principle of random sampling method to defeat the purpose of research.

Sample size: Quantitative method studies large number of population size, therefore it requires to choose larger sample sizes that are good are representative of the population.

Instruments for analysis: It follows standardized, pre-tested instruments for data collection and analysis of collected data which ensure the accuracy, reliability and validity of data. For more reliable data analysis, a normal population distribution curve is generally preferred over a non-normal distribution.

Research question: At the time of conducting this research the researcher has a clearly defined research question to which objective answers are sought.

Tools: Quantitative method uses various tools such as questionnaires, interviews, computer programmes, advanced digital or electronic instruments and so on to collect numerical data. Data are organized using tables, graphs, or figures, charts or other non-textual forms that combine large numbers of data to show relationships, trends, or differences among variables.

Scope to verify: Researchers can repeat the quantitative method to verify or confirm the findings in another setting. So the research study ensures its high reliability. This strengthens the validity of innovative discoveries or findings by eliminating the possibility of erroneous conclusions.

Outcomes: Quantitative models or formula derived from data analysis can predict outcomes. The outcomes of the research are used to generalize concepts more widely, predict future results, or investigate causal relationships.

Pre-caution: As quantitative research deals in numbers, logic, and an objective standpoint, all aspects of the study are carefully designed before conducting this study. The study focuses on numeric and unchanging data and detailed, convergent reasoning rather than divergent reasoning.

Self Check Exercise :

1. Why quantitative research method follows sampling instead of population?

Answer...

....

....

....

4.4 Advantages and Limitations

Quantitative methods presume to have an objective approach to studying research problems, where data is controlled and measured, to address the accumulation of facts, and to determine the causes of behavior. As a consequence, the results of quantitative research may be statistically significant but are often humanly insignificant. There are some advantages and specific limitations discussed in below.

4.4.1 Advantages

- This research provides estimates of populations at large.
- It is more reliable and objective and has precision, is definitive and standardized.
- The method indicates the extensiveness of attitudes held by people.
- Scope of use statistics to generalize a finding and provides results which can be condensed to statistics.
- There is a good scope to allow statistical comparison between various groups.
- It looks at relationships between variables and can establish cause and effect in highly controlled circumstances.
- This research often lessens and reorganizes a complex problem to a limited number of variables.
- Quantitative research has a scope to assume sample as the representative of the population.
- The method tests theories or hypotheses and it also measures level of occurrence, actions, trends, etc.
- In this research method subjectivity of researcher in methodology is recognized less.
- This method easily answers such questions as "How many?" and "How often?"

4.4.2 Limitations

Some specific limitations include:

- The method is good in data analysis and it is more efficient and able to test hypotheses, but may fail to spot contextual detail. So the result sometimes does not reflect real world situation.
- At the time of data collection through questionnaire, the development of standard questions by researchers can lead to "structural bias" and false representation, where the data actually reflects the researcher's view instead of the participating subject.
- This type of research introduces good results but provide less detail on behavior, attitudes, and motivation;
- As there it is flexible to data collect in terms of large or narrow data. Researcher may collect a much narrower and sometimes superficial dataset;
- Results are limited as they provide numerical descriptions rather than detailed narrative and generally provide less elaborate accounts of human perception;
- The research is often carried out in an unnatural, artificial environment so that a level of control can be applied to the exercise.
- Preset answers will not necessarily reflect how people really feel about a subject and, in some cases, might just be the closest match to the preconceived hypothesis.
- It uses a static and rigid approach and so employs an inflexible process of discovery.

Self Check Exercise :

1. Why quantitative research is statistically significant?
2. Why quantitative research never free from researcher's own bias?

Answer...
....
....
....
....
....
....
....
....

……
……
……
……
……
……
……

4.5 When do we use quantitative methods?

A lot of researchers, both quantitative and qualitative, take a pragmatist approach to research, using different methods depending on the research question they are trying to answer. In some cases, this will lead them to quantitative research, for example when they need to give a quantitative answer to a question or generalize findings to a population, or are looking to test a theory mathematically; in other cases, they will employ qualitative methods. Sometimes a mixed method approach combining quantitative and qualitative methods will be the most appropriate.

If a pragmatic approach is taken to research methods, first of all it is required to find out what kinds of questions are best answered using quantitative as opposed to qualitative methods. There are six main types of research questions that quantitative research is particularly suited to find an answer to:

- When the researcher demands quantitative answer from the study this type of research is best suited to conduct.
- Numerical change can likewise only accurately be studied using quantitative methods. For example, if a researcher wants to know the numbers of students in a university are rising or falling.
- This research is useful to conduct audience segmentation. It is done by dividing the population into groups whose members are similar to each other and distinct from other groups.
- To quantify opinions, attitudes and behaviours and find out how the whole population feels about a certain issue, this research method is best to the researcher. If a researcher wants to find out the exact number of people's attitudes regarding an issue prior to a campaign, researcher must choose this method.

- It is suitable to explain some phenomena. For example, to know the factors liable to change the student achievements in computer practice. This type of study can be successfully follows quantitative methods.
- For testing of hypotheses, quantitative research method is applied. Generally this type of research study wants to explain something, for example whether there is a relationship between students' achievement and their self-esteem and social background.

In above six research practices, first four types are descriptive research because they are trying to describe a situation. But later two types are called 'inferential research' because they explain something rather than just describe it.

Self Check Exercise :

1. In what situation a researcher selects quantitative method ?

Answer..

....

....

4.6 Quantitative Research - Approaches to Experimental Design

4.6.1 Introduction :

Experimental method is a scientific research method of quantitative study. It is a process of contribution to the already established knowledge. Thus, the researcher operates under the basic assumption that the research situation he wishes to evaluate has never existed and does not now exist. Situation here means in the sense of a programme, curriculum or method for organizing class, as well as a 'situation' created to test (Singh, 2006). Experimental research has been defined in various aspects. Sometimes it is called a method or process to test hypothesis or it is the method of verification of a hypothesis which seeks to make up two factors into a casual relationship through the study.

This method is often used where:

- there is time priority in a causal relationship;
- there is consistency in a causal relationship;
- the magnitude of the correlation is great.

4.6.2 Phases of Experimental Method

The overall structure for a quantitative design is based in the scientific method. It uses deductive reasoning, where the researcher forms a hypothesis, collects data in an investigation of the problem, and then uses the data from the investigation, after analysis is made and conclusions are shared, to prove the hypotheses not false or false. The basic procedure of a quantitative design is:

Problem Selection & Literature Review

↓

Formation of hypothesis

↓

Data collection

↓

Analysis

↓

Prove the hypothesis

↓

Conclusion

Fig. Deductive reasoning of scientific method

4.6.3 Variables:

Experimental research is usually used in sciences such as sociology and psychology, physics, chemistry, biology and medicine etc. It is a collection of research designs which use manipulation and controlled testing to understand causal processes. Generally, one or more variables are manipulated to determine their effect on a dependent variable. The experimental method is a systematic and scientific approach to research in which the researcher manipulates one or more variables, and controls and measures any change in other variables.

A variable is an element, entity, factor or measurable characteristic that varies. It may change from group to group, person to person, or even within one person over time. In other words, variables are the conditions or the characteristics the researcher manipulates, controls, or observes in order to obtain the results of an experiment.

Various types of variables are listed below.

(1) **Dependent and Independent Variables:**

An independent variable (sometimes called an explanatory, predictor, stimulus, covariate or control), is a variable that is being manipulated in an experiment in order to observe the effect on a dependent variable (sometimes called explained, outcome, response or controlled). So a dependent variable is affected by the independent variable i.e. it responds to the independent variable. For example, positive change in workers' salary increases factory production. Here salary is an independent variable and production is dependent variable.

Independent variable is categorized into two types:

(i) **Treatment variable:** The variables that the experimenter manipulates and to which he/she assign subject.

(ii) **Organismic variables:** These variables cannot be changed by researchers, as they include variables being studied. It is mainly personal characteristics of participants' gender, age, weight in behavioral sciences. These variables cannot be altered by the researchers.

(2) **Quantitative and Qualitative Variable:**

Quantitative variables are expressed in numerical terms but qualitative variable cannot be expressed in numerical figures. Qualitative variables as the abstract attributes like talent, academic achievement and creativity cannot be observed directly. But they can be given operational definition that is, precise meaning and quantifiable property.

(3) **Confounding variables:**

These variables are those aspects of a study or sample that might influence the dependent variable and whose effect may be confused with the effect of the independent variable. Confounding variables are two types:

(i) **Extraneous variables:** These are environmental elements that mystify the relationships among variables under study. It affects the dependent variables. Sometimes it causes the invalidation of the true results.

(ii) **Intervening variables:** Intervening variables include variables that are not directly observable e.g., nervousness, dullness. It is caused by the independent variable and is itself a cause of the dependent variable.

Extraneous variables can be controlled and the effects of those variables in the research study can be reduced by the researchers using the method of randomization. At the time of collection of variables researchers must select random process as well as random assignment of participants to treatment groups. To control extraneous variables, researchers hold certain variables constant throughout an experiment for both the control and treatment group. Some of the other methods are matching groups, comparing subgroups or homogenous groups, using a single group as both the control and treatment group etc. Moreover, researchers may follow other practices to minimize the effects of those variables.

4.6.4 Steps for Conducting Experimental Research

Various steps are followed in completing experimental research. These are given below:

(i) **Selecting a research problem:** Topics are typically selected based on personal interest, a review of recent literature, and the experiences of the researcher. The problems willing to testing generally can, and should, be converted into a hypothesis that can be verified.

(ii) **Literature reviews:** In this step it is needed to review the literature and specify a research question. Research questions are typically generated based on researchers' exploration of previously published literature of the topic.

(iii) **Hypothesis:** In this step a clear hypothesis is formulated. Three types of research hypotheses exist in experimental research – directional hypothesis, non-directional hypothesis and the null hypothesis.

(iv) **Defining the population:** This step includes select and assign participants to groups. After defining the population researchers randomly select and assign participants in to two groups: a control group and a treatment group.

(v) **Selection of an Instrument:** Researchers select or design instruments to use in experimental research based on which instrument will best measure the dependent variable.

(vi) **Carrying out the experiment:** It is necessary here to insist on close adherence to plans, especially as they relate to the factors of control, randomization and replication.

(vii) **Collection and analysis:** This step collects observational data and analyzes those collected data.

(viii) **Decision about the Hypothesis:** Researchers collect substantial evidences to support rejecting the null hypothesis. Null hypothesis nullify the relationships between the treatment and control group i.e. there is no difference between them. Researchers generally provide evidence to the contrary; that the two groups differed as a result of the treatment applied.

(ix) **Draw conclusions:** After statistical analyses of the data it is important to researchers to formulate the conclusions regarding the implications of their study. The conclusions of the study must be restricted to the population actually investigated.

Self Check Exercise:

1. What do you mean by experiment? Why experimental research is called 'deductive reasoning'?
2. Describe various types of variables in experimental research.

Answer...

....

....

....

....

....

....

4.7 Basic Research Design for Quantitative Studies

The overall structure for a quantitative design is based on the scientific method. It uses deductive reasoning, where the researcher frames a hypothesis, collects data, and then uses the data for analysis, after analysis is made and conclusions are shared, to prove the hypotheses not false or false. Before designing a quantitative research study, the researcher must decide whether the study will be descriptive or experimental. On the basis of the type, data gathering, analysis and interpretation will differ. Though there are different steps of research and it varies from one to one type, the basic research design is same.

4.7.1 Introduction

This part of the research is written in present tense and from the third person point of view. It includes the followings:

- Identifies the research problem
- Review of the literature
- Formulation of hypothesis
- Describes the various aspects such as define unfamiliar or complex terms, concepts, or ideas etc.

4.7.2 Methodology

Quantitative study should describe how each objective of the study will be achieved. It requires providing enough details of methodology to enable the research outcomes fruitful. The followings aspects are essentials in this section.

- Details of the study population and sampling and procedures used for their selection
- Data collection method, tools, instruments, limitations and sufficient notes regarding data collection
- Procedures for processing and analyzing the data, mathematical techniques, computer programmes etc.

4.7.3 Results

The finding of the research study should be written objectively and in a precise format. In quantitative studies, it is common to use graphs, tables, charts, and other non-textual elements to help the reader understand the data. For analyzing the data or statistical analysis, all the methods and statistical formulas should be written carefully and should be present in a logical, sequential order. Detail descriptions of the interpretations and results should be presented here.

4.7.4 Discussion

Discussions should be analytic, logical, and comprehensive. It should meld together all findings in relation to those recognized in the literature review, and located within the context of the theoretical framework of the study. The discussion should be written in the present tense. It includes the followings:

- Description of trends, comparison of groups, or relationships among variables
- Description any trends that emerged from the analysis
- Explanation of all unanticipated and statistical insignificant findings
- Highlight key findings based on the overall results and note findings that are important
- Depict any limitations or unavoidable bias in the study

4.7.5 Conclusion

At the end of the study it must be summarized and provide a final comment and assessment of the study. Basic three areas of the conclusion are:

- Summary of findings – Synthesize the answers to the research questions
- Recommendations – Key findings with recommendations or actions to be taken in practice
- Future research – Need for future research which is linked to the present study

Self Check Exercise :

1. Why quantitative research follows scientific method? Describe the steps of quantitative research design.

Answer...

....

....

....

....

....

....

....

....

....

....

....

....

4.8 Summary

In this unit the students studied that quantitative research is a type of empirical study and this research can be exciting and highly informative. That means the research focuses on verifiable observation as opposed to theory or logic. This research is about explaining phenomena by collecting quantitative data which are analyzed using mathematically based methods. Most often this type of research is expressed in numbers. A researcher represents and manipulates certain observations that they are studying. They attempts to explain what it is they are seeing and what affect it has on the subject. They also determine and what the changes may reflect. It can be used to help explain all sorts of phenomena.

Though, quantitative research is categorized into many types, only four types of quantitative research designs: descriptive, correlational, quasi-experimental and experimental are most popular. The best quantitative research gathers precise empirical data and can be applied to gain a better understanding of several fields of study. The overall goal is to convey numerically what is being seen in the research and to arrive at specific and observable conclusions.

4.9 Keywords

Bias: A loss of balance and accuracy in the use of research methods. It can appear in research via the sampling frame, random sampling, or non-response. It can also occur at other stages in research, such as while interviewing, in the design of questions, or in the way data are analyzed and presented.

Case Study: The collection and presentation of detailed information about a particular participant or small group, frequently including data derived from the subjects themselves.

Case Study: The collection and presentation of detailed information about a particular participant or small group, frequently including the accounts of subjects themselves.

Cohort Analysis: It is group by group analytic treatment of individuals having a statistical factor in common to each group. Group members share a particular characteristic.

Control Group: A group in an experiment that receives not treatment in order to compare the treated group against a norm.

Control Group: The group in an experimental design that receives either no treatment or a different treatment from the experimental group. This group can thus be compared to the experimental group.

Controlled Experiment: It is an experimental design with two or more randomly selected groups in which the researcher controls or introduces the independent variable and measures the dependent variable at least two times.

Data: Recorded observations, usually in numeric or textual form.

Dependent Variable: A variable that receives stimulus and measured for the effect the treatment has had upon it.

Discrete Variable: A variable that is measured solely in whole units, e.g., gender and siblings.

Empirical Research: The process of developing systematized knowledge gained from observations that are formulated to support insights and generalizations about the phenomena being researched.

Experimental Research: A researcher working within this methodology creates an environment in which to observe and interpret the results of a research question. A key element in experimental research is that participants in a study are randomly assigned to groups.

External validity: Concerned with the extent to which study findings can be generalized beyond the sample used in the study

Hypothesis: A tentative explanation based on theory to predict a causal relationship between variables.

Independent variable: An independent variable is a variable that is manipulated to determine the value of a dependent variable.

Internal validity: The extent to which the effects detected in the study are a true reflection of reality rather than the result of extraneous variables

Methodology: a theory or analysis of how research does and should proceed.

Null Hypothesis: The proposition, to be tested statistically, that the experimental intervention has "no effect," meaning that the treatment and control groups will not differ as a result of the intervention. Investigators usually hope that the data will demonstrate some effect from the intervention, thus allowing the investigator to reject the null hypothesis.

Population: The target group under investigation is called population. The population is the entire set under consideration. Samples are drawn from populations.

Quantitative Research: Empirical research in which the researcher explores relationships using numeric data. Survey is generally considered a form of quantitative research.

Quasi-experiment: Similar to true experiments. Have subjects, treatment, etc., but uses nonrandomized groups.

Representative Sample: Sample in which the participants closely match the characteristics of the population, and thus, all segments of the population are represented in the sample. A representative sample allows results to be generalized from the sample to the population.

Research design: It is a blueprint for conducting a study. It maximizes control over factors that could interfere with the validity of the findings and guides the planning and implementation of a study in a way that is most likely to achieve the intended goal.

Sample: A sample is a subset containing the characteristics of a larger population. Samples are used in statistical testing when population sizes are too large for the test to include all possible members or observations. A sample should represent the whole population and not reflect bias toward a specific attribute.

Treatment: The stimulus given to a dependent variable.

Validity: The degree to which a study accurately reflects or assesses the specific concept that the researcher is attempting to measure. A method can be reliable, consistently measuring the same thing, but not valid.

Variable: Observable characteristics that vary among individuals.

4.10 References and Further Readings

AECT. (2001). What is descriptive research? Retrieved Jan 7, 2017 from http://www.aect.org/edtech/ed1/41/41-01.html

Characteristics of experimental research. (n.d.). Retrieved Jan 5, 2017 from https://www.aea267.k12.ia.us/?ACT=26&fid=5&d=1468&f=experimental_research_9.22.15.pdf

Creswell, J. (2008). Educational research: Planning, conducting, and evaluating quantitative and qualitative research. New Jersey: Pearson: Merrill Prentice Hall.

Creswell, J. W. (2014). Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.). London: Sage.

Dawson, C. (2002). Practical research methods: A user-friendly guide to mastering research. Newtec Place, United Kingdom: How to Books.

Edades, J. M. M. (n.d.). Characteristics of descriptive research. Retrieved Jan 10, 2017 from https://www.scribd.com/doc/84273242/Characteristics-of-Descriptive-Research

Gilbert, N. (1995). Social research update. United Kingdom: Department of Sociology, University of Surrey. Retrieved Jan 10, 2017 from http://sru.soc.surrey.ac.uk/SRU4.html

Institute of Work and Health. (2009). Retrieved Feb 5, 2017 from https://www.iwh.on.ca/wrmb/cross-sectional-vs-longitudinal-studies

Jakobsen, T. G. (n.d.). Rules of comparative research. Retrieved Jan 11, 2017 from http://www.popularsocialscience.com/2013/01/28/some-rules-of-comparative-research/

Jowell, R. (1998). How comparative is comparative research?. The American Behavioral Scientist, 42(2).

Key elements of a research proposal: Quantitative design. (n.d.). Retrieved Jan 10, 2017 from https://www.bcps.org/offices/lis/researchcourse/develop_quantitative.html

Klazeña, A. (June 9, 2014). Types of quantitative research for students and researchers. Retrieved Jan 10, 2017 from https://blog.udemy.com/types-of-quantitative-research/

Kothari, C. R. (2004). Research methodology: Methods and techniques (2nd rev.ed.). New Delhi: New Age International Publishers.

McCusker, K & Gunaydin, S. (2015). Research using qualitative, quantitative or mixed methods and choice based on the research. Perfusion, 30(7), 537-542.

Mills, M., Bunt, G. G. & Bruijin, J. (2006). Comparative research: persistent problems and promising solutions. International Sociology, 21(5), 619-631.

Murtonen, M. (2005). Learning of quantitative research methods - University students' view, motivation, and difficulties in learning. Finland: University of Turku.

National Council for Osteopathic Research. (2014). Quantitative research methods. Retrieved Jan 7, 2017 from http://www.ncor.org.uk/wp-content/uploads/2014/ 03/Quantitative_research_methods.pdf

Ohdedar, A. K. (1993). Research methodology. Kolkata: Bengal Library Association.

Organizing Your Social Sciences Research Paper: Glossary of Research Terms. (n.d.). Retrieved Jan 10, 2017 from http://libguides.usc.edu/writingguide/researchglossary

Oskar, B. (Jul 10, 2008). Experimental research. Retrieved Jan 18, 2017 from https://explorable.com/experimental-research

Patricka, R. (Jan 3, 2015). Quantitative methods: Meaning and characteristics. Retrieved

Jan 10, 2017 from http://simplyeducate.me/2015/01/03/quantitative-methods-meaning-and-characteristics/

Polit, D. F., & Hungler, B. P. (1999). Nursing research: Principles and methods (6th ed.). Philadelphia: Lippincott.

PPA 696 Research methods. (n.d.). Retrieved Jan 10, 2017 from http://web.csulb.edu/~msaintg/ppa696/696preex.htm

Pre-experimental design (n.d.). Retrieved Jan 10, 2017 from http://sociologyindex.com/pre_experimental_design.htm

Ragin, C. (1994). Constructing social research: The unity and diversity of method. New Delhi: Pine Forge Press. Retrieved Jan 18, 2017 from http://poli.haifa.ac.il/~levi/res/mgsr4.htm

Scofield, T. L. (2012). Characteristics of a good experimental design. Retrieved Jan 6, 2017 from https://www.calvin.edu/~scofield/courses/m145/materials/overheads/goodExperimentalDesign.shtml

Sekaran, U. (2003). Research methods for business: A skill-building approach (4th ed.). New York: John Wiley & Sons.

Singh, Y. K. (2006). Fundamentals of research methodology and statistics. New Delhi: New Age International Publishers.

Sukamolson, S. (n.d.). Fundamentals of quantitative research. Retrieved Jan 7, 2017 from http://www.culi.chula.ac.th/Research/e-Journal/bod/Suphat%20Sukamolson.pdf

Suter, W. (2006). Introduction to educational research: A critical thinking approach. Thousand Oaks, CA: Sage Publications.

https://www.wikipedia.org/

Unit 5 ❑ Research Methods : Qualitative

Structure

5.0 Objective

After reading this Unit, you will be able to:

- understand the concept of qualitative research methods and difference from quantitative research methods;
- comprehend its characteristics and features;
- know the types of qualitative research design; and
- understand the triangulation and mixed methods research
- Comprehend a typical qualitative research approach
- Understand methods of collecting qualitative data and Qualitative Data Analysis
- Understand the Relevance of Qualitative Research to LIS

5.1 Introduction

Qualitative research has its roots in logical, rational and philosophical analysis. The tradition of qualitative research is as old as philosophy itself. Aristotle, Socrates, Weber, Marx, Durkhiem and Giddens are some of the many thinkers who have shaped the philosophical foundations of qualitative research. Though it is often described as an unscientific research approach, it has led to many scientific achievements. In the last century, the emergence of functionalists led to a great deal of criticism of qualitative research, and that has tended to create two distinct schools of thought about the conduct of research: one supporting quantitative research, the other supporting qualitative research. LIS is an emerging discipline, and therefore still is in the process of building a strong research foundation for itself. Library and information science is becoming increasingly important, simply because of the ever increasing informatization of society. The advent of information technology (IT) has changed the global horizons forever and created an enormous change in the way in which people acquire information and knowledge. Though libraries are changing with these times, there will be tremendous pressure on librarians and information scientists to cater to the needs of people and organizations in a most effective and efficient way. To do that takes a better understanding of many information phenomena not currently well understood, and qualitative research can play an important role in furthering that understanding. Social scientist Fidel said that qualitative research is "non controlling, holistic and case oriented, about processes, open and flexible, diverse in methods, humanistic, inductive and scientific". A review of various qualitative methods would help both researchers and practitioners to develop a broader understanding of the usefulness of these methods in Qualitative Research: A Broad Area of Inquiry The diversity of what is called qualitative research, because of its relevance to different disciplines and professions, challenges anyone to arrive at a succinct definition. Too brief a definition will seem to exclude one discipline or another. Too broad a definition will seem uselessly global. In fact, the term *qualitative research* may be like other terms of the same genre—for example, *sociological research, psychological research*, or *education research*. Within its own particular discipline or profession, each term connotes a large body of research, embracing a variety of highly contrasting methods. Think simply, for instance, of clinical and experimental psychology. Both form vigorous parts of the same field, though the methods differ markedly.

A starting point in trying to understand the collection of information for research purposes is that there are broadly two approaches: quantitative research and qualitative research. Early forms of research originated in the natural sciences such as biology, chemistry, physics, geology etc. and was concerned with investigating things which we could observe and measure in some way. Such observations and measurements can be made objectively and repeated by other researchers. This process is referred to as "quantitative" research.

Much later, along came researchers working in the social sciences: psychology, sociology, anthropology etc. They were interested in studying human behaviour and the social world inhabited by human beings. They found increasing difficulty in trying to explain human behaviour in simply measurable terms. Measurements tell us how often or how many people behave in a certain way but they do not adequately answer the question "why?". Research which attempts to increase our understanding of why things are the way they are in our social world and why people act the ways they do is "qualitative" research.

5.2 Qualitative Research

5.2.1 The word research is derived from the Middle French "recherche", which means "to go about seeking".

Qualitative research is a method of inquiry employed in many different academic disciplines, including in the social sciences and natural sciences, but also in non-academic contexts including market research, business, and service demonstrations by non-profits.[1]

Qualitative Research: Development of concepts which help us to understand social phenomena in natural (rather than experimental) settings, giving due emphasis to the meanings, experiences and views of the participants.

Qualitative Research describe or answer questions about particular, localized occurrences or contexts and the perspectives of a participant group toward events, beliefs, or practices.

Qualitative research is concerned with developing explanations of social phenomena. That is to say, it aims to help us to understand the world in which we live and why things are the way they are. It is concerned with the social aspects of our world and seeks to answer questions about:

- Why people behave the way they do
- How opinions and attitudes are formed
- How people are affected by the events that go on around them
- How and why cultures have developed in the way they have
- The differences between social groups

5.2.2 Qualitative Research VS Quantitative Research

5.2.2.1 Comparison of qualitative and quantitative research terms.

Qualitative Research	Quantitative Research
Subjective	Objective
Holistic	Reductionist
Phenomenological	Scientific
Anti positivist	Positivist
Descriptive	Experimental
Naturalistic	Contrived
Inductive	Deductive

5.2.2.2 Differences between Qualitative and Quantitative research

Qualitative	Quantitative
• The aim is a detailed description. • Researcher may only know roughly in advance what he/she is looking for. • The design emerges as the study unfolds. • Researcher is the data gathering instrument. • Data is in the form of words, pictures or objects. • Subjective - individuals' interpretation of events is important • Qualitative data is more 'rich', time consuming, and not generalizable. • Researcher tends to become subjectively immersed in the subject matter.	• The aim is to classify features, count them, and construct statistical models in an attempt to explain what is observed. • Researcher knows clearly in advance what he/she is looking for. • All aspects of the study are carefully designed before data is collected. • Researcher questionnaires or equipment to collect numerical data. • Data is numerical in nature. • Objective – seeks measurement & analysis of target concepts. • Quantitative data is more efficient, able to test hypotheses. • Researcher tends to remain separated from the subject matter.

5.3 Characteristics of Qualitative Research

- Data sources are real-world situations
- Appropriateness of methods and theories
- Variety of approaches and methods in qualitative research
- Data are descriptive
- Emphasizes a holistic approach (processes and outcomes)
- Data analysis is inductive
- Reflexivity of the researchers and the research

- Perspectives of the participants and their diversity
- Describes the meaning of research findings from the perspective of the research participants
- Reconstructing cases as starting point
- Construction of reality as basis
- Text as empirical material

5.4 Five Features of Qualitative Research

Instead of trying to arrive at a singular definition of qualitative research, you might consider five features:

1. Studying the meaning of people's lives, under real-world conditions;
2. Representing the views and perspectives of the people in a study;
3. Covering the contextual conditions within which people live;
4. Contributing insights into existing or emerging concepts that may help to *explain* human social behaviour; and
5. Striving to use *multiple sources of evidence* rather than relying on a single source alone.

Starting at the top of the list, qualitative research **first** involves studying the meaning of people's lives, under real-world conditions. People will be performing in their everyday roles or have expressed themselves through their own diaries, journals, writing, and even photography—entirely independent of any research inquiry. Social interactions will occur with minimal intrusion by artificial research procedures, and people will be saying what they want to say, not, for example, limited to responding to a researcher's pre established questionnaire. Likewise, people will not be inhibited by the confines of a laboratory or any laboratory-like setting. And they will not be represented by such statistical averages as the average American family having 3.18 persons (as of 2006)—which at once may represent accurately an entire population but in fact by definition does not speak to any single, real-life family.

Second, qualitative research differs because of its ability to represent the views and perspectives of the participants in a study. Capturing their perspectives may be a major purpose of a qualitative study. Thus, the events and ideas emerging from qualitative research can represent the meanings given to real-life events by the people who live them, not the values, preconceptions, or meanings held by researchers.

Third, qualitative research covers contextual conditions—the social, institutional, and environmental conditions within which people's lives take place. In many ways, these contextual conditions may strongly influence all human events. However, the other social science methods (except for history) have difficulty in addressing these conditions. Experiments, for instance, "control out" these conditions (hence the artificiality of laboratory experiments). Quasi-experiments admit such conditions but by design nevertheless focus only on a limited set of "variables," which may or may not fully appreciate the contextual conditions. Similarly, surveys are constrained by the need to manage carefully the degrees of freedom required to analyze the responses to a set of survey questions; surveys are therefore limited in the number of questions devoted to any contextual conditions. History does address contextual conditions, but in its conventional form studies the "dead past," not ongoing events as in qualitative research.

Fourth, qualitative research is not just a diary or chronicle of everyday life. Such a function would be a rather mundane version of real-world events. On the contrary, qualitative research is driven by a desire to explain these events, through existing or emerging concepts. For instance, one existing concept is Goffman's (1963) stigma management. In his original work, stigma management largely pertained to adaptations by individual people. However, a contemporary qualitative study applied his typology and framework to a collective group, thereby offering new insights into how the actions of nation-states also might try to overcome their own historically stigmatizing events. Similarly, qualitative research can be the occasion for developing new concepts. The concepts might attempt to explain social processes, such as the schooling of American students. An illustrative concept offered by a qualitative study is the notion of *subtractive schooling*, used to provide potentially useful explanations and to form a platform for new inquiries. In fact, studies devoid of concepts, whether existing or new, or devoid of any interpretations at all, would resemble diaries or chronicles but not qualitative research.

Fifth, qualitative research strives to collect, integrate, and present data from a variety of sources of evidence as part of any given study. The variety will likely follow from your having to study a real-world setting and its participants. The complexity of the field setting and the diversity of its participants are likely to warrant the use of interviews and observations and even the inspection of documents and artifacts.

5.5 Qualitative research designs

Six major types of qualitative research design are outlined. They are:

1) phenomenology
2) ethnography
3) grounded theory
4) case study
5) design-based Research
6) action Research

Another common research design is the survey. Surveys can be either qualitative or quantitative in their approach to data collection. A description of qualitative surveys can be found in another Trent Focus resource pack.

5.5.1 Phenomenology

The terminology used by different authors can be very confusing and the use of the term *phenomenology* is one example. Phenomenology was one of the terms used to describe qualitative research generally. However, it is also used Moustakas: "to determine what an experience means for the persons who have had the experience and are able to provide a comprehensive description of it. From the individual descriptions, general or universal meanings are derived, in other words, the essences of structures of the experience."

Phenomenology literally means the study of phenomena. It is a way of describing something that exists as part of the world in which we live. Phenomena may be events, situations, experiences or concepts. We are surrounded by many phenomena, which we are aware of but not fully understand. Our lack of understanding of these phenomena may exist because the phenomenon has not been overtly described and explained or our understanding of the impact it makes may be unclear. Back pain is another example. Correlation studies may tell us about the types of people who experience back pain and the apparent causes. Randomised controlled trials of drugs compare the effectiveness of one analgesia against another. But what is it actually like to live with back pain? What are the effects on peoples' lives? What problems does it cause? A phenomenological study might explore, for example, the effect that back pain has on sufferers' relationships with other people by describing the strain it can cause in marriages or the effect on children of having a disabled parent. Phenomenological research begins with the acknowledgement that there is a gap in our understanding and that clarification or illumination will be of benefit. Phenomenological research will not necessarily provide definitive explanations but it does raise awareness and increases insight.

5.5.2 Ethnography

Ethnography has a background in anthropology. The term means "portrait of a people" and it is a methodology for descriptive studies of cultures and peoples. The cultural parameter is that the people under investigation have something in common. Examples of parameters include:

geographical - a particular region or country
religious
tribal
shared experience

In health care settings, researchers may choose an ethnographic approach because the cultural parameter is suspected of affecting the population's response to care or treatment. For example, cultural rules about contact between males and females may contribute to reluctance of women from an Asian subgroup to take up cervical screening. Ethnography helps health care professionals to develop cultural awareness and sensitivity and enhances the provision and quality of care for people from all cultures.

Ethnographic studies entail extensive fieldwork by the researcher. Data collection techniques include both formal and informal interviewing, often interviewing individuals on several occasions, and participant observation. Because of this, ethnography is extremely time consuming as it involves the researcher spending long periods of time in the field.

Analysis of data adopts an "emic" approach. This means that the researcher attempts to interpret data from the perspective of the population under study. The results are expressed as though they were being expressed by the subjects themselves, often using local language and terminology to describe phenomena. For example, a researcher may explore behaviour which we traditionally in the westernised medical world would describe as mental illness. However, within the population under study, the behaviour may not be characterised as illness but as something else - as evidence that the individual is "blessed" or "gifted" in some way.

Ethnographic research can be problematic when researchers are not sufficiently familiar with the social mores of the people being studied or with their language. Interpretation from an "etic" perspective - an outsider perspective - may be a misinterpretation causing confusion. For this reason, the ethnographic researcher usually returns to the field to check his interpretations with informants thereby validating the data before presenting the findings.

5.5.3 Grounded theory

This methodology originated with Glaser and Strauss and their work on the interactions between health care professionals and dying patients. The main feature is the development of new theory through the collection and analysis of data about a phenomenon. It goes beyond phenomenology because the explanations that emerge are genuinely *new* knowledge and are used to develop new theories about a phenomenon. In health care settings, the new theories can be applied enabling us to approach existing problems in a new way. For example, our approaches to health promotion or the provision of care. One example of grounded theory with which many of us are familiar is theory about the grief process. Researchers observed that people who were bereaved progressed through a series of stages and that each stage was characterised by certain responses: denial, anger, acceptance and resolution. This is not a new phenomenon, people have going through these stages for as long as society has existed, but the research formally acknowledged and described the experience. Now we use our knowledge of *the grief process*, new knowledge derived from grounded theory, to understand the experience of bereavement and to help the bereaved to come to terms with their loss. We recognise when a person is having difficulty coming to terms with loss because we use the knowledge to recognise signs of "abnormal" grief and can offer help.

Various data collection techniques are used to develop grounded theory, particularly interviews and observation although literature review and relevant documentary analysis make important contributions. A key feature of grounded theory is the simultaneous collection and analysis of data using a process known as constant comparative analysis. In this process, data are transcribed and examined for content immediately following data collection. Ideas which emerge from the analysis are included in data collection when the researcher next enters the field. For this reason, a researcher collecting data through semi structured interviews may gradually develop an interview schedule in the latter stages of a research project which looks very different to the original schedule used in the first interview.

New theory begins its conception as the researcher recognises new ideas and themes emerging from what people have said or from events which have been observed.

Memos form in the researcher's consciousness as raw data is reviewed. Hypotheses about the relationship between various ideas or categories are tested out and constructs formed leading to new concepts or understandings. In this sense the theory is "grounded" in the data.

As in phenomenology where there are concepts of which we are aware but do not fully understand, there are aspects of health care which might be informed by the development of new theory. One example is spirituality. In any holistic programme of care health care professionals may talk about the need to meet the "spiritual needs" of patients. However, we understand very little of what this means. At first sight, spiritual needs might be interpreted as referring to religious beliefs but many people would say that spiritual needs are more than this. It may be an individual's sense of well being, happiness or peace of mind. Grounded theory research could provide health care professionals with a better framework for providing truly holistic care.

5.5.4 Case study

Like surveys, case study research is one of those research approaches which can take a qualitative or quantitative stance. In this resource pack, the qualitative approach to case study is described wherein the value of case study relates to the in depth analysis of a single or small number of units. Case study research is used to describe an entity that forms a single unit such as a person, an organisation or an institution. Some research studies describe a series of cases.

Case study research ranges in complexity. The most simple is an illustrative description of a single event or occurrence. More complex is the analysis of a social situation over a period of time. The most complex is the extended case study which traces events involving the same actors over a period of time enabling the analysis to reflect changes and adjustments. As a research design, the case study claims to offer a richness and depth of information not usually offered by other methods. By attempting to capture as many variables as possible, case studies can identify how a complex set of circumstances come together to produce a particular manifestation. It is a highly versatile research method and employs any and all methods of data collection from testing to interviewing.

Case study research in health care has a range of uses. For example, a case study may be conducted of the development of a new service such as a hospital discharge liaison scheme jointly run by health and social services in one locality. Another example of the case study approach would be to describe and analyse organisational change in the planning, purchasing or delivery of health services as in Total Purchasing pilot projects. One of the most common uses of the case study is the evaluation of a particular care approach. For example, an outreach teenage health service set up as an alternative to general practice based teenage clinics might be evaluated in terms of input, impact on the health of teenagers locally and the development of collaborative links with other groups involved in promoting teenage health.

One of the criticisms aimed at case study research is that the case under study is not necessarily representative of similar cases and therefore the results of the research are not generalisable. This is a misunderstanding of the purpose of case study research which is to describe *that particular case* in detail. It is particularistic and contextual. For example, the usefulness of an outreach teenage health service would be determined by a number of local factors and an evaluation of the service would take those factors into account. If the service works well it does not automatically mean that the service would work equally well in another part of the country but the lack of generalisability does not lessen the value of the service in the area where it is based. Generalisability is not normally as issue for the researcher who is involved in studying a specific situation. It is an issue for the readers who want to know whether the findings can be applied elsewhere. It is the readers who must decide whether or not the case being described is sufficiently representative or similar to their own local situation.

5.5.5 Action Research

Action-Research is today one of the more promising qualitative research approaches in Information Sciences and Technologies research.

In essence, it consists of repeatedly going through the cycle:

Planning => Action => Reflection

We start by making a plan of our action in a crude first approximation, we act following that plan, and we then reflect on the results obtained. From this reflection, we correct our previous plan, act in agreement with the new plan, and reflect on the results we have now obtained. The cycles go on, repeatedly, until we are happy with the results Action-research corresponds to what John Dewey called the *Principle of Intelligent Action*.

5.5.6 Design-based Research

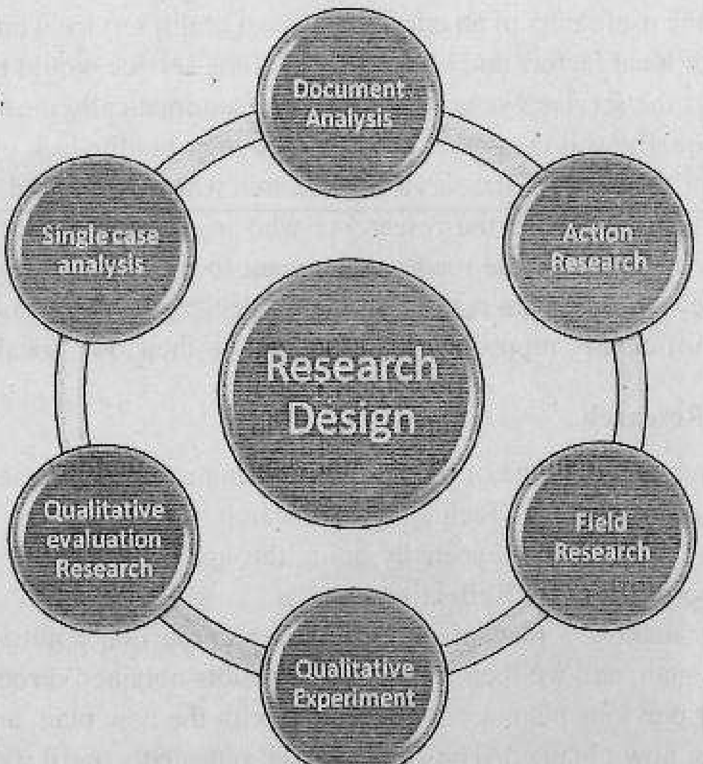
Design-based Research is a research method where knowledge is built in successive approximations while designing, building, and evaluating an artifact. The artifact may be almost anything: a piece of equipment, a software application, the solution to a social or technical problem, a theoretical framework, or even a whole theory.

In essence, it consists of repeatedly going through the cycle:

Awareness of Problem => Suggestion => Development => Evaluation => Conclusion

The successive improvements introduced in the artifact as it is put to test represent opportunities for consolidating the knowledge that emerges from its design and application.

Common Qualitative Research Design



5.6 Triangulation and Mixed Methods Research

Mixed Methods Research (also called, by some authors, Multiple Method Research or Multimethod Research) is becoming the third major research approach, along with quantitative research and qualitative research. Janice Morse (2003) provides the following distinctions: Mixed methods design - the incorporation of various qualitative and quantitative strategies within a single project, that may have either a qualitative or quantitative theoretical drive. Multimethod design - the use of two or more research methods, each conducted rigorously and complete in itself, in one project.

The results are then triangulated to form a complete whole. For multimethod designs Morse (2003) defines three principles:

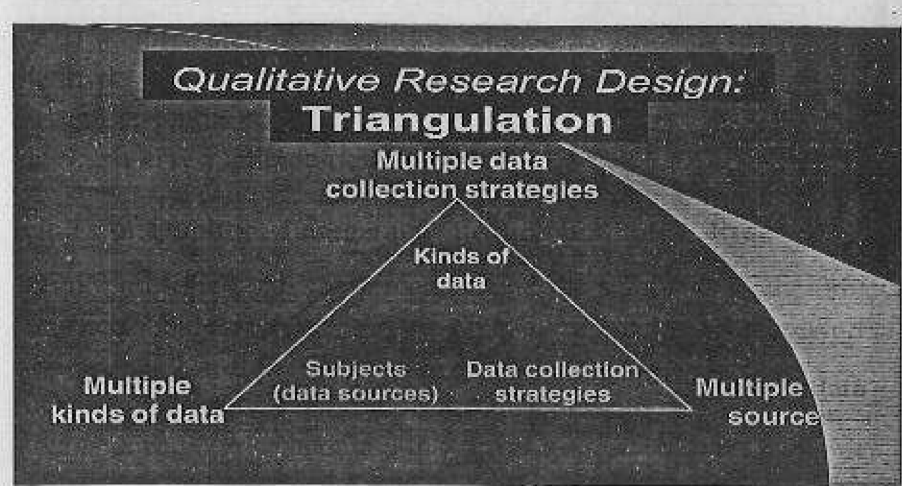
Principle 1: identify the theoretical drive (inductive or deductive) of the project.

Principle 2: develop overt awareness of the dominance (QUAN or QUAL, and simultaneous or sequential) of each project.

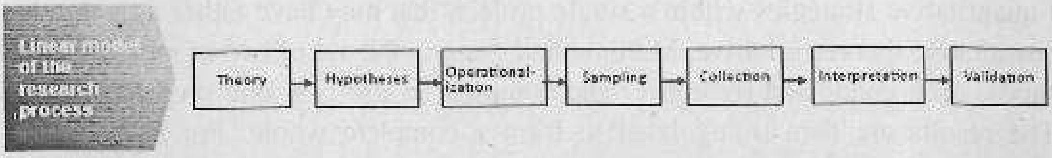
Principle 3: observe methodological integrity.

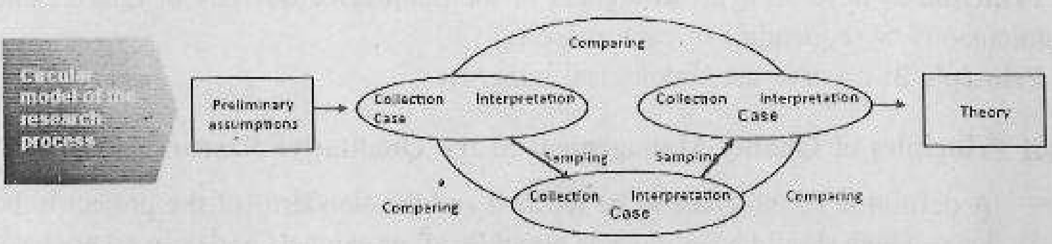
5.6.1 Principles of Quality Management in the Qualitative Research Process

- A definition of the goals to be reached and the standards of the project to be kept, which should be as clear as possible: all researchers and co-workers have to be integrated in this definition.
- A definition, how these goals and standards, and more generally the quality to be obtained, can be reached: and therefore, a consensus about the way how to apply certain methods, perhaps through joint interview training, and its analysis are preconditions for quality in the research process.
- A clear definition of the responsibilities for obtaining quality in the research process.
- The transparency of the judgement and the assessment and quality in the process



5.6.2 Models of Process and Theory

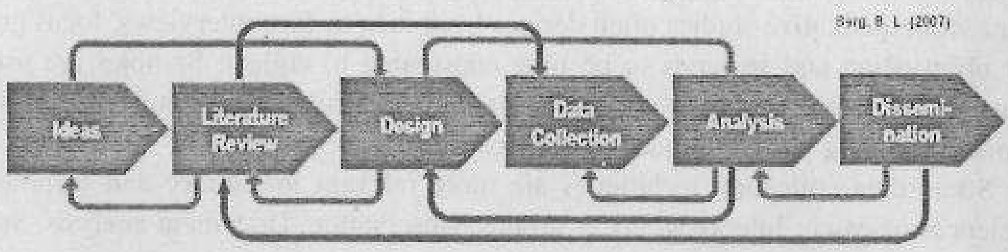




5.6.3 Research Questions in the Research Process

1. Formulation of the overall question
2. Formulation of specific research
3. Formulation of sensitizing concepts
4. Selection of research groups with which to study the question
5. Selection of appropriate designs and methods
6. Evaluation and reformulation of the specific research questions
7. Collection of data
8. Evaluation and reformulation of the specific research questions
9. Analyzing the data
10. Generalization and evaluation of the analyses
11. Formuiation of the findings

5.7 A Typical Qualitative Research Approach



- Research begins with rough *Ideas* that turn slowly into *research questions.*
- The *Literature Review* gradually builds up knowledge about the research ideas.
- The Design is the plan to be followed in order to carry out the research. It must include strategies for the selection of the *samples.*
- The *Data Collection and Organization* phase is devoted to gathering the data for the research and organizing it, so that it can be properly analyzed. These are difficult tasks, since the volume of data collected in qualitative research can be enormous.
- The Analysis includes three concurrent flows of action: *Data Reduction*, which focuses, simplifies, and transforms raw data into more manageable forms; Data Display, which presents the data as organized and compressed assemblies of information that permit conclusions to be analytically drawn; and Conclusions & Verification, where the researchers review and finalize all their conclusions and make sure that they satisfy the requirements of validity.
- Dissemination takes the form of very well written and detailed documents, so that other researchers can evaluate the analysis and conclusions obtained and decide if they trust the results and want to use them to feed their own research.

5.8 Methods of collecting qualitative data

Qualitative approaches to data collection usually involve direct interaction with individuals on a one to one basis or in a group setting. Data collection methods are time consuming and consequently data is collected from smaller numbers of people than would usually be the case in quantitative approaches such as the questionnaire survey. The benefits of using these approaches include richness of data and deeper insight into the phenomena under study.

Unlike quantitative data, raw qualitative data cannot be analysed statistically. The data from qualitative studies often derives from face-to-face interviews, focus groups or observation and so tends to be time consuming to collect. Samples are usually smaller than with quantitative studies and are often locally based. Data analysis is also time consuming and consequently expensive.

Some data collection techniques are more relevant in Library and Information Sciences research: Interview, Focus groups, Observation, Document analysis, Survey

5.8.1 The interview

Interviews can be highly structured, semi structured or unstructured. *Structured interviews* consist of the interviewer asking each respondent the same questions in the same way. A tightly structured schedule of questions is used, very much like a questionnaire. The questions may even be phrased in such a way that a limited range of responses can be elicited. For example: "Do you think that health services in this area are excellent, good, average or poor? Bearing in mind the cost of conducting a series of one to one interviews, the researcher planning to use structured interviews should carefully consider the information could be more efficiently collected using questionnaires.

Semi structured interviews (sometimes referred to as focused interviews) involve a series of open ended questions based on the topic areas the researcher wants to cover. The open ended nature of the question defines the topic under investigation but provides opportunities for both interviewer and interviewee to discuss some topics in more detail. If the interviewee has difficulty answering a question or provides only a brief response, the interviewer can use cues or prompts to encourage the interviewee to consider the question further. In a semi structured interview the interviewer also has the freedom to probe the interviewee to elaborate on the original response or to follow a line of inquiry introduced by the interviewee. An example would be:

Interviewer: "I'd like to hear your thoughts on whether changes in government policy have changed the work of the doctor in general practice. Has your work changed at all?"

Interviewee: "Absolutely! The workload has increased for a start."

Interviewer: "In what way has it increased?"

Unstructured interviews (sometimes referred to as "depth" or "in depth" interviews have very little structure at all. The interviewer goes into the interview with the aim of discussing a limited number of topics, sometimes as few as one or two, and frames the questions on the basis of the interviewee's previous response. Although only one or two topics are discussed they are covered in great detail. The interview might begin with the interviewer saying: "I'd like to hear your views on the GP role in PCTs". Subsequent questions would depend on how the interviewee responded. Unstructured interviews are exactly what they sound like - interviews where the interviewer wants to find out about a specific topic but has no structure or preconceived plan or expectation as to how they will deal with the topic. The difference with semi structured interviews is that in a semi structured interview the interviewer has a set of broad questions to ask and may also have some prompts to help the interviewee but the interviewer has the time and space to respond to the interviewees responses.

Qualitative interviews are semi structured or unstructured. If the interview schedule is too tightly structured this may not enable the phenomena under investigation to be explored in terms of either breadth or depth. Semi structured interviews tend to work well when the interviewer has already identified a number of aspects he wants to be sure of addressing. The interviewer can decide in advance what areas to cover but is open and receptive to unexpected information from the interviewee. This can be particularly important if a limited time is available for each interview and the interviewer wants to be sure that the "key issues" will be covered.

Qualitative interviews should be fairly informal. Interviewees should feel as though they are participating in a conversation or discussion rather than in a formal question and answer situation. However, achieving this informal style is dependent on careful planning and on skill in conducting the interview. More information on the skills required of the interviewer can be found in the Trent Focus Resource Pack *Using Interviews in a Research Project.*

Semi structured interviews should not be seen as a soft option requiring little forethought. Good quality qualitative interviews are the result of rigorous preparation. The development of the interview schedule, conducting the interview and analysing the interview data all require careful consideration and preparation. These matters are discussed in the Trent Focus Resource Pack: 'Using Interviews in a Research Project'.

5.8.2 Focus groups

Sometimes it is preferable to collect information from groups of people rather than from a series of individuals. Focus groups can be useful to obtain certain types of information or when circumstances would make it difficult to collect information using other methods to data collection. They have been widely used in the private sector over the past few decades, particularly market research. They are being increasing used in the public sector.

Group interviews can be used when:

- Limited resources prevent more than a small number of interviews being undertaken.
- It is possible to identify a number of individuals who share a common factor and it is desirable to collect the views of several people within that population sub group.
- Group interaction among participants has the potential for greater insights to be developed.

5.8.2.1 Characteristics of a focus group

1. The recommended size of a group is of 6 – 10 people. Smaller than this limits the potential on the amount of collective information. More than this makes it difficult for everyone to participate and interact.
2. Several focus groups should be run in any research project. It would be wrong to rely on the views of just one group. The group may be subject to internal or external factors of which the investigator is unaware. This can lead to idiosyncratic results. Individual groups may not go very well: the members may be reluctant to participate or not interact well with each other and limited insight will be gained. Sufficient groups should be run to provide adequate breadth and depth of information but a small number of groups may achieve this, as few as three or four. There is no upper limit on the number of focus group interviews that could be held although this will be limited by resources.
3. The members of each focus group should have something in common, characteristics which are important to the topic of investigation. For example, they may all be members of the same profession or they may work in the same team. They may all be patients at a practice or have experienced a similar health problem or be receiving similar treatment. Participants might or might not know each other. There are advantages and disadvantages to both.
4. Following on from (3), focus groups are usually specially convened groups. It may be necessary or even desirable to use pre formed groups but difficulties may occur. This is usually due to the pre existing purpose of the group which

can lead to the group having a particular perspective or bias which limits their potential for providing information. For example, pressure groups or groups with some political basis.

5. Qualitative information is collected which makes use of participants' feelings, perceptions and opinions. Just as in individual interviews data collection and analysis is time consuming.
5. Using qualitative approaches requires certain skills. The researchers require a range of skills: groups skills in facilitating and moderating, listening, observing and analysing.

5.8.3 Observation

Not all qualitative data collection approaches require direct interaction with people. It is a technique that can be used when data collected through other means can be of limited value or is difficult to validate. For example, in interviews participants may be asked about how they behave in certain situations but there is no guarantee that they actually do what they say they do. Observing them in those situations is more reliable: it is possible to see how they actually behave. Observation can also serve as a technique for verifying or nullifying information provided in face to face encounters.

In some research observation of people is not required but observation of the environment. This can provide valuable background information about the environment where a research project is being undertaken. For example, an action research project involving an institution may be enhanced by some description of the physical features of the building. An ethnographic study of an ethnic population may need information about how people dress or about their non verbal communication. In a health needs assessment or in a locality survey observations can provide broad descriptions of the key features of the area. For example, whether the area is inner city, urban or rural; the geographical location; the size of the population. It can describe the key components of the area: the main industries; type of housing. The availability of services can be identified: number, type and location of health care facilities such as hospitals and health centres; leisure facilities; shopping centres.

5.8.3.1 Techniques for collecting data through observation

***a. Written descriptions*:** The researcher can record observations of people, a situation or an environment by making notes of what has been observed. The limitations of this are similar to those of trying to write down interview data as it occurs. First there is a risk that the researcher will miss out on observations because he is writing about the last thing he noticed. Secondly, the researcher may find his attention focusing on a particular event or feature because they appear particularly interesting or relevant and miss things which are equally or more important but their importance is not recognised or acknowledged at the time.

b. Video recording: This frees the observer from the task of making notes at the time and allows events to be reviewed time after time. One disadvantage of video recording is that the actors in the social world may be more conscious of the camera that they would be of a person and that their behaviour will be affected. They may even try to avoid being filmed. This problem can be lessened by having the camera placed in a fixed point rather than carried around. However, this means that only events in the line of the camera can be recorded limiting the range of possible observations.

c. Photographs and artefacts: Photographs are a good way of collecting observable data of phenomena which can be captured in a single shot or series of shots. For example, photographs of buildings, neighbourhoods, dress and appearance. Artefacts are objects which inform us about the phenomenon under study because of their significance to the phenomena. For example, memorabilia in historical research. Similarly, they may be instruments or tools used by members of a sub group whether this is a population sub group or a professional or patient group.

d. Documentation: A wide range of written materials can produce qualitative information. They can be particularly useful in trying to understand the philosophy of an organisation as may be required in action research and case studies. They can include policy documents, mission statements, annual reports, minutes or meetings, codes of conduct, etc. Notice boards can be a valuable source of data. Researchers who use this method of data collection sometimes develop a reputations as a "lurker" because of their tendency to lurk around notice boards! More information about observation can be found in the Trent Focus Resource Pack *How to use observation in a research project.*

5.8.4 Document Analysis

Document Analysis covers a broad range of techniques devoted to the analysis and interpretation of the documents used as primary data sources. The term "document" is understood very broadly, including not just texts, but also sound, photos, videos, and any materials that carry relevant messages. Typical varieties of document analysis include:

- Conversational Analysis
- Discourse Analysis
- Narrative Analysis
- Objective Hermeneutics

5.8.5 Survey

In surveys the subjects complete a survey form without the intervention of the researcher. Surveys are particularly useful when the subject population is large (or distributed geographically), majority opinions are sought, and the subjects are motivated to respond. Surveys can include closed and open questions, but the number of open questions should be reduced to a minimum. Special care must be taken to minimize the length of the survey, manage the choice of words and terminology, look after the balance of its structure, fully plan the strategies for its subsequent analysis, and test-pilot it thoroughly.

5.9 Handling qualitative research data

Interviewers have a choice of whether to take notes of responses during the interview or to tape record the interview. The latter is preferable for a number of reasons. The interviewer can concentrate on listening and responding to the interviewee and is not distracted by trying to write down what has been said. The discussion flows because the interviewer does not have to write down the response to one question before moving on to the next. In note taking there is an increased risk of interviewer bias because the interviewer is likely to make notes of the comments which make immediate sense or are perceived as being directly relevant or particularly interesting. Tape recording ensures that the whole interview is captured and provides complete data for analysis so cues that were missed the first time can be recognised when listening to the recording. Lastly, interviewees may feel inhibited if the interviewer suddenly starts to scribble: they may wonder why what they have just said was of particular interest.

The ideal tape recorder is small, unobtrusive and produces good quality recording. An in built microphone makes the participants less self-conscious. An auto reverse facility means that the tape will automatically "turn itself over" if the interview lasts longer than the recording time available on one side of the tape: this prevents an interruption in the flow of conversation. A tape recorder with a counter facility can be useful when analysing the taped data.

Transcribing qualitative data

Transcribing is the procedure for producing a written version of the interview. It is a full "script" of the interview. Transcribing is a time consuming process. The estimated ratio of time required for transcribing interviews is about 5:1. This means that it can take two and a half hours or more to transcribe a thirty minute interview. It also produces a lot of written text as one interview can run to up to 20 pages.

It may not be essential to transcribe every interview. It is possible to use a technique known as tape analysis which means taking notes from a playback of the tape recorded interview. If tape analysis is used the counter facility can be useful because the researcher can listen to the tape and make a note of the sections which contain particularly useful information and key quotations and return to these sections of the tape for fuller analysis. However, the previously mentioned problems of bias can occur if inexperienced qualitative researchers attempt tape analysis. It is certainly preferable to produce full transcripts of the first few interview data. Once the researcher becomes familiar with the key messages emerging from the data tape analysis may be possible.

The researcher should consider the question "who should do the transcribing? If the research is funded or supported by an employer there may be resources to pay an audio typist. This is usually more cost effective than a health care professional who will take longer and is more highly paid. However, if the transcriber is unfamiliar with the terminology or language contained in the interviews this can lead to mistakes or prolong the transcribing time.

Another procedure sometimes adopted when interviews are used in qualitative research is constant comparative analysis. This is a process whereby data collection and data analysis occur on an ongoing basis. The researcher conducts the first interview which may be unstructured or semi structured. The interview is transcribed and analysed as soon as possible, certainly before the next interview takes place, and any interesting findings are incorporated into the next interview. The process is repeated with each interview. When using this procedure it is quite possible that the initial interviews in a research projects are very different to the later interviews as the interview schedule has been continuously informed and revised by informants.

5.10 Qualitative Data Analysis

Qualitative data analysis involves summarizing data and presenting the results in a way that communicate the most important features. As quantitative research we are interested to discover the big picture in qualitative research as well, but by using different techniques. We start labeling or coding every item of information to recognize differences and similarities between all different items. It needs a method of identifying and coding items of data which appear in the text of transcript. All the items of data from one interview should be compared with other interviews. Same procedures are used for qualitative data collected through interviews, FGDs, observation and documentary analysis.

It involves the following processes:

- Coding data
- Finding Patterns
- Labeling Themes
- Developing Category Systems
- Looking for emergent patterns in the data

5.10.1 Content analysis

a. Procedure for categorization of verbal or behavioral data

b. It involves coding and classifying data

c. Analysis done at two levels:

- Basic or manifest level: descriptive – what was actually said
- Higher or interpretative level: what was meant by response – also called latent level of analysis

Content analysis involves the following steps:

1. Read the transcript and make brief note of interesting or relevant information
2. Make a list of the different type of information from the notes
3. Categorization of the listed items
4. Identify the categories that are some how linked to each other (major categories or themes)
5. Compare and contrast various categories
5. Repeat the process from stage 1-5 on next transcripts
   - Identify new categories of information

- Accommodate data in the existing categories
- Color code different categories and review

7. Collect together all the extracts from the transcribed interviews that you have put into one category
8. Review different categories and move items if required from one category to another
9. Review and check if two or more categories can fit together
10. Check the initial notes, consider if any previously excluded data is relevant and should be included in results

5.11 Basic Steps In Qualitative Research

- Write a tentative research proposal
- Intensive participation in a field setting
- Collect detailed data from field activities
- Synthesize and interpret the meanings of the field data
- Write the research report

5.12 Advantages

- Greater data accuracy than direct questioning, in natural settings people behave naturally;
- Problems of refusal, not at home, false response, non-cooperation etc. are absent;
- No recall error;

5.13 Limitations

- Time consuming, — too many things to observe,
- May not be representative,
- Difficulty in determining root cause of the behavior.

5.14 The Relevance of Qualitative Research to Library and Information Science

Among researchers who employed qualitative methods, only a few adopted the methodological approach. Almost all projects focused on the user's perceptions looked at processes, collected qualitative data and used more than one method. The three interrelated factors brought qualitative methods to the attention of LIS researchers:

a. the failure of quantitative methods to produce what is expected of them

b. the move toward a user-centred approach and

c. the growing interest in qualitative methods

A user centered approach implies that information related phenomena are studies from the user perspective that system design and evaluation are centered on the user not the system. This means that each user is unique, operating within a certain context and affected by situational conditions.

The growing interest in qualitative method in social sciences has also influenced researchers in LIS. LIS should join the move away from the positivism typical of contemporary social sciences because it cannot satisfy requirements that are essential to the LIS research, namely, holistic, reflective, empirical and dialectic.

Although it is a social science, and an interdisciplinary one. LIS need to develop its own qualitative methods and procedures. It would be useful if researcher describes not only the findings but also their methodological path, challenges, retreats, and successes. With such description, LIS can inductively develop a methodological framework and set of procedures adequate for its own research.

References & select Readings

Berg, B. L. (2007). *Qualitative Research Methods for the Social Sciences* (6th Ed.). Pearson Education.

Bernard, HR. (2002) *Research Methods in Anthropology: Qualitative and Quantitative Approaches.* Walnut Creek: AltaMira Press

Bernard, H.R. (2000). *Social Research Methods: Qualitative and Quantitative approaches.* Thousand Oaks, CA: Sage publication.

Boehm, B.W. (1981). *Software Engineering Economics.* Englewood Cliffs, N.J.: Prentice-Hall.

Bryman, A & Burgess, R (Eds.) (1993). *Analysing Qualitative Data.* London, Routledge.

Creswell, J.W. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Traditions.* Thousand Oaks, CA: Sage Publication.

Creswell, J. W. (2007). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3rd Ed.), Sage Publication.

Denzin, Norman K.; Lincoln, Yvonna S., eds. (2005). The Sage Handbook of Qualitative Research (3rd ed.). Thousand Oaks, CA: Sage Publication. ISBN 0-7619-2757-3

Flick, U W E. (2009). *An Introduction to Qualitative Research.*New Delhi: Sage Publication.

Gibbs, G. R. (2007). *Analyzing Qualitative Data. In U. Flick* (Ed.). *The Sage Qualitative Research Kit.* Sage Publication.

Hennink, Monique; Hutter, Inge & Bailey, Ajay (2011). *Qualitative Research Methods.* New Delhi: Sage Publication.

Kothari, C.R.(2009). *Research Methedology: Mehods & Techniques*. New Delhi: New Age International.

Krueger, RA and Casey, MA. (2000). *Focus Groups: A Practical Guide for Applied Research.* Thousand Oaks: Sage Publications.

Miles M, Huberman A (1994) *Qualitative Data Analysis.* Thousand Oaks, Sage Publication.

Munhall, P.L. (2001). *Nursing Research: A Qualitative Perspective*, 3rd Edition. Sudbury, MA: Jones and Bartlett

Patton, M.Q. (2002). *Qualitative Research & Evaluation Methods*, 3rd Edition. Thousand Oaks, CA: Sage

Puchta, C. and Potter, J. (2004) *Focus Group Practice.* London: Sage Publications

Stake RE (1995). *The Art of Case Study Research* .Thousand Oaks: Sage publication.

Yin, Robert K.(2011). *Qualitative Researchfrom Start to Finish.* New York: Guilford Press.

Sample Questions

1. Distinguish between Qualitative and quantitative research.
2. Define Qualitative research.
3. Mention the characteristics of Qualitative Research.
4. Describe the features of Qualitative Research.
5. Enumerate the different kinds of qualitative research design.
5. What is Triangulation and Mixed Methods Research?

7. State principles of quality management in the Qualitative Research Process.
8. State Models of research Process of qualitative research.
9. How do you proceed research through Qualitative Research Process?
10. Explain a typical qualitative research approach.
11. Explain the methods of collecting qualitative data.
12. What is the characteristics of focus group?
13. Mention and describe the Techniques for collecting data through observation.
14. What processes are involved in Qualitative Data Analysis?
15. Which steps are involved in Content analysis?
15. Describe the relevance of Qualitative Research to Library and Information Science.

Unit 6 ❑ Problems, Hypothesis / Research

Structure

6.1 Defining the research problem

In scientific inquiry, the first and foremost step is that of selecting and defining a research problem. A researcher must find the problem and formulate it so that it becomes susceptible to research.

A research problem refers to some difficulty which a researcher experiences in the context of either a theoretical or practical situation and wants to obtain a solution for the same. A research problem does exist if the following conditions are met with :

(i) There must be an individual to whom the problem can be attributed. The individual or the organisation, as the case may be, occupies an environment which is defined by uncontrolled variables.

(ii) There must be at two courses of action. A course of action is defined by one or more values of the controlled variables.

(iii) There must be atleast two possible outcomes from two courses of action of which one should be preferable to the other. Otherwise it means there must be atleast one outcome that the researcher wants, i.e., an objective.

(iv) The courses of action available must provide some chance of obtaining the objective, but they can not provide the same chance, otherwise the choice would not matter.

An individual or a group of persons can be said to have a problem which can be technically described as a research problem, if they (individual or group), having one or more desired outcomes, are confronted with two or more courses of action that have some but not equal efficiency for the desired objective (s) and are in doubt about which course of action is best.

6.1.1 Components of a research problem

The formulation of a research problem is vital in research process in order to obtain a suitable solution. The selection and formulation of a research problem is based on the following components identified by the researcher.

(i) There must be an individual or a group which has some difficulty or the problem.

(ii) There must be some objective(s) to be attained at. If one wants nothing, one cannot have a problem.

(iii) There must be alternative means for obtaining the objective(s) one wants to achieve. This means that there msut be least two means available to a researcher for if he has no choice of means, he cannot have a problem.

(iv) There must remain some doubt in the mind of researcher with regard to the selection of alternatives. This means that research must answer the question concerning the relative efficiency of the possible alternatives.

(v) There must be some environments to which the difficulty pertains.

Thus, a research problem is one which requires a researcher to find out the best solution for the given problem, i.e. to find out by which course of action the objective can be attained optimally in the context of a given environment.

6.1.2 Selection of research problem

The research problem undertaken for study must be carefully selected. The problem is selected by careful and extensive reading of appropriate research reports, articules in periodicals covering related areas or other literature. A research guide may help in choosing a research problem. A problem must spring from the researcher's mind. However, the following points may be observed in selecting a research problem or a subject for research or investigation :

(i) Subject which is overdone should not be normally choosen, for it will be a difficult task to throw any new light in such case.

(ii) Controversial subject should not the choice as research problem.

(iii) Too narrow or too vague problems should be avoided.

(iv) The subject selected for investigation should be familiar and feasible so that the related material or sources of research are within one's reach. A researcher may contact subject experts who are engaged in research work. He may discuss with others engaged in researches in the subject.

(v) The importance of the subject, the qualifications and training of a researcher, the costs involved, the time factor are few other criteria that must also be considered in selecting a problem.

(vi) The selection of the problem must be preceded by a preliminary study. This preliminary study is undertaken to justify the problem for conducting research work.

(vii) The problem so selected for research should be interesting to the researcher himself. If he is not interested in it, he will not be able to face and overcome the hurdles which may come at every step in research.

Goods and Hatt have given the following criteria for the selection of a research problem : (a) researcher's interest, intellectual curiosity and drive ; (b) practicability and implementation worthiness ; (c) urgency of the problem ; (d) anticipation of expected outcomes and their importance for the field represented ; and (e) resources, training and personal qualification of the researcher, availability of speical equipment, if any, data, method and sponsorship and finally administrators cooperation.

Thus, from the above mentined points it is clear that the selection of the research problem for a scientific inquiry is a key element in the research process.

6.2 Delimitation of the problem

Defining a research problem properly and clearly is a crucial part of a research study and must in no case be done hurriedly. However, in practice this is frequently overlooked which causes a lot of problems later on. Hence, the research problem is defined in a systematic manner by giving due weightage to all relating points discussed above. Dalimitation of a research problem determine the limits or boundaries of the research project undertaken. It will mention the geographical limits of the study. By limiting the boundaries of a research problem, the research will be on the track and it will help to discriminate relevant data from the irrelevant ones. The technique for the purpose involves the undertaking of the following steps generally one after another in doing a worthwhile research with an appropriate research design and carry it out meaningfully to its logical conclusion.

6.2.1 Stepts involved in Delimiting the problem

Thes step by step decisions in delimiting a problem are the following :

(i) **Statement of the problem in a general way** : First of all the problem must be stated in a broad general way, keeping in view either some practical concern or some scientific or intellectual interest. For this purpose, the researcher must immerse himself thoroughly in the subject matter concerned and if necessary a preliminary survery be made. The problem stated in a broad general way may contain various ambiguities which must be resolved by thinking again and again. At the same time the feasibility of a particular solution has to be considered and the same should be kept in view while stating the problem.

(ii) **Understanding the nature of the problem** : The next step in defining the problem is to understand its origin and nature clearly. The best way of understanding the problem is to discuss it with those who first raised it and what are its objectives. For a better understanding of the nature of the problem involved, the researcher can have discussion with those who have a sound knowledge of the problem concerned or similar other problems. The researcher should also know the environment within which the problem is to be studied and understood.

(iii) **Surveying the available literature** : All available literature relating to the problem concerned must necessarily be surveyed and examined before a definition of the research problem is given. This means that the researcher must be well acquianted with the relevant theories in the subject field, reports and records, and also other relevant sources. The researcher must give due weightage to the researches already done on related studies and make a review of it. This is done to find out what data and other materials, if any, are available for operational purposes. Knowing what data are available often serves to narrow the problem itself as well as the technique that might be used. This may help the researcher to know if there are certain gaps in the theories, or whether the existing theories applicable to the problem under study are inconsistent with each other, or whether the findings of the different studies do not follow a pattern consistent with the theoretical expectations and so on. At times such studies may also suggest useful and even new lines of approach to the present problem.

(iv) **Developing the ideas through discussions** : A researcher may get useful information by disusing the subject with his colleagues or other persons having enough experience in the same area or similar problems. This is quite often known as an 'experience survey'. People with rich experience can enlighten and can give advice the researcher on different aspects for his proposed study. They help the researcher to sharpen his focus of attention on specific aspects within the field.

(v) **Rephrasing the research problem** : Finally, the researcher must rephrase the research problem into a working proposition. Once the nature of the problem has been clearly understood, the environment has been defined, discussions over the problem have taken place and the available literature have been surveyed and examined, rephrasing the problem into analytical or operational terms is not a difficult task. By rephrasing the research problem at hand, it may become operationally viable and may help in the development of working hypothesis.

In addition to what has been stated above, the following points must also be observed while defining a problem :

(a) Technical terms or words or phrases, with special meanings used in the statement of the problem, should be clearly defined.

(b) Basic assumptions or postulates (if any) relating to the research problem should be clearly stated.

(c) A straight forward statement of the value of the investigation, i.e. the criteria for the selection of the problem should be provided.

(d) The suitability of the time period and the sources of data available must also be considered by the researcher in defining the problem.

(e) The scope of the investigation or the limits within which the problem is to be studied must be mentioned explicitly in defining a research problem.

Thus, a research problem very often, follows a sequential pattern – the problem is stated in a general way, the ambiguities are resolved, thinking and rethinking process results in a more specific formulation of the problem so that it may be a realistic one in terms of the available data and resources and is also analytically meaningful. All this results in a well defined research problem that is not only meaningful from an operational point of view, but is equally capable of paving the way for the development of a working hypothesis and for means of solving the problem itself.

6.3 Exercise

1. Define hypothesis and indicate its difference with the theory.
2. What are the components of a research problem ?
3. How a research problem is delimited ?
4. Discuss the method of testing the hypothesis.

6.4 Summary

Hypothesis is an important component in a research problem. A research problem is defined in the light of the hypothesis. A defined hypothesis must be tested to validity of facts or information. Thus hypothesis plays an important role in formulation of a research problem.

6.5 References

1. Ghosh, B. N. Scientific method and social research. New Delhi : Sterling Publishers, 1987.
2. Kothari, C. R. Research methodology : Methods and Techniques. New Delhi : Wishwa Prakashan, 1990.
3. Kumar, P. S. G. Research methods and statistical techniques. New Delhi ; B. R. Pub. 2004.
4. Goode, W. J. and Hatt, P. K. Methods in social research. London : Mc Graw Hill Book, 1981.
5. Young, P. V. Scientific social surveys and research. New Delhi : Prentice Hall of India, 1984.

Unit 7 ❑ Literature Review

Structure

7.0 Objectives

This unit will give you an understanting of the following aspects :

— defining the problem ;
— meaning and definition of the hypothesis ;
— difference between hypothesis and theory ;,
— technique involved in delimiting the problem ;
— testing the hypothesis ;

7.1 Introduction

Hypothesis is usually considered as an important instrument in research. It is a formal affirmative statement predicting a single research outcome, a tentative explanation of the relationship between two or more variables. It is an assumption based statement of a proposition or a guess which the researcher proves through his study. J. S. Mill pointed out that a scientific hypothesis is "of such a nature as to be either proved or disproved by comparison with observed facts."

The researcher or the investigator cannot enter the field with a blank mind. Normally, the researcher begins the task of investigation with some ideas about the subject matter vaguely formulated. Then the searcher proceeds to find out whether the ideas conceived by him are totally correct or partially correct or totally false. But these ideas may be useful in understanding the problem. These preliminary ideas which guide the investigator in his study may be termed as hypothesis.

The important functions that hypothesis serve in scientific inquiry are the development of theory and the statements of parts of an existing theory in testable form. A hypothesis is formulated in such a way that this hunch can be tested. A hypothesis is derived from a theory, from past experience, observations, or information gained from others. For the purpose of testing the hypothesis, the variables must be defined and the researcher specifies what operations are to be conducted or tests used to measure each variable. Then, the hypothesis focuses the investigation on a definite target and determines what observations, or measures, are to be used.

7.2 Meaning and definitions of hypothesis

Hypothesis is defined by various scientists in their own ways. The term hypothesis has two parts : hypo means 'less than' and thesis means a proposition to be proved. Thus 'hypothesis' is less than generally held view'.

P. V. Young defines hypothesis as 'a provisional central idea which becomes the basis for fruitful investigation, is known as working hypothesis.'

According to J. S. Mill, 'a hypothesis is only an unproved supposition, a weak form of proof.'

Goode and Hatt defines hypothesis in a more elaborate way, "A hypothesis states what we are looking for. It is a proposition which can be put to a test to determine its validity. It may prove to be correct or incorrect."

Barr and Scaties states, "a hypothesis is a statement temporarily accepted as true..., when the hypothesis is fully established, it may take the form of facts, principles or theories."

From the above definitions it is clear that a hypothesis is nothing but a series of assumptions, presumptions made on the basis of probablities and profound hunches. In other sense, hypothesis a tentative generalisations and a statement temporarily accepted as true, the validity of which remains to be tested. It is a statment in research which the study or investigation may prove or disprove.

7.2.1 Hypothesis and theory

Hypothesis and theory are closely related to each other. In a scientific inquiry, a hypothesis is formulated as a tentative supposition or guess. But when a hypothesis is tested and validated and found to be true, it forms a theory. In other words, if a hypothesis is tested and established through the process of deductions and further, if the conclusions thus reached and found true, for further testing over a number of investigations, it becomes a true. This theory when it works satifactorily is generally accepted and acts as an instrument for further explanation till it is disproved or rejected.

Hypothesis is not the same as theory though they are closely related. According to W. H. George, "in practice a theory is an elaborate hypothesis which deals with more type of facts than does the simple hypothesis...". The hypothesis is drawn from the theory but the visc-versa is not true. When the hypothesis is put to test and found correct, it becomes a part of the theory.

Goode and Hatt explain, "the formulation of the deduction, however, constitutes a hypothesis, if verified, it becomes part of a future theoretical ...." The early stages of a theory is a hypothesis. The distinction between a theory and a hypothesis is one of degree, rather than of kind. Thus the relationship between the two is very close indeed.

7.2.2 Testing the Hypothesis

The function of the hypothesis is to state a specific relationship between phenomena in such a way that this relationship can be empirically tested. That is, the hypothesis must be empirically demonstrated as either probable and not probable. The basic method of this demonstration is to design the research so that logic will require the acceptance or rejection of the hypothesis, on the basis of resulting data. This requires control of the observation in order to eliminate other possible relationships. A basic aspect of research design, therefore, is setting up the research so as to allow logical conclusions to be drawn.

The basic design of logical proof were formulated by John Stuart Mill and is considered as the foundation of experimental procedure, although many refinements have been made. He explained two methods of arriving at logical conclusion to test the hypothesis :

(a) Method of agreement ; and

(b) the Negative canon of agreement.

Method of agreement may be stated as, when two or more cases a given phenomena have one and only one condition in common, then that condition may be regarded as the cause (or effect) of the phenomenon. More simply, if we can make observation 'Z' in every case that we find condition 'C', we can conclude that they are causally related.

Elements of situation X

A	B	C

—— Produce —— Z

Elements of situation Y

C	D	E

—— Produce —— Z

From the above figure, if it is known that all the conditions in the two situations, X and Y, are described and designated by A, B, C, D and E and further if it is known both sets of conditions result in observation 'Z', then it must be concluded that C and Z are related as cause and effect.

The negative canon of agreement is as follows : When lack of certain phenomena are always or nearly always found along with lack of certain other phenomena, we may conclude that there is a causal relationship between them.

To test a hypothesis using a set of data, a sample selected at random is subjected to a detailed analysis, for it may not be possible to evaluate the universe as a whole. The result obtained from the observed data is then compared with the expected one if the hypothesis is true. The difference in sample result and the expectation, the doubt of validity of the original hypothesis is noticed. At some level of doubt, the hypothesis may be rejected and this level of doubt may be expressed as a probablity. Thus, on the basis of probablity estimates, a hypothesis may be rejected or accepted after it has been tested.

Hypothesis testing helps to decide on the basis of a sample data, whether a hypothesis about the population under investigation is likely to be true or false. Several statistical tools are used to test the hypothesis which are classified as :

(a) Parametric tests or standard tests of hypothesis ; and

(b) Non-parametric tests or distribution-free test of hypothesis.

Parametric tests usually assume certain properties of the parent population from which we draw samples. Statistical methods used for testing hypothesis which are called non-parametric tests because such tests do not depend on any assumption about the parameters of the parent population.

Literature review is an important step in any research process. Any scholarly communication is incomplete rather erroneous without literature review. Literature review surveys scholarly articles, books and other sources (e.g. dissertations, conference proceedings) relevant to a particular issue, area of research, or theory, providing a description, summary and critical evaluation of each work. The purpose is to offer an overview of significant literature published on a topic.

Besides enlarging knowledge about the topic, writing a literature review let to gain and demonstrate skills in two areas-

a. **Information Seeking:** the ability to scan the literature efficiently, using manual or computerized methods, to identify a set of useful articles and books

b. **Critical Appraisal:** the ability to apply principles of analysis to identify unbiased and valid studies.

A literature review must do these things-

- be organized around and related directly to the thesis or research question
- synthesize results into a summary of what is and is not known
- identify areas of controversy in the literature
- formulate questions that need further research

According to Caulley (1992) of La Trobe University, the literature review should:

- compare and contrast different authors' views on an issue
- group authors who draw similar conclusions
- criticise aspects of methodology
- note areas in which authors are in disagreement
- highlight exemplary studies
- highlight gaps in research
- show how the present study relates to previous studies
- show how the present study relates to the literature in general
- conclude by summarising what the literature says

In assessing each piece, consideration should be given to:

a. **Provenance**—What are the author's credentials? Are the author's arguments supported by evidence (e.g. primary historical material, case studies, narratives, statistics, recent scientific findings)?

b. **Objectivity**—Is the author's perspective even-handed or prejudicial? Is contrary data considered or is certain pertinent information ignored to prove the author's point?

c. **Persuasiveness**—Which of the author's theses are most/least convincing?

d. **Value**—Are the author's arguments and conclusions convincing? Does the work ultimately contribute in any significant way to an understanding of the subject?

Organizing the body in Literature Review

Once researchers have the basic categories in place, they must consider present sources within the body of the paper. Create an organizational method to focus this section even further. To come up with an overall organizational framework for the review, consider the following scenario and then three typical ways of organizing the sources into a review:

- Chronological review-

It follows the chronological method, one can write about the materials above according to when they were published.

- By publication-

Order the sources by publication chronology, then, only if the order demonstrates a more important trend.

- By trend-

A better way to organize the above sources chronologically is to examine the sources under another trend, such as the history of topic.

- Thematic-

Thematic reviews of literature are organized around a topic or issue, rather than the progression of time. However, progression of time may still be an important factor in a thematic review.

But more authentic thematic reviews tend to break away from chronological order. For instance, a thematic review of material on the topic might examine how they are portrayed as evolutionary development. The subsections might include how they are personified, how their proportions are exaggerated, and their principles. A review organized in this manner would shift between time periods within each section according to the point made.

- Methodological-

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

Sometimes though, the researcher might need to add additional sections that are necessary for his/her study, but do not fit in the organizational strategy of the body. Whatever it is necessary to focus-

Current Situation: Information necessary to understand the topic or focus of the literature review.

History: The chronological progression of the field, the literature, or an idea that is necessary to understand the literature review, if the body of the literature review is not already a chronology.

Methods and/or Standards: The criteria used to select the sources in his/her literature review or the way in which information is presented. For instance, they might explain that his/her review includes only peer-reviewed articles and journals.

Examples of Literature Review

Example-1

Define Topic: Statistical validity of h-index

Literature Review: Concerning the first stream of research, Glanzel (2006) analyzed the basic mathematical properties of the h index thanks to the adoption of the Paretian distribution for the citation count, stressing the strength of such index when the available set of papers is small (that is the case for young researchers mainly). Iglesias and Pecharroman (2007) proposed to use a simple multiplicative correction to the h index able to take into account the differences among researchers coming from different science citation index (SCI) fields and thus allowing a fair and sustainable comparison. Indeed these authors offer a table with such normalizing factors according to specific distributional assumptions of the citation counts (power law or stretched exponential model). Burrell (2007) made a step ahead since he proposed to employ a stochastic model for an author's production/citation patterns. In that framework it is possible to consider different situations according to the level of production and citation or the length of a researcher's career. Beirlant and Einmahl (2010) and Pratelli et al. (2012). Beirlant and Einmahl (2010) demonstrated the asymptotic normality of the empirical h index for the Pareto-type and Weibull-type distribution families, allowing the construction of asymptotic confidence intervals of each author and evaluating the statistical significance of the difference between two authors with the same academic profile (in terms of career length and SCI field). Very recently Pratelli et al. (2012) investigated, in a full statistical perspective, the distributional properties of the h index and the large sample expressions of its relative mean and variance, in a discrete distributional context.

Annotated Bibliography:

Beirlant, J. & Einmahl, J.H.J. (2010). Asymptotics for the Hirsch index. Scandinavian Journal of Statistics, 37, 355–364.

Burrell, Q.L. (2007). Hirsch's h-index: A stochastic model. Journal of Informetrics, 1, 16–25.

Glanzel, W. (2006) On the h-index—A mathematical approach to a new measure of publication activity and citation impact. Scientometrics, 67, 315–321.

Iglesias, J.E. & Pecharroman C. (2007). Scaling the h-index for different scientific ISI fields. Scientometrics, 73, 303–320.

Pratelli, L., Baccini, A., Barabesi, L. & Marcheselli, M., (2012). Statistical analysis of the Hirsch index.Scandinavian Journal of Statistics, 39, 681–694.

Example-2:

Define Topic: Developing a Holistic Model for Digital Library Evaluation

Literature Review: Few studies report their DL evaluation at this level. Essentially, criteria are employed to assess four types of digital content: digital object, metadata, information, and collection. Among these four types, digital objects seem to be the unique type to DLs and, hence, be evaluated under DL specific criteria, such as fidelity (Kenney, Sharpe, & Berger,1998) and suitability to original artifact (Goodrum, 2001;Jones & Paynter, 2002). The remaining three types have been evaluated with conventional criteria, including accuracy,

clarity, cost, ease of understanding, informativeness, readability, timeliness, and usefulness. Additionally, scalabilityfor user communities (Kengeri, Seals, Reddy, Harley &Fox, 1999; Kenney et al., 1998; Larsen, 2000) tackles a crucial issue in DL innovation, which involves more diverse user communities with various backgrounds and changing needs. Digital technology evaluation has two foci: hardware and software. The latter uses primarily conventional relevance based effectiveness measures, while several studies adapt them to fit into digital and hypermediated circumstances (Hee, Yoon &Kim, 1999; Salampasis & Diamantaras, 2002). As for hardware evaluation, display quality and robustness for digital information are frequently used to evaluate electronic and communication devices. Meanwhile, reliability, cost, and response time are used for both hardware and software evaluations. Interface is the most heavily evaluated DL level. Moreover, compared with the other five DL levels, interface evaluations tend to have more ready-to-use frameworks and criteria checklists, such as Wesson's (2002) multiple view, Nielson's (1993) five measures and 10 principles, Dillon's (1999) TIME framework, and Mead and Gay's (1995) evaluation

tool. Nevertheless, only Nielsen's (1993) usability test attributes (learnability, efficiency, memorability, errors, and satisfaction) receive wide adoptions (e.g., Prown 1999; Peng, Ramaiah, & Foo, 2004).

Annotated Bibliography:

Dillon, A. (1999). TIME-A multi-level framework for the design and evaluation of digital libraries. International Journal of Digital Libraries, 2 (2/3), 170–177.

Goodrum, A.A. (2001). Multidimensional scaling of video surrogates. Journal of American Society for Information Science, 52(2), 174–182.

Hee, M.,Yoon,Y.I.,&Kim, K.C. (1999). Unified video retrieval systems supporting similarity retrieval. In T.J.M. Bench-Capon, G. Soda,&A.M. Tjoa(Eds.), Proceedings of the Tenth InternationalWorkshop on Database and Expert System Applications (pp. 884–888). NewYork: Springer.

Jones, S., & Paynter, G.W. (2002). Automatic extraction of document keyphrases for use in digital libraries: Evaluation and application. Journal of the American Society for Information Science & Technology, 53(8), 653–677.

Kengeri, R., Seals, C.D., Reddy, H.P., Harley, H.D., & Fox, E.A. (1999). Usability study of digital libraries: ACM, IEEE-CS, NCSTRL, and NDLTD. International Journal on Digital Libraries, 2(2/3), 157–169.

Kenney, A.R., Sharpe, L.H., & Berger, B. (1998). Illustrated book study: Digital conversion requirements of printed illustration. In C. Nikolaou & C. Stephanidis (Eds.), Research and advanced technology for digital libraries: Proceedings of the Second European Conference (pp. 279–293).Berlin, Germany: Springer.

Kenney, A.R., Sharpe, L.H., & Berger, B. (1998). Illustrated book study: Digital conversion requirements of printed illustration. In C. Nikolaou & C. Stephanidis (Eds.), Research and advanced technology for digital libraries: Proceedings of the Second European Conference (pp. 279–293).Berlin, Germany: Springer.

Larsen, R. (2000). The DLib test suite and metrics working group: Harvesting the experience from the digital library initiative. Retrieved onApril 13,2004, from http://www.dlib.org/metrics/public/papers/The_Dlib_Test_Suite_and_Metrics.pdf

Mead, J.P., & Gay, G. (1995). Concept mapping: An innovative approach to digital library design and evaluation. ACM SIGOIS Bulletin, 16(2),10–14.

Nielsen, J. (1993). Usability engineering. Boston MA: Academic Press.

Peng, L.K., Ramaiah, C.K., & Foo, S. (2004). Heuristic-based user interface evaluation at Nanyang Technological University in Singapore. Program-Electronic Library and Information System, 38(1),42–59.

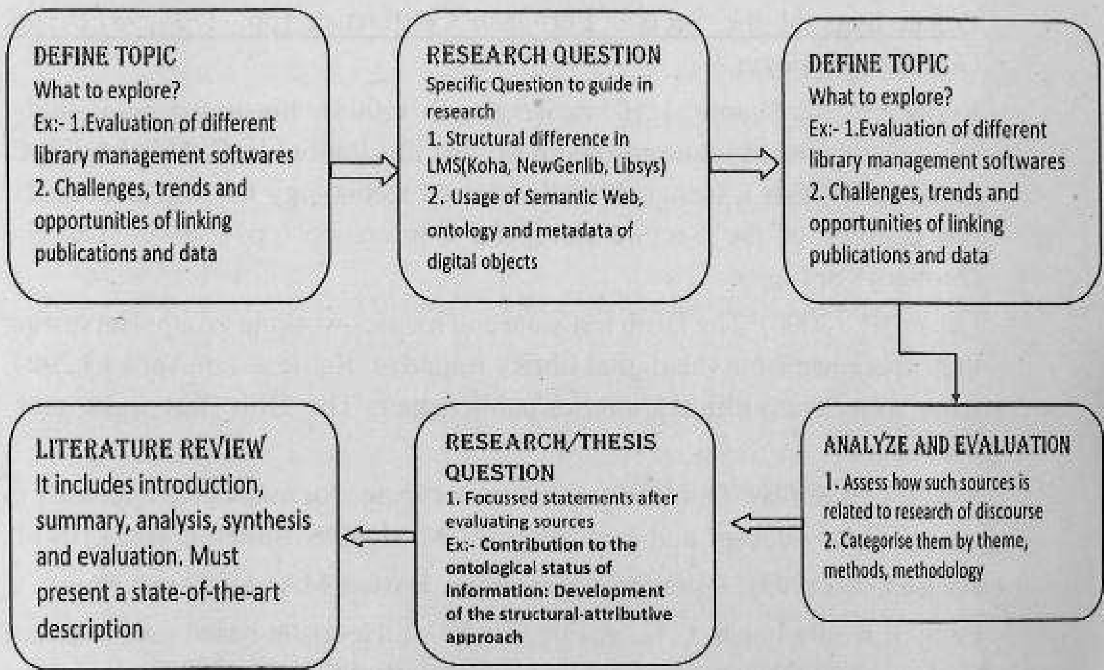
Prown, S. (1999). "Detecting 'broke': usability testing of libraryWeb sites." Retrieved on December 12, 2004, at http://www.library.yale.edu/<"prowns/nebic/nebictalk.html

Salampasis, M., & Diamantaras, K.I. (2002). Experimental usercentered evaluation of an open hypermedia system and Web information seeking environments. Journal of Digital Information, 2(4). Retrieved September 29, 2009, from http://journals.tdl.org/jodi/article/viewArticle/57/60

Wesson, J., & Greunen, D.V. (2002). Visualization of usability data: Measuring task efficiency. In P. Kotzé, L. Venter, & J. Barrow (Eds.),Proceedings of the 2002 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on Enablement through Technology (pp. 11–18). New York: ACM Press.

Appendix-1

Literature Review in Schematic Representation



Unit 8 ❑ Data Collection : Tools and Techniques

Structure

8.1 Objectives

The objective of this unit is to aware researchers about the sources of relevant data including tools and techniques required for obtaining the data. This unit will deal with the most frequently used data collection techniques such as observation, interview, questionnaire, case study,schedule etc. These methods for gathering data are most commonly, but not exclusively used in research. These can be otherwise called instruments but not research methodologies and more than one techniques can be used for the purpose of research.

8.1.1 Introduction

Data can be considered as small cells of an organism with which research question or argument is constructed. It is data which sustain the progress of a research, function as evidence to validate what is proposed or stated by the researchers, and bring about the realisation of the final result. Data may be primary or secondary.

Primary data are first-hand data being produced by the participants in or witnessed of an event or a phenomenon. Sometimes primary data are called hands-on data, data one gets by one's own personal efforts through experiments, laboratory investigations or field work.

Secondary data are produced through making use of primary data; that is, secondary data are based on or derived from primary data. Secondary sources may be used for back ground study,for peripheral or fringe information.

Data collection usually involves measuring some research phenomenon, whether it is a process, an object, or a human subject's behaviour. The objects of measurement will differ, of course, from one research projects to another, depending upon the purpose of the inquiry and the availability of suitable instruments. Physical scientists use some kinds of devices e.g. thermometers, barometers, Geiger counters electron microscopes, yard sticks and rulers. The instruments or tools used by behavioural or social scientists include such as tools as questionnaires, written or oral tests, checklists, attitude scales, and similar other indexes. Thus experiments in case of scientific research, interviews and questionnaires for social or behavioural sciences research are generators of primary data. Data gathering methods used in case studies are primarily based on direct observation; both participant and nonparticipant observation can be used. When necessary, these methods are supplemented by structured techniques such as interviews and questionnaires. It can be concluded thatinterviews and questionnaires are powerful tools for generating primary data about the current status of events or phenomena.

Data collection represents the key point of any research project. Some of the methods of data collection such as interviewing and questionnaires, are probably more popular. Some methods entail a rather structured approach to data collection—that is, the researcher establishes in advance the broad contours of what he or she needs to find out about and designs research instruments to implement what needs to be known. The questionnaire

is an example of such an instrument; the researcher establishes what he or she needs to know to answer the research questions that drive the project and designs questions in the questionnaire that will allow data to be collected to answer those research questions. Similarly, something like a structured interview—the kind of interview used in survey investigations—includes a host of questions designed for exactly the same purpose. There is a difference between research questions and the kinds of questions, thatare posed in questionnaires and interviews. They are very different: a research question is a question designed to indicate what the purpose of an investigation is; a questionnaire question is one of many questions that are posed in a questionnaire that will help to shed light on and answer one or more research questions.

There are methodsof data collection that are less structured or, to put itanother way, that are more unstructured. Research methods will be encountered thatemphasize a more open-ended view of the research process,so that there is less restriction on the kinds of things that can be found out about. Research methods suchas participant observation and semi-structured interviewingare used so that the researcher can keep more of an open mind about the contours of what he or she needsto know about, so that concepts and theories can emerge out of the data. This is the inductive approach to theorizingand conceptualization that was referred to above.

Techniques for collecting data can be classified as follows:

8.2 Observation

One of the most widely used data collection technique as well as tool in research is observation. It is the process of acquiring knowledge and information through senses. Observation simply means to 'see' a situation or a phenomenon. The term observation may be defined as, the object or subject of an investigation is being subjected to over on the basis of usual surveillance and that the information obtained will then be related to more general propositions or theories. Observation becomes a scientific tool and the method of data collection when it serves a formulated research purpose which is systematically planned and recorded and is subjected to checks and controls on validity and reliability. According to this technique the information is sought by way of investigator's own direct observation without asking from the respondents. This method is particularly suitable in studies which deal with subjects who are not capable of giving verbal reports of their feelings for one reason or the other.It is widely used in descriptive survey method.

Gathering information or data through observation of social behaviour is an ancient practice. Stories of ancient kings who in disguise would stroll in the streets to observe and thereby know how their subjects lived and what opinions they held about their king and his rule were very popular.

Observation is the process by which a researcher or a team of researchers observes what is going on in some real life situation, and collects relevant or needed data(i.e., by recording pertinent happenings) according to some plan formulated to meet the need of the research undertaken. A researcher while he conducts his study observes numerous things around him but takes into account only those that are relevant for his study. It is the technique used for collecting facts through the use of sensory organs like ears, eyes and nose. Along with sensation concentration for observing things is also required. Attention is thus an important component of observation. Facts are recognised in observation only relationship are drawn from previous experience and knowledge. This is known as perception.Three components of observation are:

- Sensation
- Attention
- Perception

Observational methods have been employed in natural history-a field that owes much to the scientists who observed, recorded and classified what they saw. Much of social science research, particularly educational research, take recourse to the observational technique. It is used to evaluate the overt behaviour of individuals in controlled and uncontrolled situation.

The direct surveillance of dimensions of a phenomenon that is to be measured or evaluated. This close perception of a phenomenon facilitates a detailed and exact explanation of how the phenomenon behaves under known conditions. An example of observation is watching students use the library online public access catalogue to determinethe effectiveness of their search habits.

8.2.1 Kinds of Observation

Observation may be divided into two kinds:

- **Participant Observation**
- **Non-participant Observation**

8.2.1.1Participant Observation

In this observation method the researcher becomes a member of a group or sample under observation. He /she(or the team) may play the role, in varying degrees of participating, of a visiting stranger, an attentive listener, an eager learner, or a more complete participant observer.

8.2.1.2 Non-participant Observation

In this observation the observer carries on his job without making his presence disturbing to the group or sample under observation. He or she may profitably make use of a one-way vision screen that permits him to see the subject but prevents the subject to see him.

Observation may also be classified as:

- **Structured Observation**
- **Unstructured Observation**

8.2.1.3 Structured Observation

It is put on aformal basis and is designed to test casual hypotheses. It is mainly carried on in controlledconditions, such as in library in the case of library and information science research. Interaction analysis of the reference section verbal behaviour of any librarypersonnel may be taken as an example.

Structured observation begins with specific formulation or plan. The content of observation is more or less defined and has very little scope for change. The observer fixes in advance categories of behaviour in respect of which he requires to analyse the problem in hand, and keeps in mind the time limit within which he is to execute the observation.

Steps that can be taken to increase the reliability of structured observation include the following:

- Developing adequate definitions of the kinds of behaviour that are to be recorded, and being certain that they correspond to specific concepts to be studied.
- Carefully training the observers to ensure that they are adequately prepared and that they have confidence in their ability or judgement to check the appropriate categories.

- Avoiding observer bias, generally the observer should take behaviours at their face value and not attempt to interpret their "real" meaning at least not at the time the observations are made.

8.2.1.4 Unstructured Observation

Unstructured observation is mainly associated with the participant observation and often takes the form of exploratory technique. In the unstructured observation categorization of behaviour may not be made in advance. Free from predetermined categories, the observer can consider aspects of behaviour according to their context or the situation of which these are part.

There are steps which can be taken to increase the accuracy of unstructured observation:

- Using two or more observational techniques, such as sound and visual recordings, and then comparing results.
- Having two or more people observe the same behaviour, with the same technique, and then comparing the results.
- Being careful to distinguish between actual behaviour and perceptions or interpretationsof the behaviourwhen taking notes. Researcher bias can easily creep in during this stage.
- Avoiding becoming involved in the activity being observed can also help to increase accuracy and reduce bias.
- Being careful not to take behaviour for granted can improve accuracy. For example, a researcher observing online public access catalogue and the use could not assume that every time a patron is searching different title. The patron may be browsing for the same title under one or more heading or search terms.
- Obtaining reactions from the participants regarding the accuracy of the observations can be useful in situations where the subjects are fully aware of the role of the researcher. But one would have to be careful that doing so did not affect or bias future behaviour of the subject.

Also Observation can be classified as:

- Controlled Observation
- Uncontrolled Observation

8.2.1.5 Controlled Observation

Observation when takes place with defined pre-arranged plans, involving experimental procedure is known as controlled observation. This type of observation is conducted in the laboratories.

8.2.1.6 Uncontrolled Observation

Observation when takes place in natural settings it may be termed as uncontrolled observation. Here no attempt is madeto use precision instruments. The major aim is to get a spontaneous picture of life and persons. Uncontrolled observation is done in case of exploratory studies.

8.2.1.7 Random Observation

Sometimes researchers select appropriate environments for conducting their research. A library,for example, is a natural setting for research studies in library systems and services and a researcher, placing himself or herself in it, may record observations on a sample of events he happens to observe. These samples are random observations. Random observations are essentially free from bias and neutrality is maintained. Here samples are considered as representative which lead to valid conclusions about the phenomenon.

8.2.2 Recording of Observation

The recording of observation data may be done either simultaneously with the observation, or soon after the observation has been completed. In the former method, phenomena or incidents are recorded during the time when they occur. In the latter method, the observer does not record what he observes then and there, but records immediately or soon after what he has observed while the details of his observation are still fresh or vivid in mind.

The recording must be un biased and objective.

8.2.3 Advantages of Observational Research

As a data collection technique, observation has several important advantages they include the following:

- The use of observation makes it possible to record behaviour as it occurs.
- Observation allows one to compare what people actually did with what they said they did. Participants in a study may consciously or unconsciously report their behaviour as different than it in fact occurred; the observed behaviour may well be more valid.
- Observational techniques can identify behaviour, actions. etc. that people may not think to report because they seem unimportant or irrelevant.

- With observational techniques, a researcher can study subjects who are unable to give verbal reports.
- The use of observation is generally independent of subjects' willingness to participate. For example, one could observe how library users are using Online Public Access Catalogue without asking each user beforehand if he or she were willing to be observed. There are ethical and sometimes legal implications that should be explored before deciding to observe persons without their permission, or at least awareness.

8.2.4 Limitations of Observational Research

Observational techniques do suffer from a few limitations; some of the more important ones are:

- It is not always possible to anticipate a spontaneous event and thus be prepared to observe it. Some of the most critical activity at the catalogue may even take place when no one is there to observe.
- The duration of an event affects the feasibility of observing it. The activities at a catalogue are generally short enough to be easily observed; such would not be the case in trying to observe how a faculty member conducts research.
- Some types of behaviour are obviously too private or personal in nature to be observed. This is less of a disadvantage in library related research,however than it is in the behavioural sciences.
- It is generally somewhat more difficult to quantify observational data than other kinds. Behaviour simply cannot always be broken down into neat categories.

8.3 THE QUESTIONNAIRE

Data collection in the form of questionnaire is the most popular research method. According to Webster's new Collegiate dictionary questionnaire is "a set of questions for submission to a number of persons to get data..."

A questionnaire is a formulated series or list of questions systematically prepared which are meant for collecting data on a problem under investigation. The person who answers a question is the respondent. Questionnaires are sent to respondents (i) personally, (ii) by messenger and (iii) by mail or post. Usually mail or post is the means of sending questionnaires. The response to a questionnaire consists of written answers from the respondent who also puts his signature on the questionnaire document. Obviously the questionnaire involves the literate respondent.

8.3.1 Types of Questionnaire

Questionnaires are of two types such as:

- Open (sometimes called Open-end) or Unstructured
- Closed or Structured

8.3.1.1 Open Questionnaires

Open questionnaires require the respondents to write their answers at some length- in many sentences-at least for a few or several questions. Closed questionnaires, on the other hand, provide the respondents with generally readymade alternative answers words provided by the researcher. A respondent has only to choose from the alternatives the ones he thinks are the right answers; only occasionally he is required to write answers and that only in single words, phrases, or sentences.

8.3.1.2 Closed Questionnaire

A closed or structured questionnaire may itself be of two types. One is Yes/No type, in which the respondent answers by ticking either "Yes" or "No" as the case may be. In the other type, two alternative answers are given, and the respondent ticks one or as many as he thinks pertinent.

Sometimes a closed questionnaire may take the form of pictorial questionnaire, in which the alternative answers are in pictures, and the respondent answers by ticking the pertinent picture(s). It may be used when information is sought from children, or from illiterates, or when the input of questions can best be expressed in pictures. But the pictorial questionnaire is very costly and also bulky in form.

8.3.1.3 Mixed Questionnaire

In practice the researcher usually has to use what may be called a mixed questionnaire. Here both closed and open questionnaires are mixed. It is one in which the closed and the open questionnaires, and even the pictorial questionnaire, are mixed.

8.3.2 Types of Questions

There are three basic types of questions:

1. **Descriptive.**When a study is designed primarily to describe what is going on or what exists. Public opinion polls that seek only to describe the proportion of people who hold various opinions are primarily descriptive in nature. For instance, if any one wants to know what percent of the population in a library for the

purpose of book selection would vote for a hardcopy books and journals or softcopy of the same.Here simply interest is on describing something.

2. **Relational.**When a study is designed to look at the relationships between two or more variables. A user survey that compares what proportion of males and females say they would prefer for hard copy of books or soft copy of books in the next book selection is essentially studying the relationship between gender and book selection preference.
3. **Causal.**When a study is designed to determine whether one or more variables (e.g., a program or treatment variable) causes or affects one or more outcome variables.In a library book purchase procedure an exhibition on new arrivals can be arranged to try to determine the changes of user preferences. This will study whether the exhibition(cause) changed the proportion of user who would select hard copy books or stick to electronic soft copy books.

On the other hand survey questions can be divided into two broad types: **structured** and **unstructured**. From an instrument design point of view, the structured questions pose the greater difficulties. From a content perspective, it may actually be more difficult to write good unstructured questions.

Varieties of Structured Questions

Dichotomous Questions

When a question has **two** possible responses, it can be considered **dichotomous.** Surveys often use dichotomous questions that ask for a Yes/No, True/False or Agree/ Disagree response. There are a variety of ways to lay these questions out on a questionnaire:

Do you believe that the penalty is justified for overdue books?

———Yes

———No

Questions Based on Level of Measurement

Questions can be classified in terms of their level of measurement. For instance, in case ofoccupation using a **nominal** question. Here, the number next to each response has no meaning except as a placeholder for that response. The choice of a "2" for a

Professor and a "1" for a **Student** is arbitrary — from the numbering system used it cannot be infer that a lawyer is "twice" something that a truck driver is.

Types of user:

1=Student

2=Professor

Respondents can be asked on the basis of rank order:

Rank the users in order of preference:

————M.Phil Students

————Doctoral research scholars

————Post Doctoral Research Scholars

Survey questions can be constructed that attempt to measure on an **interval** level. One of the most common of these types is the traditional 1-to-5 rating (or 1-to-7, or 1-to-9, etc.). This is sometimes referred to as a **Likert response scale**. Here an opinion question can be asked on a 1-to-5 bipolar scale (it's called bipolar because there is a neutral point and the two ends of the scale are at opposite positions of the opinion):

Penalty is justified for overdue books under some circumstances.

1. **Strongly disagree**
2. **Disagree**
3. **Neutral**
4. **Agree**
5. **Strongly agree**

8.4.1.1 Scaling

A variety of questions utilize a scale. Selltiz (1959) categorizes such scale as differential, summated and cumulative.

Differential Scale: Often referred to as Thurstone type scale, utilizes a series of statements or items with equidistances between them. Each item, therefore, receives a value point equidistant from those immediately preceding and following it. The respondent is asked to check each statement with which he or she agrees or the two or three statements that best represent his or her position.

Cumulative Scale: Consists of a series statements with which the respondent indicates agreement or disagreement. The items are related to each other. One "should" respond to subsequent manner similar to preceding ones. Cumulative scales represent the ordinal level of measurement. Eg. Guttman Scale.

Summated Scale: Consist of series of statements or items, but no effort is made to distribute them evenly along a continuum. Only items that are favourable or unfavourable to the attitude of interest are used (unless a "no opinion" type of response is included).Likert Scale is one of the most commonly used summative scales.

Likert Scale or summated scale:

Likert scale, named after Rensis Likert, who developed the method. The Likert scale is essentially a **multiple-indicator** or **multiple-item measure** of a set of attitudes relating to a particular area. The goal of the Likert scale is to measure intensity of feelings about the area in question. In its most common format, it comprises a series of statements (known as 'items') that focus on a certain issue or theme. Each respondent is then asked to indicate his or her level of agreement with the statement. Usually, the format for indicating level of agreement is a five-point scale going from 'strongly agree' to 'strongly disagree', but seven-point scale and other formats are used too. There is usually a middle position of 'neither agree nor disagree' or 'undecided' indicating neutrality on the issue. Each respondent's reply on each item is scored, and then the scores for each item are aggregated to form an overall score. Normally, since the scale measures intensity, the scoring is carried out so that a high level of intensity of feelings in connection with each indicator receives a high score (for example, on a five-point scale, a score of 5 for very strong positive feelings about an issue and a score of 1 for very negative feelings). The measure of commitment to work referred to in Research in focus 7.2 is an example of a Likert scale. Variations on the typical format of indicating degrees of agreement are scales referring to frequency (for example, 'never' through to 'always') and evaluation (for example, 'very poor' through to 'very good').

There are several points to bear in mind about the construction of a Likert scale. The following are particularly important.

- The items must be statements and not questions.
- The items must all relate to the same object (job, organization, ethnic groups, unemployment, sentencing of offenders, etc.).
- The items that make up the scale should be interrelated.

Another interval question uses an approach called the **semantic differential**. Here, an object is assessed by the respondent on a set of bipolar adjective pairs (using 5-point rating scale):

Please state your opinion on insurance of library materials:

Insert Table

	Very much	Somewhat	Neither	Neither	Very much	
Interesting						**Boring**
Simple						**Complex**
Uncaring						**Caring**
Useful						**Useless**

In a **cumulative or Guttman scale**.Here, the respondent checks each item with which they agree. The items themselves are constructed so that they are cumulative — if you agree to one, you probably agree to all of the ones above it in the list:

Please check each statement that you agree with:

—Are you willing to permit outsiders to use the library of your institution?

—Are you willing to permit outsiders to be the member of the institution?

——Arè you willing to permit outsiders to be the member of other libraries in your neighbourhood?

——Are you willing to have an outside person to be the user of neighbouring library?

Filter or Contingency Questions

Sometimes questions are asked to respondent one question in order to determine if they are qualified or experienced enough to answer a subsequent one. This requires using a **filter or contingency question**. For instance, if the respondent has ever read Oxford English Dictionary and a different question if he or she has not. In this case, a filter question have to be constructed to determine whether they've ever read Oxford English Dictionary:

Have you ever read Oxford English Dictionary?

☐ **Yes.**

☐ **No.**

If yes, about how many times you read Oxford English Dictionary?

☐ **Once**

☐ **2 to 5 times**

☐ **6 to 10 times**

☐ **More than ten times**

Question content:

- The questioner should ask him or herself if a specific question is actually necessary.
- If yes, how many questions are particularly needed for a topic.
- Question should be more specific, concrete, or related to the respondent's personal experiences.
- Misleading questions should be avoided.
- Each questions ask one question. A question containing more than one concept presents difficulties for subsequent analysis and interpretation.

8.3.2 Pre-Questionnaire Planning

Before preparing a questionnaire pre planning for designing the questionnaire is essential. The steps by which a brief out line can be drawn is as follows:

a. Define the problem (and purpose)

b. Consider previous, related research, the advice of experts, etc.

c. Hypothesize a solution to the problem (or at least identify research questions, the answers to which will shed some light on the problem).

d. Identify the information needed to test the hypothesis. This step should include deciding which aspects of the problem will be considered, and planning ahead to the presentation and analysis of the data. Deciding how the data will be organized, presented and analysed can significantly influence what types of data will have to be collected. It may be useful at this point to construct so called "dummy- tables," or tables presenting the important variables with hypothetical values, to help detect possible problems regarding presentation and analysis. As for example the hypothesis can be tested on information need of social scientists of research institute library.

e. Identify the potential respondents or subjects. Such as in a research institutes library potential respondents are research scholars. Here a question can be incorporated are these research scholars aware about the resources of that particular library.

f. Select the best or most appropriate technique for collecting the necessary data. Researcher should consider the relevant advantages and disadvantages of the questionnaire, interview, observation, and other techniques.

8.3.2.1 Advantages of the Questionnaire

There are several advantages of questionnaire as follows:

a. The questionnaire, especially mail questionnaire, tends to encourage frank answers.it Is easier for the researcher to guarantee anonymity for the respondent.The respondent can complete the questionnaire without the researcher's being present.

b. The characteristics of the questionnaire that help to produce frank answers also help to eliminate interviewer bias. Here it is meant that the style of verbal presentation cannot influence the response. Personal influence of the interviewer is avoided: sex, ethnic origin and perceived social status may all influence the accuracy of the interview, but can eliminated in the questionnaire

c. Fixed format of the questionnaire tends to eliminate variation in the questioning process. This does not rule out the possibility of respondents interpreting the same questions in different way.

d. The manner in which a mail questionnaire is distributed and responded to also allows it to be completed, within limits, at the leisure of the participants. This encourages well thought out, accurate answers.

e. Questionnaires can be constructed so that quantitative data are relatively easy to collect and analyze.

f. Questionnaire can facilitate the collection of large amounts of data in a relatively short period of time.

g. Questionnaire are usually relatively inexpensive to administer.

h. Properly constructed, questionnaire can also elicit a great deal of information on the survey topic,particularly if quantitative data only are sought, and relative costs of the survey are reduced.

8.3.2.2 Disadvantages of Questionnaire

There are few disadvantages of questionnaire also:

There are few disadvantages of questionnaire also:

a. Precludes personal contact with respondents, perhaps causing the investigator to gain insufficient knowledge about participant in a study
b. Does not allow respondent to qualify ambiguous questions
c. If the prepared instrument does not arouse respondent emotions (i.e. when the questionnaire is too impersonal), valid responses might not be elicited.
d. Poorly worded or direct questions might arouse antagonism or inhibitions on the part of respondents
e. Difficulty in obtaining responses from a representative cross section of target population
f. Because opinionated respondents might be more likely than other subjects to complete and return it ,use of a questionnaire might lead to nonresponse bias
g. Some potential respondents may be antagonistic toward mail surveys, regardless of the purpose or quality of the instrument distributed
h. Verification of the accuracy of questionnaire responses might sometimes be difficult, or even impossible
i. Uneducated subjects might not respond to a list of printed questions
j. Most questionnaires cannot be designed to uncover causes or reasons for respondents' attitudes, beliefs,or actions.

8.4 The Interview

Interview means "Conversation with a purpose". As a research technique the interview is conversation carried out with the definite purpose of collecting data on the problem under investigation by means of the spoken word through a person contact between an interviewer (researcher) and an interviewee (i.e. respondent). Interviews are akin to questionnaires. Like questionnaires interviews are also designed to collect valid and reliable data through the response of the respondent to a planned sequence of questions, but the difference lies in the fact that whereas in questionnaires the respondent reads the questions and replies in written words, in interviews the questions are read to the respondent who answers in spoken words which are recorded by the interviewer.

"The purpose of interviewing is to find out what is in or on someone else's mind."(Patton,1990)Also according to Patton(2002)"quality of information obtained during an interview is largely dependent on the interviewer"

8.4.1 Stages of Interview

The whole process or operation of interview takes place in the following three stages:

- Rapport Building

- Probing
- Recording

8.4.1.1 Rapport Building

It means the establishment of harmonious and effective contact between the interviewer and interviewee (respondent). This depends on the capacity of the interviewer,by dint of which he is able to motivate the interviewee to communicate.

8.4.1.2 Probing

It implies the asking of the right type of questions in the right manner and the right language.If the one of these goes wrong, the respondent cannot be expected to communicate to the desirable extent or depth. The question or question-complex, therefore, should be such and so put that it is capable of gathering all the relevant information.

8.4.1.3 Recording

Recording, that is,recording of response, may be accomplished in two ways:

(a) at the time of interview; and

(b) after the interview is over.

(a) **Recording at the time of interview**: assures accuracy, since all the data obtained from the respondent can be noted down then and there. But in practice situations may arise in which the respondent, on seeing that his statements or opinions are being noted, gets psychologically reserve and cautions and does not make himself free in his response.

Recording on tape is convenient, and it dispenses with the necessity of writing during the interview, which is a job that may be of a distracting influence both to the interviewer and the respondent. Interview recorded on tape may be replayed as often as necessary at a later time. In addition to words, the tone of voice and emotional impact of the response is preserved by the tape.

(b) **Recording of response after the interview is over:** impliesrecording based on memory, which presumably affects accuracy. This draw-back may be much reduced if recording is done immediately after the interview has been taken.

8.4.2 Types of Interview

Like questionnaires, interviews also are primarily of two types:

- **Unstructured (also called Non Directive, Uncontrolled, Unguided)**

This type of interview sometimes called 'semi-structured' interview, seeks to collect more detailed information from the interviewees, through questions, for the answers of which the respondents would require to speak at some length. In this type of interview questions put to the interviewees are smaller in number and of open-ended type. It is flexible in the method of questions, and the sequence of questions. The interviewer is more free to choose the form of question depending on the specific situation, the level and condition of the respondent. The wording and the sequence of the questions may change according to the response pattern. Because of its little insistence on form it is often called Informal interview.

The unstructured interview is suitable for in-depth studies. The objective is to obtain discursive and complex information, more of qualitative and subjective type than of objective and quantitative type, containing a high proportion of opinion, attitude and personal experience. Used at the early stage of an investigation. This interview helps the researcher in getting insight into the problem. Because of its flexibility, the investigator may enable himself to pursue a lead, which may throw new light on the problem and make it more meaningful and significant. It can yield by-products, that is, generate such data, or insight into a direction, as could suggest quite a new problem or topic for elaborate investigation at a later time.

But to pilot the unstructured interview the interviewer requires much skill, versatility and on the-spot adaptability, otherwise he/she will fall short of the necessary rapport and fail to get adequate information. The interviewer needs to put much effort into ensuring that he stays fairly close to the point and is not drawn into distracting side-issues, that he records the data in an objective and consistent manner and that his interview has covered all the facts or aspects of the problem.

• **Structured (also called Directive, Controlled, Guided)**

This kind of interview seeks to obtain specific or pinpointed data from the interviewees.With this end in view the interviewer pilots his interview with the help of an interview schedule which is essentially a questionnaire often to the last detail. Here the number and nature of questions, the order of putting them to the interviewee, wording of the questions, recording system and every other thing involved in the interview process is standardized. The response patterns are also standardized in the form such as Know/Don't know; Yes/No etc. Thus the questions are mostly close ended, where the alternative responses expected are given for the choice of the respondent. Because of the insistence of the standardized form, this method is called Formal Interview. This design obviously enables the interviewer easily to record the data while the interview takes place, the recording being mostly in the form of ticking or circling.

The structured interview is more methodical and easier to administer. It requires less efforts, less on the spot adaptability from the interviewer to establish rapport with the respondent. It is adopted to derive more precise generalizations. But it should be adopted when a good knowledge of the problem has been gained, so that the questions can be framed to function effectively and without the necessity of modifications. Therefore the structured interview is expedient at the later stage of the investigation, when a careful exploratory study of the problem has enabled the researcher to structure the field and to devise specific and adequate questions from which there may be little deviations.

The classification is based on the types of data required for the problem under investigation.

The above two types relate to structure. Irrespective of structure, interviews may be of the following types:

- **Group Interview/Individual interview**

Group Interview

When more than one person is interviewed at a time, the interview is called group interview. Usually in a group interview, the persons forming the are asked different questions pertaining to a theme; but sometimes they may be required to answer the same question so that different views of the same question are obtained.

The size of the group is a matter to be considered. According to C.V. Good,"The size of the group should not be so large that it is unwieldy or inhibits participation by most members and should not be so small that it lacks substantially greater coverage than in the individual interview. The optimum size is approximately 10 to12 persons, Social, intellectual and educational homogeneity are important for effective participation of all group members. A circular seating arrangement, with the interviewer as one of the group," is collective to full and spontaneous reporting and participation.

Individual Interview

When only one person is interviewed at a time, it is individual interview.

- **Focussed Interview**

It takes place with persons known to have been involved in a particular situation, and is focussed on the subjective experience-attitudes, opinions and emotional responses in respect of the particular situation under study.

- **Repeated Interview**

It implies interviewing the same respondent(s) more than once pertaining to the same subject. It makes an attempt to trace the specific development of a social or psychological process, that is observes the changes in actions or attitudes which determine a given behaviour pattern or social situation, e.g. for a study of how a voter may be subjected to repeated interview.

- **Depth Interview**

It is mostly applied in psycho-analytical study. It attempts to elicit data about the unconscious motivations as well as all other aspects of personality dynamics. It is a lengthy procedure and is designed to encourage free and uninhibited response. The major purpose of such interviews are diagnosis and treatment of an individual's problem. This type of interview is often designated as "clinical interview"

8.4.3 Conducting the Interview

- In conducting an interview, the interviewer should attempt to create a friendly nonthreatening atmosphere. Much as one does with a cover letter, the interviewer should give a brief, casual introduction to the study; stress the importance of the person's participation; and assure anonymity, or at least confidentiality, when possible. The interviewer should answer all legitimate questions about the nature of the study,and produce appropriate credentials upon request. Approach all respondents as individuals; assure them that their views are valuable and of significance to the survey being conducted.
- Researcher should set up the interview well in advance, and interviewer should appear for the interview promptly. Sending the list of questions in advance is inadvisable, it will reduce candidness. If the interview has been scheduled ahead of time,it is a good idea to confirm the date in writing and to send a reminder several days before the interview.
- Use of audio visual aids helps to improve the recording the responses. But for that permission of the respondent has to be taken.
- Interviewer should avoid asking more than one question at a time. The interviewer must be careful not to show surprise etc. as a result of any of the interviewee's responses. Such responses can bias future responses of the participant.
- The interviewee's responses should represent his or her thoughts alone not a combination of his or hers and interviewer's. To obtain as many responses as

possible, the interviewer must also learn how to deal with "don't knows". Often people need a certain amount of encouragement before they will response fully to particular question.There is a fine line,however,between encouraging a response and helping to word it or forcing an answer where there should not be one. Consequently, one should be conservative or cautious in encouraging a response when the interviewee seems reluctant to provide one.

- Interviewer should avoid directly dispute respondents, even though you may know or suspect that their replies are inaccurate.
- Arguing with respondents or condemn their views should be avoided.
- Interviewer should be neutral in recording responses so that the collected data are accurate and objective.
- Interviewer should express gratitude to the respondent for their cooperation and assistance upon completion of interview session.

8.4.4 Bias in Interview

Bias presents a real threat to the validity of interviews. Particularly critical is the bias that may be introduced by the interviewer. As was indicated earlier, some interviewer bias can be avoided by ensuring that the interviewer does not overreact to responses of the interviewee. Other steps that can be taken to help avoid or reduce interviewer bias include:

- Having the interviewer dress inconspicuously and/or appropriately for the environment;
- Holding the interview in a private setting;
- Keeping the interview as informal as possible

8.4.5 Advantages of Interview

- Interview produces better response rate. Here sample of persons actually participating in the study tends to represent a large percentage of the original sample, and is therefore more representative of the population than would be a sample representing a relatively low response. Apparently the personal contact of the interview helps to encourage, or put more pressure on, persons to fully respond. Consequently, it is also possible to employ interview schedules of greater length than comparable questionnaires, without jeopardizing a satisfactory response rate.
- Personal contact provides a greater capacity than the mail questionnaire for the correction of misunderstandings by participants.
- It is generally believed that the interview is better at revealing information that is complex and/or emotionally laden. The use of visual aids can sometimes facilitate presenting and/or recording complicated information.

8.5 Case Study

The case study is a way of organizing social data for the purpose of viewing social reality. It examines a social unit as a whole. The unit may be a person, a family, a social group, a social institution, or a community. The purpose is to understand the life cycle or an important part of the life cycle of the unit. The case study probes deeply and analyses interactions between the factors that explain present status or that influence change or growth. It is a longitudinal approach,showing development over a period of time.

8.5.1 Examples of Case Studies

Some examples of case study are given below:

- Knowledge River: A Case Study of a Library and Information Science Program Focusing on Latino and Native American Perspectives /Patricia Montiel-Overall and Sandra Littletree

  This article discusses the development of Knowledge River, a program at the University of Arizona School of Information Resources and Library Science established through several Institute of Museum and Library Services grants designed to recruit Latino and Native American students to the library and information science (LIS) profession. Knowledge River (KR) was designed as a national model for increasing diversity in information organizations and LIS programs. The article describes the KR model and elements of the program that have increased its success. Included are participation in a residential cohort, real-world library work experiences, and formal mentoring by KR graduates and other ethnic minorities in the field. Knowledge River has served as a catalyst for increasing awareness of diversity issues and multiple perspectives in addressing issues in the LIS field. Knowledge River has also resulted in a requirement that all LIS students enroll in at least one diversity course. This article also provides a retrospective analysis of the KR model and presents a theoretical framework for developing future LIS diversity programs such as KR.
- A Case Study of Periodical Use by Library and Information Science Students/ Tammy Ivins

  There is a lack of information in the literature about the sources used for research by modern Master of Library and Information Science students in the United States, and so the objective of this project is to understand the use of periodical articles by these students. Specifically: do articles play a major role in student research, how current are the articles cited, and can a core group of periodicals be identified? 192 capstone papers from 2005–2010 at the University of North Carolina at Chapel Hill were sampled for a bibliographic analysis. The results show that periodical articles do play a significant role (making up 48% of all references) and are fairly current (49% of all article citations were from within five years). This study identified four core LIS periodicals among its results: The Journal of Academic Librarianship, College & Research Libraries, Library Journal, and Communications of the ACM. Finally, 85% of all periodicals were cited by only one student, indicating that MLIS students use a broad variety of periodicals for their research.

8.6 Government Agencies Engaged in Conducting Survey

8.6.1 Census of India

The Indian Census is the largest single source of a variety of statistical information on different characteristics of the people of India. With a history of more than 130 years, this reliable, time tested exercise has been bringing out a veritable wealth of statistics every 10 years, beginning from 1872 when the first census was conducted in India non-synchronously in different parts. To scholars and researchers in demography, economics, anthropology, sociology, statistics and many other disciplines, the Indian Census has been a fascinating source of data. The rich diversity of the people of India is truly brought out by the decennial census which has become one of the tools to understand and study India.

The responsibility of conducting the decennial Census rests with the Office of the Registrar General and Census Commissioner, India under Ministry of Home Affairs, Government of India. It may be of historical interest that though the population census of India is a major administrative function; the Census Organisation was set up on an ad-hoc basis for each Census till the 1951 Census. The Census Act was enacted in 1948 to provide for the scheme of conducting population census with duties and responsibilities of census officers. The Government of India decided in May 1949 to initiate steps for developing systematic collection of statistics on the size of population, its growth, etc., and established an organisation in the Ministry of Home Affairs under Registrar General and ex-Officio Census Commissioner, India. This organisation was made responsible for generating data on population statistics including Vital Statistics and Census. Later, this office was also entrusted with the responsibility of implementation of Registration of Births and Deaths Act, 1969 in the country.

8.6.2 National Sample Survey Organisation

The National Sample Survey (NSS) which came into existence in the year 1950, is a multi-subject integrated continuing sample survey programme launched for collection of data on the various aspects of the Indian economy required by different agencies of the Government, both Central and States. Although, at the beginning, the NSS started with the objective of collecting data for the construction of national accounts and its area of operation was kept restricted only to the rural areas of the country during the two rounds, it gradually expanded its geographical coverage and the scope of its enquires to cover, by and large,all the important socio-economic aspects influencing the life of the population in rural as well as urban areas.

8.6.3 National family Health Survey

The National Family Health Survey (NFHS) is a large-scale, multi-round survey conducted in a representative sample of households throughout India. Three rounds of the survey have been conducted since the first survey in 1992-93. The survey provides state and national information for India on fertility, infant and child mortality, the practice of family planning, maternal and child health, reproductive health, nutrition, anaemia, utilization and quality of health and family planning services.

Each successive round of the NFHS has had two specific goals: a) to provide essential data on health and family welfare needed by the Ministry of Health and Family Welfare and other agencies for policy and programme purposes, and b) to provide information on important emerging health and family welfare issues. The Ministry of Health and Family Welfare (MOHFW),Government of India, designated the International Institute for PopulationSciences(IIPS)

Mumbai, as the nodal agency, responsible for providing coordination and technical guidance for the survey. IIPS collaborated with a number of Field Organizations (FO) for survey implementation. Each FO was responsible for conducting survey activities in one or more states covered by the NFHS.

Technical assistance for the NFHS was provided mainly by ORC Macro (USA) and other organizations on specific issues. The funding for different rounds of NFHS has been provided by USAID, DFID, the Bill and Melinda Gates Foundation, UNICEF, UNFPA, and MOHFW, GOI.

8.6.4 List of Surveys Done by Different Agencies of Government of India

- **Directorate of Economics & Statistics, D/O Agriculture and Cooperation, Ministry of Agriculture, New Delhi**
- Improvement of Agricultural Statistics
  (a) Estimation of area of principal agricultural crops under Timely Reporting Scheme

(b) Estimation of area and production of principal agricultural crops under Establishment of an Agency for Reporting of Agricultural Statistics
(c) Crop Estimation Survey on Fruits & Vegetables

- Improvement of Crop Statistics
- The Comprehensive Scheme for Studying the Cost of Cultivation of Principal Crops in India ?
- Agriculture Census 2010-11 and
- Input Survey 2011-12 By Ministry of Agriculture
- **Labour Bureau, Ministry of Labour**
- Rural Labour Enquiry
- Working Class Family Income & Expenditure Survey
- Occupational Wage Surveys
- Socio-economic conditions of Different Segments of Labour.
- Annual Survey of Industries
- Survey of Labour Conditions
- Contract Labour Survey
- Annual Survey of Industries
- **Reserve Bank of India**
- Industrial Outlook Survey
- Order Books Inventory and Capacity Utilisation Survey
- Credit Condition Survey
- Inflation Expectations Survey of Households
- Consumer Confidence Survey
- Survey of Professional Forecasters
- Basic Statistical return (BSR)-4
- Composition and Ownership Pattern of Deposits with Scheduled Commercial Banks (SCBs)
- BSR-6: Debits to Deposit and Credit Accounts with Scheduled Commercial Banks ? Survey of Small Borrowal Accounts

Miscellaneous Surveys conducted by Central Government Offices

- Sample Registration System(SRS)
- National Family Health Survey(NFHS)
- Annual Health Survey
- Socio Economic Surveys
- Annual Survey Of Industries
- Agriculture Surveys

8.6.5 Sample questionnaire

QUESTIONNAIRE-1

FOR LIBRARIAN

1. **General Information**:

(a) Name of the University : ____________________

(b) Year of Establishment : ____________________

(c) Address : ____________________

____

Phone__________ Fax__________

E-mail__________ Website__________

(d) Name of the University Library : ____________________

(e) Name of the Librarian/Incharge : ____________________

(f) Year of Establishment : ____________________

(g) Library Hours : ____________________

(h) Address : ____________________

____

Phone__________ Fax__________

E-mail__________ Website__________

(i) Total **Collections** (numbers in year wise)

Sr. No.	Items	2002-03	2003-04	2004-05	2005-06	2006-07
1	Books					
2	Journals (current)					
3	Journals (bound vol.)					
4	Ph.D. Theses					
5	Manuscripts					
6	Rare Books					
7	Audio Cassettes					
8	CD-ROMs/Floppies/Database					
9	Microfilms/Microfiches/Slides					
10	Magnetic Tape/Films					
11	Others (please specify)					

(j) Total **Staff** (numbers in year wise)

Sr. No.	Staff	2002-03	2003-04	2004-05	2005-06	2006-07
1	Librarian					
2	Deputy Librarian					
3	Assistant Librarian					
4	Doc. Officer/Inf. Officer					
5	Sr. Prof. Assistant					
6	Cataloguer					
7	Professional Assistant					
8	Library Assistant					
9	LDC/DEO/Clerk					
10	Library Attendant					
11	Janitor					
12	Book Lifter					
13	Peon					
14	Mender/Binder					
15	Others					

(k)Total enrolled **Users/Members** of Library (numbers in year wise)

Sr. No.	Users	2002-03	2003-04	2004-05	2005-06	2006-07
1	Faculties					
2	Research Scholars					
3	PG Students					
4	UG Students					
5	Admin. Staff					
6	Supporting Staff					
7	Outsiders					

(l) Annual **Budget** allocated to the Library (Rs. in Lakh in year wise)

Sr. No.	Users	2002-03	2003-04	2004-05	2005-06	2006-07
1	Books					
2	Journals					
3	Non Book Materials					
4	Accessories					
5	Other (please specify)					

(m) Is there any provision for any separate budget for e-resources? Yes/No

If yes, please give the % of total budget: ________________ (n)

Library Automation: Are you using any library automation package? Yes/No

If yes, please mention:

(i) Name of the automation package: ________________

(ii) Date of implementation of automation package: ______________

(iii) Do you have any server to run the library automation package? Yes/No

If yes, please give the details of Server: ________________

2. Networking

(a) Do you have LAN Facilities? Yes/No

(b) Do you have Internet facilities: Yes/No

If yes, please tick the Mode of Connectivity:

(i) Dial-up Yes/No

(ii) Leased Line Yes/No

(iii) BVSAT Yes/No

(iv) V-SAT Yes/No

If yes, please give the details of V-SAT:

V-SAT	Capacity	MB	Bandwidth

3. e-Resource Section:

(a) Do you have any separate section for e-Resources? Yes/No

If yes, please mention how many computers are available to access the e Resources for users?

PCs	Pentium IV	Pentium III	Pentium II	Any other

How many printers are available in e-Resource section?

Printers	Laser Jet	Ink Jet	Desk Jet	DMP

(b)

(c) Please tick the other Hardware Devices those are available in e-Resource section:

(i) Bar Coding Instrument (ii) Digital Camera

(iii) Scanner (iv) Web Camera

(v) Other (please specify) ____________________

(d) Please tick the Storage Medium those are the using by the users in eResource section:

(i) CD ROM	Yes/No	(ii) PEN DRIVE	Yes/No
(iii) ZIP Disk	Yes/No	(iv) CTD	Yes/No
(v) DAT	Yes/No	(vi) DVD	Yes/No

(e) How many staffs are working in this section?

(i) No. of Professionals with computer knowledge __________

(ii) No. of Professionals without computer knowledge __________

(iii) No. of Non-Professionals __________

(f) Are you using any software for e-Resource Section? Yes/No If yes, please mention which software is available in your e-Resource Section:

(i) DSpace (ii) Greenstone (iii) CDSware

(iv) Bepress (v) Eprints (vi) Fedora

(vii) Kepler (viii) I-Tor (ix) MPG eDoc

(x) MyCoRe (xi) Ingenta (xii) Ebrary

(xiii) Digital Library Project

(xiv) Other (please specify) ____________________

4. Status of UGC-INFONET Programme
Are you a member of UGC-INFONET Programme? Yes / No
If yes, please give the details:
(a) Date of Implementation: ______________________________
(b) How many e-Resources are you accessing (in number): ____________
(i) e-Journals ________________(ii) Full Text Database ________
(iii)Bibliographical Database______(iv) Other (please specify)______

5. Connectivity & Usage of National/International Consortia
Is your library a member of Library Consortia for e-Resources other than UGC-INFONET? Yes / No. If yes, please give the status of under mentioned:

Name of the Organiza-tion	Starting year of Member-ship	Yearly Subscr-iption in Rs.	Average No. of Request/ Month	Retrieval proportion Against Request	No. of online. Journals Sub	Usage of online journals in %
INDEST						
DELNET						
NISCAIR						
EDUSAT						
ERNET						
OTHER						

6. Details of e-Resources
(a)Collection of e-Resources (in numbers):

Sl. No.	e-Resources	2002-03	2003-04	2004-05	2005-06	2006-07
1	CD-ROM Titles					
2	e-Database					
3	e-Journals					
4	e-Reports					
5	e-Content pages					
6	e-Clippings					
7	e-Books					
8	Others					

(b) Mode of e-Resources: Do you provide the e-Resource services through?

(i)	Internet via their website	Yes/No
(ii)	CD-ROM Network	Yes/No
(iii)	Commercial Online service Vendors	Yes/No
(iv)	Others (please specify)	Yes/No

(c) Which types of e-Resources are subscribing for e-Resource section, please tick the under mentioned:

(i) Full text	Yes/No	(ii) Abstract	Yes/No
(iii) Bibliographic	Yes/No	(iv) Numeric	Yes/No
(v) Graphic	Yes/No	(vi) Others (please specify)_____	

(d) Criteria for selecting of e-Resources:

a.	Quantity to meet user need	Yes/No
b.	Subject relevance	Yes/No
c.	Cost effectiveness	Yes/No
d.	Authenticity of information	Yes/No
e.	After sale maintenance	Yes/No
f.	Currency of information	Yes/No
g.	Back Issues Facility	Yes/No
h.	Distributed access	
i.	Period of Access	Yes/No
j.	Added Value	Yes/No
k.	Ease of accessibility	Yes/No
l.	Legal issues	Yes/No
m.	Preservation	Yes/No
n.	Vendor reliability	Yes/No
o.	Period of Access	Yes/No

(e) Name of e-Resources those are accessing by the users among the full text and bibliographic in your library:

Full-text e-Resources

(i)	Emerald	Yes/No
(ii)	Citation Index	Yes/No
(iii)	Elsevier s Science Direct	Yes/No

(iv)	Bio MedNet Reviews	Yes/No
(v)	Pub Med, Math Science	Yes/No
(vi)	INIS, IBP	Yes/No
(vii)	UNDP Human Development Report	Yes/No
(viii)	World Development Report	Yes/No
(ix)	Human Development Database	Yes/No
(x)	CRISNIFAC, Vans	Yes/No
(xi)	AGRIS, Prowess, Copex, Socio file	Yes/No
(xii)	IEEE/IEE Electronic Library Online (IEE)	Yes/No
(xiii)	Project Mouse Journals	Yes/No
(xiv)	American Chemical Society	Yes/No
(xv)	Annual Reviews	Yes/No
(xvi)	American Physical Society	Yes/No
(xvii)	American Institute of Physics	Yes/No
(xviii)	Biological Abstracts	Yes/No
(xix)	Cambridge University Press Journals	Yes/No
(xx)	Institute of Physics	Yes/No
(xxi)	JSTOR	Yes/No
(xxii)	Nature	Yes/No
(xxiii)	Royal Society of Chemistry	Yes/No
(xxiv)	Science Online	Yes/No
(xxv)	INSIGHT	Yes/No
(xxvi)	OCLC	Yes/No
(xxvii)	Oxford University Press	Yes/No
(xxviii)	Taylor & Francis	Yes/No
(xxix)	Springer Verlag s Link	Yes/No
(xxx)	ABI/Inform Complete	Yes/No
(xxxi)	ACM Digital Library	Yes/No
(xxxii)	ASCE Journals	Yes/No
(xxxiii)	ASME Journals (+A M R)	Yes/No
(xxxiv)	CRIS INFAC Ind. Inf.	Yes/No

(xxxv) EBSCO Database Yes/No

(xxxvi) Euromonitor (GMID)

Bibliographic Database

(i) COMPENDEX on EI Village Yes/No

(ii) INSPEC on EI Village Yes/No

(iii) JET Yes/No

(iv) MathSciNet Yes/No

(v) SciFinder Scholar Yes/No

(vi) Web of Science Yes/No

(vii) J-Gate Custom Content for Consortia (JCCC) Yes/No

(f) What types of criteria are you following for selection of e-Resources?

(i) User recommendation Yes/No

(ii) Scanning catalogue Yes/No

(iii) Surfing e-information research website Yes/No

(iv) List services Yes/No

(v) Consulting other libraries Yes/No

(vi) News group Yes/No

(vii) Free Online Trial Access Yes/No

(viii) Recommendation of Faculty members, Research scholars and students Yes/No

(g) Do you following the under mentioned to evaluate the e-Resources for purchase?

(i) Trial before use Yes/No

(i) Performance Yes/No

(ii) Product review by expert Yes/No

(iii) Cost Yes/No

(iv) Coverage Yes/No

(v) Scope Yes/No

(vi) Limitations Yes/No

(vii) Time lag Yes/No

(viii) Access facility Yes/No

(ix) Single Users/Networking Yes/No

(x) Evaluation of content & performing under existing environment Yes/No

(h) Do you collect the users opinions to evaluate the subscribed e-Resources to take the decision __________ to continue or to cancel the subscription to these resources? Yes/No

Do you have Collection Development Policies for e-Resources? Yes/No If yes, please tick the Elements of Collection Development Policies for eResources those are following by your library:

(i) Short and Long Term Objectives Yes/No

(ii) Selection Responsibility Yes/No

(iii) Need Assessment & Users requirement Yes/No

(iv) Selection responsibility Yes/No

(v) Levels of Collections Yes/No

(vi) Selection Criteria

(vii) Acquisition Procedures Yes/No

(viii) Security, Authentication Yes/No

(ix) Coordination of Libraries Resources Yes/No

(x) Balance Between Print & e-Collections Yes/No

(xi) Others (Please specify) ________________________________

(j) Collection Development Policies includes:

Outline the Present Collection s Strength & Weakness:

(i) Make the Policy Decisions involved in Purchasing Activities [ ]

(ii) Policy should be open to changes & constantly involved & update [ ]

(k) Cataloguing of e-Resources

Do you catalogue the e-Resources? Yes/No. If yes, please mention:

(i) Subscribed e-Resources [ ]

(ii) Free online e-Resources [ ]

(iii) CD-ROMs [ ]

(iv) Others (Please specify): __________________________

(l) Library has own website: Yes/No. If yes, please give the URL _________

(m) Preservation of e-Resources in: Do you preserver the e-Resources? Yes/No If yes, please tick the format which you prefer for preservation:

(i) PDF Format [ ]

(ii) HTML Format [ ]

(iii) Others (please specify) ______________________________

(n) Are you following the Licensing Agreements? Yes/No If yes, please tick the under mentioned:

(i)	Authorized users will use the resources	Yes/No
(ii)	Cost of Access	Yes/No
(iii)	Access	Yes/No
(iv)	Archival Backup	Yes/No
(v)	Fair Use	Yes/No
(vi)	Confidentiality	Yes/No
(vii)	Multiple Format	Yes/No
(viii)	Negotiations	Yes/No
(ix)	Electronic Links	Yes/No
(x)	Inter Library Loan	Yes/No
(xi)	Protection of Increase of Price	Yes/No
(xii)	Dispute Resolution	Yes/No
(xiii)	Download Records for In-house Database	Yes/No

(xiv) Others(please specify) ______________________________

(o) Wedding policy of e-Resources :

(i)	Longer Availability of e-Resources	Yes/No
(ii)	Difficult to maintain the e-Resources	Yes/No
(iii)	More Comprehensive Coverage is offering by other	Yes/No

(iv) Collection of users opinion to evaluate the subscribed e-Resources to take the decision either to continue or to cancel the subscription to these resources Yes/No

(v) Is your library continue to flew weeding out of the e-Resources?

7. Users of e-Resources

(a) How many hours remains open the e-Resource section to access the eResources for users?

(i) Below 6 hours [ ]

(ii) 6-8 hours [ ]

(iii) 8-10 hours [ ]

(iv) Above 10 hours [ ]

(b) Please mention the % of opening hours of e-Resource section with general library hours ______________________

(c) Please mention the total attendance of users per day using the e-Resource section ______________________

(d) Please mention the % of e-Resource users with the normal library users ______________________

(e) Do you provide the printing facility to the users for search materials? Yes/No. If yes, please mention the charges for printout per page in paisa: ______________________

(f) e-Resources are using by per user (in hours)

S.No.	Users	Per Month	Per Year
1	Faculties		
2	Research Scholars		
3	PG Students		
4	UG Students		
5	Administrative Staff		
		Total	**Total**

(g) Are you charging any cost to use the e-Resources for users? Yes/No. If yes, please mention the charges per hour/per month according to user wise in Rs.

S.No.	Users	Per Hour	Per Month
1	Faculties		
2	Research Scholars		
3	PG Students		
4	UG Students		
5	Administrative Staff		
		Total	**Total**

(g) Please specify the user level of satisfaction:

S.No.	Items Used	Relevance in %	Highly Satisfied	Satisfied	Neutral	Dissatisfied	Highly Dissatisfied
1	Biblio.						
2	Abstract						
3	Full Text						
4	Graphic						

(h) Are the users of e-Resources less satisfied? Yes / No. If no, please tick the reasons:

(i) Less Opening Time [ ]

(ii) Charges to access e-Resources [ ]

(iii) Lack of proper guidance [ ]

(iv) Lack of proper e-Resources [ ]

(v) Lack of Printing facilities [ ]

(vi) Insufficient e-Resources [ ]

(vii) Technical problems [ ]

(viii) Lack of portability in contrast with original print materials [ ]

(ix) Failure of Hardware & Software affect the functioning of e-Resources section [ ]

(x) Lack of knowledge about tools & tech. used for searching and retrieving of e-Resources [ ]

Suggestions if any: ______________________________

____

Name of the Respondent: ______________ Designation __________

Signature with Seal ______________________________

Thank you for your cooperation

FOR USERS (STUDENTS, RESEARCHERS AND FACULTY MEMBERS)

Personal Information:

Name:	First ___________	Middle ___________	Last ___________
Sex:	Male [ ]	Female [ ]	
Age Group (years):	Below 20 [ ]	21-25 [ ]	26-30 [ ]
	31-35 [ ]	Above 35 [ ]	
Status:	Teacher [ ]	Researcher [ ]	PG Student [ ]
	UG Student [ ]	Other [ ]	

Name of the University: ______________________________

Department/Branch: ______________________ Session ___________

Use of e-Resources

1. **Experience of using the e-Resource section**: How long have you been using the e-Resources section?
   Less than 6 months [ ]   6 months-1 year [ ]   1-2 years [ ]
   2-4 years [ ]   More than 4 years [ ]
2. **Frequency of e-Resources section**: How often do you use e-Resource section?
   Daily [ ]   2-3 times a week [ ]
   2-3 times a month [ ]   Once in a month [ ]
3. **Time spent in the e-Resource section**: How many hours you spend in a week to use e-Resources?
   Less than 1 hr a week [ ]   2-3 hrs a week [ ]   5-6 hrs a week [ ]
   7-9 hrs a week [ ]   10-20 hrs a week   Over 20 hrs a week[ ]
4. **Most frequently used location of e-Resource**: From which place do you most frequently use for e-Resource?
   At university library [ ]   Other place in university [ ]
   At home [ ]   At other place [ ]
5. **Methods of Learning to use of e-Resources:** How did you learn to handle the eResources?
   Training from university library [ ]   Guidance from colleagues and friends [ ]
   Self instruction [ ]   External courses [ ]   Any other [ ]
6. **Purpose of using e-Resources**: The purpose(s) you mainly use the e-Resources for?
   Research [ ]   Education [ ]   Any other [ ]

7. **Problems Encountered while using the e-Resources**: What troubles face you mostly to use the e-Resources?
Slow access speed [ ] [ ]
Difficulty in finding relevant information [ ]
Overload of information on the Internet [ ]
It takes too long to view/download pages [ ]
Privacy problem [ ]
Any other please specify [ ]

8. **Ways to Browse Information from the e-Resource Section**: How do you browse the required information from the e-Resources section?Type the Web address directly [ ] Use search engines [ ]
Use subscribed e-Resources [ ] Any other please specify ________

8. **Collection of the search matter**: Do you collect the search matter from eResources? Yes / No. If yes, please mention through:
Pen Drive [ ] CD [ ] Floppy [ ]
Any other please specify____

10. **Satisfaction with e-Resource Facilities:** Upto what extent, are you satisfied with the e-Resource facilities provided by the university library?
Fully [ ] Partially [ ] Lease satisfied [ ] No comments [ ]

11. **Please specify your satisfaction level**: How much are you satisfied with the:

S.No.	Items Used	Relevance in %	Highly Satisfied	Satisfied	Neutral	Dissatisfied	Highly Dissatisfied
1	Biblio.						
2	Abstract						
3	Full Text						
4	Graphic						

12. **Satisfaction with e-Resource Section**: Are you satisfied with e-Resources section? Yes / No. If no, please tick the reasons:
(i) Less opening Time Yes / No
(ii) Charges to access e-Resources Yes / No
(iii) Lack of paper guidance Yes / No
(iv) Lack of proper e-Resources Yes / No
(v) Lack of Printing facilities Yes / No
(vi) In-sufficient e-Resources Yes / No
(vii) Technical problems Yes / No

(viii) Lack of probability in contrast with original print materials: Yes / No

(ix) Failure of Hardware and Software affect the functioning of e-Resources section: Yes / No

13. **Comparison of Conventional Documents and e-Resources**: In your opinion, using e-Resources as compared to use of conventional document is:

(i)	Time saving	[ ]	or Time consuming	[ ]
(ii)	More informative	[ ]	or Less informative	[ ]
(iii)	More expensive	[ ]	or Less expensive	[ ]
(iv)	Easy to use	[ ]	or Complicated	[ ]
(v)	More preferred	[ ]	or Less preferred	[ ]
(vi)	More flexible	[ ]	or Less flexible	[ ]
(vii)	easy to handle	[ ]	or Complicated	[ ]
(viii)	More effective	[ ]	or Less effective	[ ]

14. **Influences on Academic Efficiency?** How the use of e-Resources has influenced your academic efficiency?

Use of conventional documents has increased	[ ]
Dependency on the e-Resources has increased	[ ]
Expedited the research process	[ ]
Improved professional competence	[ ]

15. **Printing Facility**: Do you prefer to take the printout for search materials from the library? Yes / No.

Any suggestions: ______________________________

Signature

Thank you for your time and completing this questionnaire.

8.7 Summary

In conducting research, the researcher has several data collection techniques at his or her disposal. Three of most commonly used techniques- the observation, the questionnaire, the interview has been considered in this unit.

Observational methods are best suited for describing and understanding behaviour as it occurs. They are less effective for gathering information about a person's perceptions, beliefs, attitudes, etc.

Questionnaires and interviews are frequently used for obtaining the latter type of information. With these techniques, heavy reliance is placed on the respondent's report for information;normally, the investigator has not observed the events in question. This raises the issue.of the validity of verbal reports in particular, but the question of validity(and reliability) is important in deciding upon any technique of data collection.

Throughout the design process the researcher must be particularly careful not to introduce bias into the study. Even sequencing of questions.

In developing any of the three data collection techniques, the researcher also must make decisions regarding the specific types of questions to be used or the types of behaviour to be observed. These decisions are greatly affected by the kinds of information needed and whether the study is exploratory in nature.

Finally, in selecting a specific technique or instrument, the researcher should be aware of the various pros and cons of each method. For example, if one were particularly anxious to achieve a high response rate, he or she might choose the interview over the questionnaire. But if cost were the major concern, then the questionnaire would be obvious choice, other considerations being roughly equal.No one method is likely to be perfect for a given situation, but it should be possible to select one technique as the best alternative, given the objectives, subject, priorities, and limitations of the investigation.

8.7.1 Questions /SelfAssessment Questions

1.What are different types of observation methods? Discuss briefly.

2. What is scaling? Discuss its different types?

3.What are different stages of interview? Discuss.

8.8 References/Bibliography/Select Reading

Bibliography

American Library Association. (1966). *Library statistics:a handbook of concepts,definitions,and terminology.* Chicago: American Library Association.

Bryman, A. (2012). *Social research methods.* Oxford: Oxford University Press.

Busha, C. H., & Harter, S. P. (1980). *Research methods in librarianship:techniques and interpretation.* Orlando: Academic Press.

Dooley, D. (2004). *Social research methods* . New Delhi: Prentice Hall.

Good, C. (1966). *Essentials of educational Research.*

Hafner, A. W. (1989). *Descriptive statistical techniques for librarians.* Chicago: American Library Association.

J.L.Sardana, & Sehgal, R. (n.d.). *Statistical methods for librarians.* New Delhi: ESS ESS.

Ohdedar, A. (1993). *Research methodology.* Calcutta: Bengal Library Association.

Patton, M. (1990). *Qualitative evaluation and research methods.* Thousand Oaks,CA: Sage.

Patton, M. (2002). *Qualitative evaluation and research method.* Thousand Oaks,CA: Sage.

Powell, R. R. (1985). *Basic reserach methods for librarians.* New Jersey: Ablex Publishing Corporation.

Sehgal, R. (1998). *Applied statistics for library science research* (Vol. 2v). New Delhi.

Simpson, I. S. (1983). *Basic statistics for librarians.* London: Clive Bingley.

Slater, M. (1990). *Research methods in library and information studies.* (M. Slater, Ed.) London: The Library Association.

Trochim, W. M. (2016, JULY 15/07/2016). *Types of question.* Retrieved from Research methods knowledge base: http://www.socialresearchmethods.net/kb/questype.php

Unit 9 ❑ Presentation of Data : Techniques

Structure

9.1 Learning Objectives

After reading this unit you will be able to learn

1. The concept of variable and how it differs from constant
2. What characteristics a variable should possess
3. The categories of variable and there interrelationships with reference to statistical analysis
4. Different categories of statistical analysis
5. Types of data analysis
6. The precise steps of data organisation
7. An overview of different aspects of frequency distribution

9.2 Introduction

The raw data collected through research investigation cannot be used as it is for further analysis to reach the findings and concrete conclusion. These are to be organised and presented properly to help statistical analysis. Again, the findings from statistical analysis should also be presented appropriately. This chapter gives some basic concepts related to data collection, organisation and analysis, an overview of the steps of data organisation and frequency distribution.

9.1. Basic Concepts

9.1.1. Understanding Statistical Data and Variables

After conducting an investigation, the researcher has handful of raw data collected from his sample population. These data are collected through a sequence of observations, based on a set of objectives of the research work. Such data are collected, particularly in case of social science research, from the sample drawn from population and is

[1] http://www.businessdictionary.com

known as statistical data. Data are the values (measurements or observations) that the variables can assume.

9.1.1.1. What is a variable? How does it differ from a constant?

A variable is a data item which may change in value. Anything that can vary can be considered a variable. According to Business Dictionary[1],

"a variable is a characteristic, number, or quantity that increases or decreases over time, or takes different values in different situations." Variables always bear some values, may be quantitative/ numerical and/or text values. For instance, age can be considered a variable because age can take different quantitative or numerical values for different people. Similarly, country can be considered a variable because a country can hold text values by different country-names.

On the contrary, a *constant* is a number or symbol that has a fixed and unchanged value. **Variables differ from constants in the ways as stated below:**

- Variables are quantities with changing magnitude, hence can assume different values based on the application; Constants are quantities with unchanging values, and used to represent numbers with significance.
- Constants are used to represent quantities in nature which are fixed, and the variables are used to represent unknowns.

9.1.1.2. Characteristics of a variable

A variable has two distinguished characteristics:

- A variable is an attribute that describes a person, place, thing, or idea.
- The value of the variable can "vary" from one entity to another.

Now, what is attribute?

Attribute: An attribute is a specific value of a variable. For instance, the variable sex or gender has two attributes: male and female. Or, the variable agreement might be defined as having five attributes denoting different levels of gradual variation:

- 1 = strongly disagree
- 2 = disagree
- 3 = neutral
- 4 = agree
- 5 = strongly agree

In addition, a variable should possess the following characteristics.

- Each variable should be exhaustive; it should include all possible answerable responses. For instance, if the variable is "religion" and the only options of attributes are "Hinduism", "Jewish", and "Muslim", there are quite a few religions that may have not been included. The list does not exhaust all possibilities. The way to deal with this is to explicitly list the most common attributes and then use a general category like "Other" to account for all remaining ones.

- The attributes of a variable should be mutually exclusive, no respondent should be able to have two attributes simultaneously. While this might seem obvious, it is often rather tricky in practice. For instance, you might be tempted to represent the variable "Employment Status" with the two attributes "employed" and "unemployed." But these attributes are not necessarily mutually exclusive — a person who is looking for a second job while employed would be able to check both attributes! But don't we often use questions on surveys that ask the respondent to "check all that apply" and then list a series of categories? Yes, we do, but technically speaking, each of the categories in a question like that is its own variable and is treated dichotomously as either "checked" or "unchecked", attributes that are mutually exclusive.

9.1.1.3. Types of statistical data and/variables

Before any statistical analyses or calculations of data, it is required to decide what type of data one is dealing with. There are a number of typologies. But for the convenience to understand these are broadly outlined below.

9.1.1.3.1. Qualitative and Quantitative data

On the basis of the character or nature, the data collected to investigate a research question may be of two types-

Qualitative data: which cannot be measured numerically and it is also called descriptive/ categorical data. These can only be organised under suitable categories for further descriptive and analytical study.e.g. name, religion, marital status, socioeconomic status, awareness. Such data cannot be expressed in numerical forms.

Ex. nature of job of the user community of a library.

Quantitative data: A variable takes different 'values' which can be measured numerically in suitable units. Ex. different ages of users of a public library. A variable whose values depend on chance and cannot be predicted is called random variables. A random variable is a function that associates a unique numerical value with every outcome of an experiment. The value of the random variable will vary from trial to trial as the experiment is repeated.

9.1.1.3.2. Continuous and Discrete or Discontinuous Variables

Though variables can be numerical or textual, for statistical computation only the data with quantitative character are considered and these technically are called the 'random variables' or 'variates', or usually are considered as 'variables' in practice. These are of two types,

a) Continuous: A variable is called continuous when it can take any value within a specified interval. Ex age of a library user may be any value in between a range, say, 35 years or 38 years or 56 years, or it may be 35.5 years or 42.25 years etc., or even 48.125 years

b) Discrete or Discontinuous: A discrete variable can only take some definite or specific values from a continuous value set. Ex. Number of books issued in a library. The number of books issued in a library per day can take only the values 0, 1, 2, 3, ..., 10, ...25, ..., etc. i.e. only the whole number and/or specified value and no fractional value, or unspecified value. E.g. monthly subscription of a journal may be Rs. 450.25 or Rs. 450.26, but not in a fraction of Paisa.

Usually quantitative data are classified as continuous or discrete. If a variable can take on any value between its minimum value and its maximum value, it is called a continuous variable; otherwise, it is called a discrete variable.

Some examples will clarify the difference between discrete and continuous variables.

- Suppose the fire department mandates that all fire fighters must weigh between 150 and 250 pounds. The weight of a fire fighter would be an example of a continuous variable; since a fire fighter's weight could take on any value between 150 and 250 pounds.
- Suppose we flip a coin and count the number of heads. The number of heads could be any integer value between 0 and plus infinity. However, it could not be any number between 0 and plus infinity. We could not, for example, get 2.3 heads. Therefore, the number of heads must be a discrete variable.

9.1.1.3.3. Ungrouped and Grouped Data

The data obtained in original form are called raw data or ungrouped data. These data have not been arranged in any systematic order.

Say as example,

The marks obtained by 25 students in a class in a certain examination are given below:

25, 8, 37, 16, 45, 40, 29, 12, 42, 25, 14, 16, 16, 26, 20, 10, 36, 33, 24, 25, 35, 11, 30, 45, 48.

Grouped Data:

Grouped data is a statistical term used in data analysis. Data presented in the form of frequency distribution is called grouped data. A raw dataset can be organized by constructing a table showing the frequency distribution of the variable (whose values are given in the raw dataset). Such a frequency table is often referred to as grouped data.

The idea of grouped data can be illustrated by considering the above raw dataset in the following way (Table 1):

Table 1: Grouped Data

Marks obtained	No of students (frequency)
0-10	2
11-20	7
21-30	7
31-40	5
41-50	4

9.1.1.3.4. Independent and Dependent variable

A variable is not only something that we measure, but also something that we can manipulate and something we can control for. Keeping in view this aspect, variables can be divided as independent variables and dependent variables.

The independent variable is what is natural and not affected or manipulated by the experiment, rather it is the treatment or the program or the cause that reacts on the dependent variable. It is such a variable that is being manipulated in an experiment in order to observe the effect on a dependent variable. Most experiments consist of observing the effect of the independent variable(s) on the dependent variable(s). The dependent variable is what is affected by the independent variable or by the effects of the experiment. It measures the experimental outcome depending upon on an independent variable(s). The independent variable is the variable we're using to explain why the dependent variable varies. In most experiments, the effects of the independent variable on the dependent variables are observed.

For example, say a tutor asks 100 students to complete a mathematical test. The tutor wants to know why some students perform better than others. Whilst the tutor does not know the answer to this, s/he thinks that it might be because of two reasons:

(1) some students spend more time revising for their test; and

(2) some students are naturally more intelligent than others.

As such, the tutor decides to investigate the effect of revision time and intelligence on the test performance of the 100 students. The dependent and independent variables for the study are:

Dependent Variable: **Test Mark** (measured from 0 to 100)

Independent Variables : **Revision time** (measured in hours) and **Intelligence** (measured using IQ score)

Therefore, the aim of the tutor's investigation is to examine whether these independent variables - revision time and IQ - result in a change in the dependent variable, the students' test scores. However, it is also worth noting that while this is the main aim of the experiment, the tutor may also be interested to know if the independent variables - revision time and IQ are also connected in some way.

9.1.1.3.5. Depending upon scales of measurement

a) Nominal scale data:

Nominal scale data are divided into categories that are only distinguished by their name and labels and cannot be classified one above another e.g. race, name, sex, name of country, name of crops, type of blood. In this type of data there is no implication of order or ratio. Nominal data that falls into two groups are called dichotomous data e.g. male/ female, black/white, rural/ urban.

b) Ordinal scale data:

When the categorical data can be placed in meaningful order on the basis of their quality, it is known as ordinal data. In this the exact difference between the two groups cannot be estimated e.g. scoring of students categorized as A (70% and above), B (60-69 %), C (50-59 %). In this the exact difference between the students placed in grade A and B cannot be estimated.

c) Interval scale data:

Interval scale data are like ordinal data in that they can be placed in a meaningful order. The categories are arranged in equally spaced units and there is no absolute zero point. e.g. temperature where 0° C(Celsius) does not mean no temperature, rather it equals to 32° F (Fahrenheit). /In addition, they have meaningful intervals between items, which are usually measured quantities. E.g., on the Celsius scale the difference between 100°C and 90°C is the same as the difference between 50°C and 40°C.However as because interval scales do not have an absolute zero, ratio of scores are not meaningful e.g. 100°C is not twice as hot as 50°C, and as because 0°C does not indicate a complete absence of heat.

d) Ratio scale data:

A ratio scale has the same properties as an interval scale; but it has an absolute zero. So, meaningful ratios do exist. e.g. weight in grams or pounds, time in seconds or days and pulse rate in beats per minute are all ratio scale data, etc. In the measurement of temperature, the Kelvin scale is the ratio scale, in which 0Ú K (Kelvin) indicates an absolute absence of heat, just as a zero pulse rate indicates an absolute lack of heartbeat. Therefore, it is correct to say that a pulse rate of 120 beats/min is twice as fast as pulse rate of 60 beats / min, or that 300 K is twice as hot as 150 K.

9.1.1.3.6. Univariate, Bivariate and Multivariate DataUngrouped Data:

Statistical data are often classified according to the number of variables being studied.

- **Univariate data: Single-variable** or **univariate** data refers to data where one aspect of a variable at a time is being observed. When we conduct a study that looks at only one variable, we say that we are working with univariate data.

  Suppose, for example, that we conducted a survey to estimate the average weight of high school students. Since we are only working with one variable (weight), we would be working with univariate data.

  This is commonly used to distinguish a distribution of one variable from a distribution of several variables. It is perhaps the simplest form of statistical analysis.

- **Bivariate data**: In statistics, bivariate data is that data-set that has two variables. When we conduct a study that examines the relationship between two variables, we are working with bivariate data.

  For example, suppose we conducted a study to see if there were a relationship between the height and weight of school students. Since, here we are working with two variables (height and weight), we would be working with bivariate data.

  The quantities from these two variables are often represented using a scatter plot (See Section of Unit 12). This is done so that the relationship (if any) between the variables is easily seen.

- **Multivariate data:** Multivariate data is the data in which analysis are based on more than two variables per observation. Usually multivariate data is used for explanatory purposes. Multivariate statistics concerns understanding the different aims and background of each of the different forms of multivariate analysis, and how they relate to each other. The practical implementation of multivariate statistics to a particular problem may involve several types of univariate and multivariate analyses in order to understand the relationships between variables and their relevance to the actual problem being studied.

Suppose, for example, if the above study consists of heights and weights of children, collected over several years.

Multivariate analysis is essentially the statistical process of simultaneously analysing multiple independent variables with multiple dependent variables. Certain types of problem involving multivariate data, for example simple linear regression and multiple regression, are not usually considered as special cases of multivariate statistics because the analysis is dealt with by considering the (univariate) conditional distribution of a single outcome variable given the other variables.

9.1.2. Understanding Statistical Procedures

Statistical analysis techniques can be used to describe data, generate hypotheses, or test hypotheses. Generally, statistics can be divided into three categories: descriptive, inferential, and explanatory

9.1.2.1. Descriptive statistics

Descriptive statistics is the term given to the analysis of data that helps to describe, show or summarize data in a meaningful way such that, for example, patterns might emerge from the data. It is also known as summary statistics. It is used to describe the basic features of the data in a study. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data. Descriptive statistics do not, however, allow us to make conclusions beyond the data we have analysed or reach conclusions regarding any hypotheses we might have made. They are simply a way to describe our data.

For example, sometimes, it is to be reported that what percentage of students in a college borrows books from library or search for online articles? To describe a set of data, we use tables, graphs, and simple statistical procedures.

This part of statistics is very important because if we simply present our raw data it would be hard to visualize what the data was showing, especially if there was a lot of it.

For example, if we had the results of 100 pieces of students' coursework, we may be interested in the overall performance of those students. We would also be interested in the distribution or spread of the marks. Descriptive statistics allow us to do this. How to properly describe data through statistics and graphs is an important topic. Typically, there are two general types of statistic that are used to describe data:

- **Measures of central tendency:** these are ways of describing the central position of a frequency distribution for a group of data. In this case, the frequency distribution is simply the distribution and pattern of marks scored by the 100 students from the lowest to the highest. We can describe this central position using a number of statistics, including the mode, median, and mean.
- **Measures of spread or dispersion:** These are ways of summarizing a group of data by describing how spread out the scores are. For example, the mean score of our 100 students may be 65 out of 100. However, not all students will have scored 65 marks. Rather, their scores will be spread out. Some will be lower and others higher. Measures of spread help us to summarize the dispersion of these scores. To describe this spread, a number of statistics are available to us, including the range, quartiles, absolute deviation, variance and standard deviation.

(The measures will be discussed in Unit 11)

When we use descriptive statistics it is useful to summarize our group of data using a combination of tabulated description (i.e., tables), graphical description (i.e., graphs and charts) and statistical commentary (i.e., a discussion of the results).

When analysing data, such as the marks achieved by 100 students for a piece of coursework, it is possible to use both descriptive and inferential statistics in the analysis of their marks. Typically, in most research conducted on groups of people, you will use both descriptive and inferential statistics to analyse your results and draw conclusions.

9.1.2.2. Inferential Statistics

Inferential statistics are used to make generalizations or inferences about a population based on findings from a sample. We have seen that descriptive statistics provide information about our immediate group of data. For example, we could calculate the mean and standard deviation of the exam marks for the 100 students and this could provide valuable information about this group of 100 students. Any group of data like this, which includes all the data you are interested in, is called a **population**. A population can be small or large, as long as it includes all the data you are interested in. For example, if you were only interested in the exam marks of 100 students, the 100 students would represent your population. Descriptive statistics are applied to populations, and the properties of populations, like the mean or standard deviation. These are called **parameters** as they represent the whole population (i.e., everybody you are interested in). Often, however, you do not have access to the whole population you are interested in investigating, but only a limited number of data instead.

> For example, you might be interested in the exam marks of all postgraduate students of NSOU. It is not feasible to measure all exam marks of all students in all the postgraduate programmes of NSOU. So you have to measure a smaller **sample** of students (e.g., 100 students), which are used to represent the larger population of all the target students.

Properties of samples, such as the mean or standard deviations, are not called parameters, but **statistics**. Inferential statistics are techniques that allow the investigator to use these samples to make generalizations about the populations from which the samples were drawn. It is, therefore, important that the sample accurately represents the population. The process of achieving this is called sampling (sampling strategies are discussed in detail in the previous unit). Inferential statistics arise out of the fact that sampling naturally incurs sampling error and thus a sample is not expected to perfectly represent the population. The methods of inferential statistics are (1) the estimation of parameter(s) and (2) testing of statistical hypotheses. These procedures are a bit more complicated, but very essential to analyse the population. We use inferential statistics to make specific conclusions to a large population, even though we haven't surveyed everyone in this population.

> For example, looking at the example of postgraduate students in the preceding paragraph, it might be possible to investigate every student's result in NSOU, but it might not be possible to enquire every student to find out their individual reactions. So it is logical to enquire a sample instead the whole population to get students' reactions.

9.1.2.3. Explanatory statistics

The last set of statistics is explanatory. Often, descriptive statistics raise interesting causal questions. For example, some students do better than others in the examinations.

Why? Does their socio-economic background affect in their academic performance? – This implies causality: something causes the variations in grades.

With this set of statistical procedures, we do not just have variables, rather we have independent and dependent variables (See in section 9.1.1.3.4.). The dependent variable is what we're trying to explain. Why do some students get higher grades than others? Does their socio-economic status have an effect on their grades? That is, do grades depend on socio-economic status? That is why, in this situation, we would call grades the dependent variable. In this situation, we're saying that socio-economic status is not dependent on anything; therefore, we call it the independent variable. The independent variable is the variable we're using to explain why the dependent variable varies: we think that grades vary by socio-economic status. The dependent or independent nature of a variable relies on the situation we're investigating. Grades are the dependent variable in this example, but in a completely different example, we could use grades as the independent variable to explain why some students pursue certain types of careers.

Of course, all three sets of statistical procedures are related, and any given statistical project likely will involve all three.

Data Analysis:

Data collected for statistical analysis are to be organised in a suitable way to facilitate their analytical study. To understand the characteristics of variables and how these are used in research we should know the following aspects of analysis of variance.

ANOVA: An "Analysis of Variance" (ANOVA) tests three or more groups for mean differences based on a continuous (i.e. scale or interval) dependent variable. It may sefem odd that the technique is called "Analysis of Variance" rather than "Analysis of Means." But, the name is appropriate because inferences about means are made by analyzing variance. ANOVA is used to test general rather than specific differences among means.

ANCOVA: The obvious difference between ANOVA and ANCOVA is the letter "C", which stands for 'covariance'. Like ANOVA, "Analysis of Covariance" (ANCOVA) has a single continuous response variable. Unlike ANOVA, ANCOVA compares a dependent variable by both a factor and a continuous independent variable (e.g. comparing test score by both 'level of education' and 'number of hours spent studying'). The term "factor" refers to the variable that distinguishes this group membership. The term for the continuous independent variable used in ANCOVA is "covariate". ANCOVA is also commonly used to describe analyses with single dependent variable, continuous independent variables, and no factors. Such an analysis is also known as a regression.

MANOVA: Multivariate analysis of variance (MANOVA) extends the analysis of variance to cover cases where there is more than one dependent variable to be analyzed simultaneously. In basic terms, A MANOVA is an ANOVA with two or more continuous response variables. The obvious difference between ANOVA and a "Multivariate Analysis of Variance" (MANOVA) is the "M", which stands for multivariate. Like ANOVA, MANOVA has both a one-way flavour and a two-way flavour. The numbers of factor variables involved distinguish a one-way MANOVA from a two-way MANOVA. A more subtle way that MANOVA differs from ANOVA is that MANOVA compares levels of a factor that has only two levels (also known as binary). When dealing with a single response variable and binary factor (e.g. gender), one uses an independent sample t-test. However, a t-test cannot estimate differences for more than one response variable together, thus a MANOVA fills that need. When comparing two or more continuous dependent variables by a single factor, a one-way MANOVA is appropriate (e.g. comparing 'test score' and 'annual income' together by 'level of education'). A two-way MANOVA also entails two or more continuous response variables, but compares them

9.2. Data Organization

The collection and organization of data are an integral and critical part of the research process. Once piles of data are collected from the sample population, they have to be presented in a systematic form and order so as to bring into focus their salient features. According to Techopedia (www.techopedia.com),

"Data organization, in broad terms, refers to the method of classifying and organizing data sets to make them more useful. Some IT experts apply this primarily to physical records, although some types of data organization can also be applied to digital records".

Data can be presented in the text, in a table, or pictorially as a chart, diagram or graph. But data in textual form cannot reveal the features easily and may not be convenient for statistical analysis if the values of a variable exceed a nominal range of three or four numbers. Sets of numerical results should usually be presented as tables or pictures rather than included in the text. Well-presented tables and graphs can concisely summarise information which would be difficult to describe in words alone. On the other hand, poorly presented tables and graphs can be confusing or irrelevant.

Principles of data presentation:

i. Data should be arranged in such a way that it should create interest in the reader's mind at the first sight.
ii. The information is to be presented in a compact and concise form without losing important details.
iii. The presentation should be in a simple form so as to draw the conclusion directly by viewing at the data.
iv. The presentation should be done in such a way that it can help in further analysis of data.

Organisation of statistical data involves the step of activities of scrutiny of collected data, classification of the data in an orderly and logical manner, presenting the classified data in tabular forms for ease of analysis and ultimately graphical presentation of the analysed data.

9.2.1. Scrutiny of Data

Scrutiny of Data: data collected may require a thorough scrutiny before making it ready for presentation. **'Scrutiny'** refers to the process of removing or correcting invalid, faulty and unreliable data from a dataset. This is very important for obtaining an error-free and accurate inferential result.

Statistical data are collected either by observation or measurement. It is quite impossible to attain absolute accuracy in collected data, as well as in final results. Certain inaccuracies may be readily detected, e.g. inaccuracies that arise from the dropping and shifting of a decimal point. Again, there may be figures which may be very unlikely to consider; e.g. it is not very comfortable to accept the age of a father is 45, while his son is 30 years old. Or, it is not possible to assure absolute accuracy regarding an analytical study of a population based on the analysis of samples from it.

Many statistical analyses try to find out a pattern in a data series, based on a hypothesis or assumption about the nature of the data. Scrutiny of collected data helps in removing those data points which are either-

a) Obviously disconnected with the effect or assumption due to some other factors which apply only to some particular data points; these particular data points are to be ignored, and analysis may be conducted on the remaining data.

b) Obviously erroneous, i.e. some external error is reflected in that particular data point, either due to a mistake during data collection, reporting etc.

The inaccuracies arising in the course of statistical investigations are called 'statistical errors'. Statistical errors may be considered as

a) Biased Errors: Biased errors arise due to personal bias or prejudices of investigators or respondents, and defects in the measuring instrument. The total effects of biased errors increases as the number of observations increases.

b) Unbiased errors: These enter into a statistical enquiry due to chance causes. With increase in in the number of observations, the total effects of unbiased errors diminish. The magnitude of such errors can be predicted based on probability.

9.2.2. Classification of Data

Classification is the process of arranging the collected piles of statistical information under different categories or classes according to some common characteristics possessed by the individual entity. Statistical data collected during the course of an investigation may be so varied that it is not possible to analyse and get the true significance of the values, unless they are arranged properly in a defined way. So, as a rule the primary step in the analysis is to classify and rearrange these into desired homogeneous categories. This grouping of data is determined based on the purpose for which these are to be used. Without this arrangement, the collected data remain useless. Orderly and logical classification of data helps the investigator to appropriately compare, analyse and draw inference of the experiment. Ex. Population census (N G das, 19).

Techopedia defines Data classification as "the process of sorting and categorizing data into various types, forms or any other distinct class. Data classification enables the separation and classification of data according to data set requirements for various business or personal objectives. It is mainly a data management process"[2]. The primary objectives of data classification are:

[1] https://www.techopedia.com/ retrieved on 08.07.2016

- To arrange the collected data in a logical and purposeful way
- To condense and simplify the data
- To readily distinguish similarities and dissimilarities of data
- To facilitate the comparative and analytical study of data
- To reveal the relationships among the set of data
- To pinpoint most *significant features* of the data at a glance
- To have the basis of tabulation
- To be the basis for data visualisation in tabular and/ or graphical method

9.2.2.1. Types of Data classification:

Data collected for statistical analysis may generally be recognised as the following four types

- Geographical: This is the classification of data on the basis of geographical regions. For example, number of college libraries in different states of the country may be classified under the name of states in the following way

Table9.1 : Number of colleges in Top 10 districts in India

District	Number of Colleges
Bangalore, Karnataka	924
Jaipur, Rajasthan	544
Hyderabad, AP	533
Pune, Maharashtra	470
Rangareddy, Telangana	448
Nagpur, *Maharashtra*	433
Mumbai, *Maharashtra*	358
Bhopal, MP	323
Chittoor, AP	290
Indore, MP	290

- Chronological: this classification relates to a certain period or point of time at which the events in question have occurred.

Table 9.2: Use of Electronic and Printed Books by the faculty of NSOU

Year	E-Books	Print Books
2010		
2011		
2012		
2013		
2014		
2015		

- Qualitative: In qualitative classification data are classified on the basis of the character or attribute of the subject matter such as sex, religion, literacy, etc. in this type of classification the attribute under study cannot be measured. For example, a sample population of library users may be classified as follows.
  1. Male and female
  2. Bengalis and non-Bengalis
  3. Employed and unemployed

- Quantitative: Classification of total population on the basis of quantitative class intervals.

Table 9.3: Ages of the users of a town library

Age (in Years)	No. of Users
11-20	50
21-30	200
31-40	260
41-50	360
51-60	90
61-70	40
Total	**1000**

9.2.2.2. Principles of Classification:

There are no hard and fast rules for deciding the class interval, however it depends upon:

- Knowledge of the data
- Lowest and highest value of the set of observations
- Utility of the class intervals for meaningful comparison and interpretation
- The classes should be collectively exhaustive and non-overlapping i.e. mutually exclusive and independent.
- The number of classes should not be too large otherwise the purpose of classification i.e. summarization of data will not be served.
- The number of classes should not be too small either, for this also may obscure the true nature of the distribution.
- The classes should preferably be of equal width. Otherwise the class frequency would not be comparable, and the computation of statistical measures will be laborious.
- More specifically Sturges' formula can be used to decide the number of class interval;

  $K=1+3.322(\log_{10} n)$

  Where k = no. of classes, n=no. of observation
- The width of the class interval may be determined by dividing the range (R) by k

  $$W = \frac{R}{k}$$

  Where,

  R= difference between the highest and the lowest observation.

  w = width of the class interval
- When the nature of data makes them appropriate class interval width of 5 units,

10 units and width that are multiple of 10 tend to make the summarization more comprehensible.

9.2.3. Tabulation

Tabulation may be defined as the logical and systematic organisation of statistical data in rows and columns according to their logical classification. A statistical table thus formed is a systematic and sensible arrangement of quantitative data as per their appropriate grouping. It simplifies the presentation of data and facilitates comparisons and computations.

The purposes of a statistical table are

- To arrange data in appropriate order
- To enable the significance of data readily understood
- To facilitate swift comparison of statistical data
- To facilitate the detection of errors, omissions and repetition of data
- To reveal the characteristics of data.

9.2.3..1. Structure of a statistical table

A statistical table has at least four key parts and some other minor parts. The first four items stated below are the key parts of a statistical table.

1. The Title
2. The Box Head (column captions)
3. The Stub (row captions)
4. The Body
5. Prefatory Notes
6. Foot Notes
7. Source Notes

The general sketch of table indicating its necessary parts is shown below:

..... **THE TITLE** .....

..... **Prefatory Notes** .....

	..... **Box Head**.....		
..... **Row Captions**.....	..... **Column Captions**.....		
..... **Stub Entries**.....	..... **The Body**.....		

Foot **Notes...**

Source Notes...

1. **The Title:**

   A title is the main heading written in capital shown at the top of the table. It must explain the contents of the table and throw light on the table as whole different parts of the heading can be separated by commas there are no full stop be used in the little.

2. **The Box Head (column captions):**

   The vertical heading and subheading of the column are called columns captions. The spaces were these column headings are written is called box head. Only the first letter of the box head is in capital letters and the remaining words must be written in small letters.

3. **The Stub (row captions):**

   The horizontal headings and sub heading of the row are called row captions and the space where these rows headings are written is called stub.

4. **The Body:**

   It is the main part of the table which contains the numerical information classified with respect to row and column captions.

5. **Prefatory Notes:**

   A statement given below the title and enclosed in brackets usually describe the units of measurement is called prefatory notes.

6. **Foot Notes:**

   It appears immediately below the body of the table providing the further additional explanation.

7. **Source Notes:**

   The source notes is given at the end of the table indicating the source from when information has been taken. It includes the information about compiling agency, publication etc...

9.2.3.2. General Rules of Tabulation:

- A table should be simple and attractive. There should be no need of further explanations (details).
- Proper and clear headings for columns and rows should be need.
- Suitable approximation may be adopted and figures may be rounded off.
- The unit of measurement should be well defined.
- If the observations are large in number they can be broken into two or three tables.
- Thick lines should be used to separate the data under big classes and thin lines to separate the sub classes of data.
- Proper syntax is to be used for any missing cell value or missing data within a table

9.2.4. Diagrammatic and graphical presentation

The distribution of data can be presented graphically or pictorially for easy understanding and for quick interpretation. Diagrams and graphs give visual indications of magnitudes, groupings, trends and patterns in the data. These parameters can be more simply presented in the graphical manner. The diagrams and graphs help for comparison of the variables.

9.2.4.1. Diagrammatic presentation

A diagram is a visual form for presentation of statistical data. The diagram refers various types of devices such as bars, circles, maps, pictorials and cartograms etc.

a) Importance of Diagrams

1. They are simple, attractive and easy to understand
2. They give quick information
3. It helps to compare the variables
4. Diagrams are more suitable to illustrate discrete data
5. It will have more stable effect in the reader's mind.

b) Limitations of diagrams

1. Diagrams shows approximate value
2. Diagrams are not suitable for further analysis
3. Some diagrams are limited to experts (multidimensional)
4. Details cannot be provided fully
5. It is useful only for comparison

c) Rules for drawing diagrams

i. Each diagram should have suitable title indicating the theme with which diagram is intended at the top or bottom.

ii. The size of diagram should emphasize the important characteristics of data.

iii. Approximate proposition should be maintained for length and breadth of diagram.

iv. A proper / suitable scale to be adopted for diagram

v. Selection of approximate diagram is important and wrong selection may mislead the reader.

vi. Source of data should be mentioned at bottom.

vii. Diagram should be simple and attractive

viii. Diagram should be effective than complex.

d) Choice or selection of diagram

There are many methods to depict statistical data through diagram. No single diagram is suited for all purposes. The choice of diagram for describing a given set of data requires skill, knowledge and experience. This choice is purposive and depends upon the nature of data. The nature of data will help in taking a decision as to one-dimensional or two-dimensional or three-dimensional diagram. It is also required to know the purpose of presentation to make the diagram effective.

The following points are to be kept in mind for the choice of diagram.

1. To common man having less concepts of statistics, cartogram and pictograms are appropriate.
2. To present the components apart from magnitude of values, sub-divided bar diagram are most suited.
3. To show a large number of components, pie diagram is more suitable.

9.2.4.1. Types of diagrams

1. One dimensional diagrams (line and bar)
2. Two-dimensional diagram (rectangle, square, circle)
3. Three-dimensional diagram (cube, sphere, cylinder etc.)
4. Pictogram
5. Cartogram

9.2.4.1.1. One dimensional diagrams (line and bar)

In one-dimensional diagrams, the length of the bars or lines is taken into account. Widths of the bars are not considered. Bar diagrams are classified mainly as follows.

a) Line diagram

b) Bar diagram: Bar diagrams may be of following types-

i. Vertical bar diagram

ii. Horizontal bar diagram

iii. Multiple (compound) bar diagram

iv. Sub-divided (component) bar diagram

v. Percentage subdivided bar diagram

a) Line diagram

This is simplest type of one-dimensional diagram. On the basis of size of the figures, heights of the bar / lines are drawn. The distances between bars are kept uniform. The limitation of this diagram are it is not attractive cannot provide more than one information.

For example, draw the line diagram for the data in the following table:

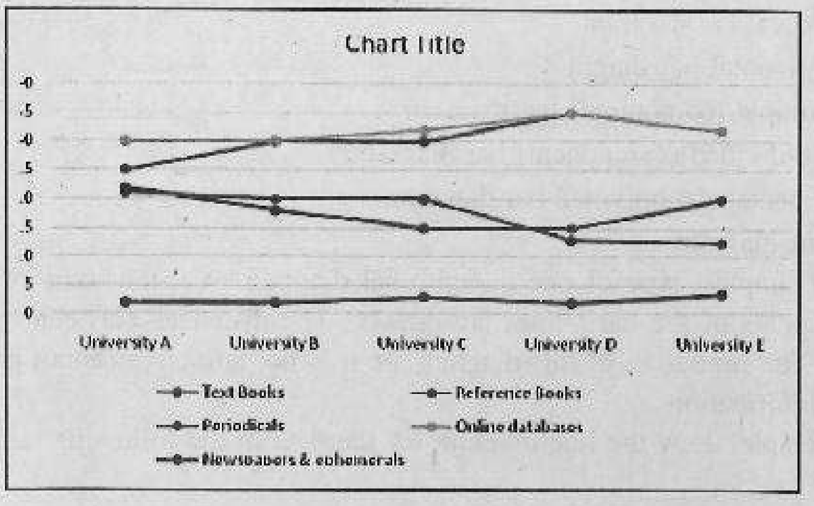
These can be exemplified with the data set presented in the following table:

Table: Distribution of funds of Library budget at five Universities in the Year 2016 (in Lakhs)

	University A	University B	University C	University D	University E
Text Books	22	18	15	15	20
Reference Books	21	20	20	13	12.5
Periodicals	25	30	30	35	32
Online databases	30	30	32	35	32
Newspapers & ephemerals	2	2	3	2	3.5
Total	100	100	100	100	100

a) Line diagram

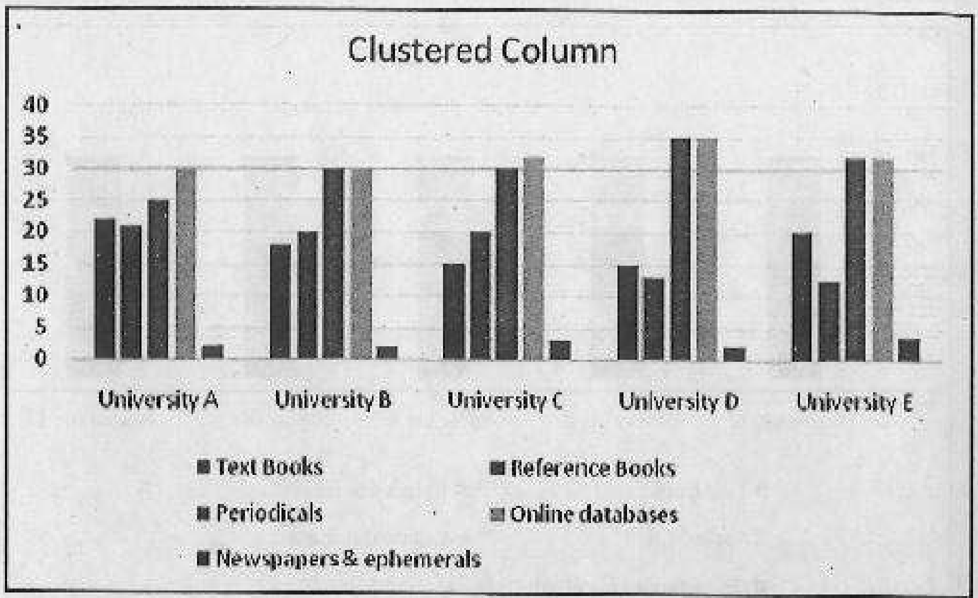
This is simplest type of one-dimensional diagram. On the basis of size of the figures, heights of the bar / lines are drawn. The distances between bars are kept uniform. The limitation of this diagram are it is not attractive cannot provide more than one information



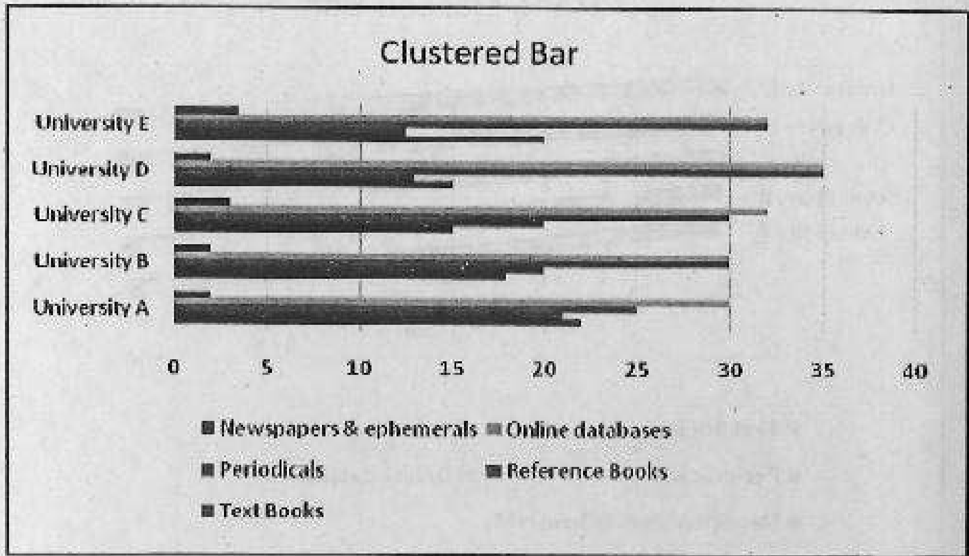
b) Bars diagrams

A simple bar diagram can be drawn using horizontal or vertical bar. In business and economics, it is very a common diagram.

i. Vertical bar diagram:

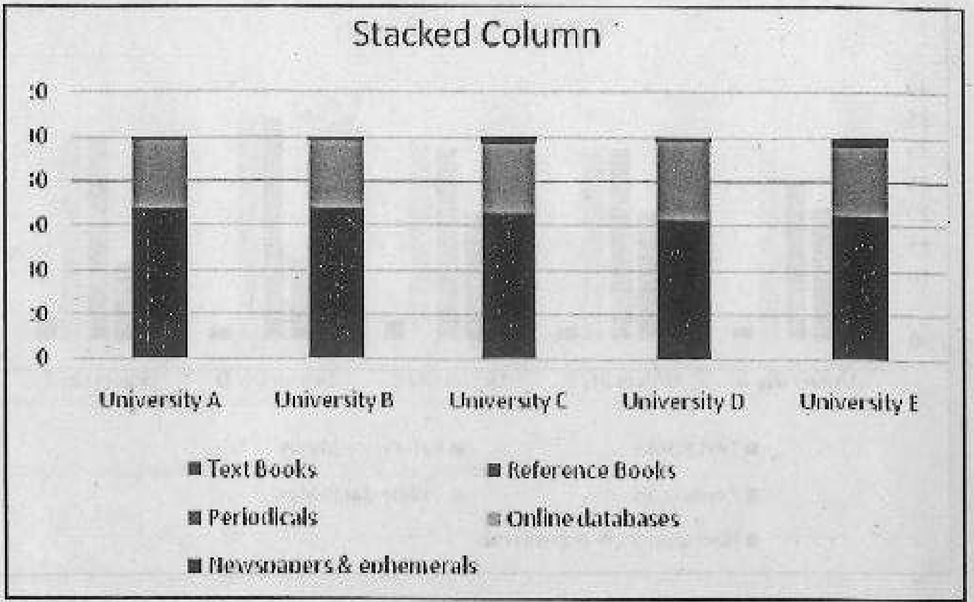


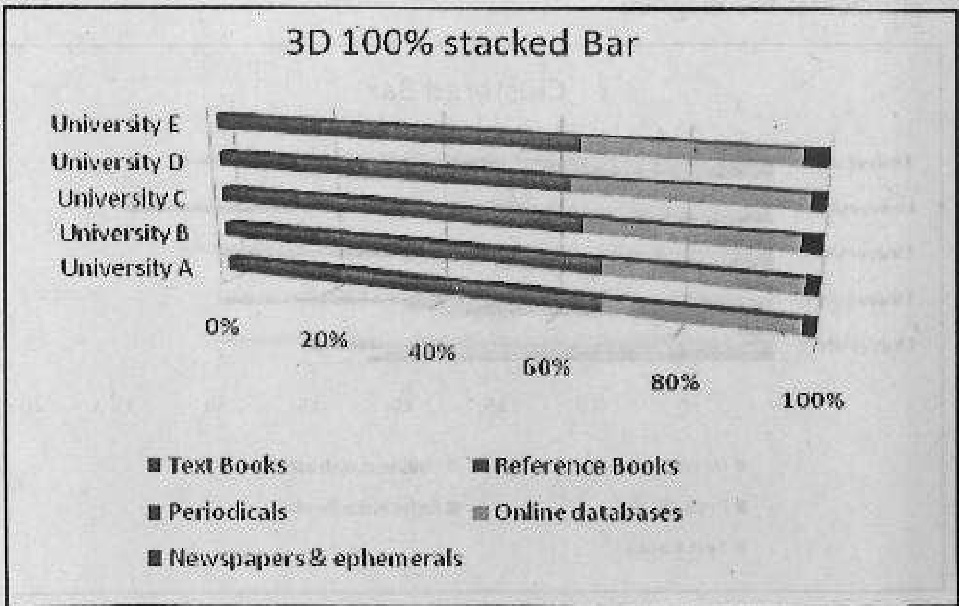
i. Horizontal bar diagram



iii) Compound bar diagram (Multiple bar diagram)

Multiple bar diagrams are used to provide more information than simple bar diagram. Multiple bar diagram provides more than one phenomenon and highly useful for direct comparison. The bars are drawn side-by-side and different columns; shades hatches can be used for indicating each variable used.

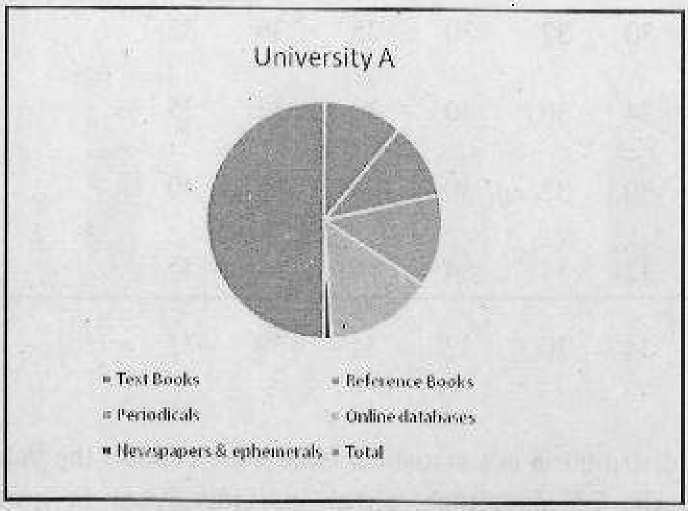




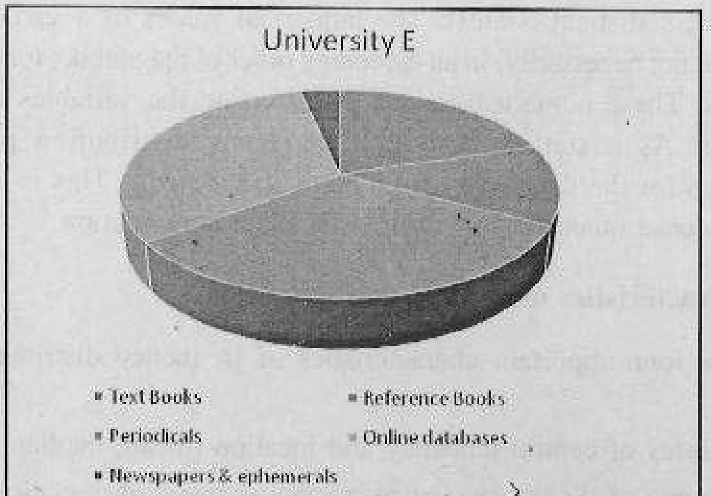
Pie diagram

Pie diagram helps us to show the portioning of a total into its component parts. It is used to show classes or groups of data in proportion to whole data set. The entire pie represents all the data, while each slice represents a different class or group within the whole. Following illustration shows construction of pie diagram.

2-D Pie diagram:



3-D Pie diagram:



9.3. Frequency Distribution

Frequency of a value of the variable is the number of times it occurs in a given series of observations.

For Example, say, Daily number of Book issue in a library over a period of 30 days are listed below

Table: Daily number of issuing books

30	32	30	35	36	32
34	30	40	30	32	35
30	32	30	35	36	40
32	34	34	30	32	35
34	30	32	34	30	32

Frequency distribution is a statistical table which shows the values of the variable arranged in order of magnitude, either individually or in groups, and also the corresponding frequencies side by side. It is an approach to organize the raw numerical values of a variable collected for an investigation, in an orderly manner. It is typically used within a statistical context. The numerical values of a variable are arranged, generally (but not necessarily) in an ascending order of magnitude, for easy manipulation and analysis. These numerical values of occurring the variables are recognised as observations. As a statistical tool, a frequency distribution provides a visual representation for the distribution of a particular variable. This is a common way to list and condense quantitative data for statistical computation.

10.3.1. Characteristics of frequency distribution

There are four important characteristics of frequency distribution. They are as follows:

i. Measures of central tendency and location (mean, median, mode)

ii. Measures of dispersion (range, variance, standard deviation)

iii. The extent of symmetry/asymmetry (skewness)

iv. The flatness or peaked-ness (kurtosis).

10.3.2. Objectives of frequency distribution: The objectives of frequency distribution are

- Condensing the raw numerical data
- Summarizing a large mass of data
- Organisation of the data for interpretation
- Easily obtaining the important characteristics of the data set
- Simplifying the computation
- Comparing different data sets
- Approximating the nature of probability distribution of the population.

10.3.3. Types of frequency distribution :

There are two types of frequency distributions:

1. **Simple frequency distribution:** It shows the values of the variable individually. Data in this type of distribution are ungrouped.

E.g.

Table: Simple frequency distribution

Daily number of book issue	Frequency (No. of Days)
30	9
32	8
34	5
35	4
36	2
40	2
Total	**30**

2. Grouped frequency distribution: Grouped frequency distribution shows the values of the variable in groups or intervals. This can be used when the range of values in the data set is very large. The data must be grouped into classes that are more than one unit in width.

Table: Grouped frequency distribution

Age group(in years)	Frequency(No. of users)
14-19	20
20-24	22
25-29	40
30-34	30
35-39	25
40-49	38
Total	175

10.3.4. Formulation of frequency distribution:

The following technical components are required for formulation of frequency distribution

1. Class/ Class intervals
2. Class Frequency
3. Class limits
4. Class boundaries
5. Mid value/ Class Mark/ Mid-point
6. Class width:

The class width for a class in a frequency distribution is found by subtracting the lower (or upper) class limit of one class minus the lower (or upper) class limit of the previous class.

7. Frequency density
8. Relative frequency

Table: components of frequency distribution

Class	Class Frequency	Class limits		Class boundaries		Mid value	Class width	Frequency density	Relative frequency
		Lower	Upper	Lower	Upper				
15-19	15	15	19	14.5	19.5	17	5	15/5=3	15/150=0.1
20-24	22	20	24	19.5	24.5	22	5	22/5=4.5	22/150=0.15
25-29	40	25	29	24.5	29.5	27	5	25/5=5	40/150=0.26
30-34	30	30	34	29.5	34.5	32	5	30/5=6	30/150=0.2
35-39	25	35	39	34.5	39.5	37	5	25/5=5	25/150=0.17
40-49	18	40	49	39.5	49.5	42	5	18/5=3.6	18/150=0.12
Total	150								1.00

10.3.2.4.1. Class / Class intervals

While arranging large amount of data, they are grouped into different groups according to the size of values to get an idea of the distribution. Each of these groups in the said range of data is called the Class. For example, in the Table ... the Column 1 arranges the age groups (in years) of library users. Each of these groups is called class. Class interval is the difference between the upper and lower class boundaries of any class.

a) Open Ended and Closed Ended Classes

Open-end class :

When one end of a class is not specified, the class is called an open-end class. If, in a frequency distribution, the initial class interval is indeterminate at its beginning and/or the final class interval is indeterminate at its end, the distribution is said to possess "open ended" classes. For example,

Table: Both-side open ended
Table: One-side open-ended

Class	Class Frequency	Class	Class Frequency
Upto 19	20	Upto 19	20
20-24	22	20-24	22
25-29	40	25-29	40
30-34	30	30-34	30
35-39	25	35-39	25
Above 40	38	40-49	38
Total	175	Total	175

b) Methods of forming class-interval

i) Inclusive method (non-overlapping)

When the lower and the upper class limit is included, then it is an *inclusive class interval*

Ex:

Marks	No. of students
20 – 29	5
30 – 39	15
40 – 49	25

A student whose mark is 29 is included in 20 – 29 class interval and a student whose mark in 39 is included in 30 – 39 class interval.

a) Exclusive method (overlapping)

In this method, the upper limits of one class-interval are the lower limit of next class. This method makes continuity of data.

Ex:

Marks	No. of students
20 – 30	5
30 – 40	15
40 – 50	25

A student whose mark is between 20 to 29.9 will be included in the 20 – 30 class. Better way of expressing is

Marks	No. of students
20 to less than 30	
(More than 20 but less than 30)	5
30 to less than 40	15
40 to less than 50	25
Total Students	**50**

Magnitude of class interval

The magnitude of class interval depends on range and number of classes. The range is the difference between the highest and smallest values is the data series. A class interval is taken generally in the multiples of 5, 10, 15 and 20.

Sturges' rule is applied to find the number of classes. It a rule for determining the desirable number of groups into which a distribution of observations should be classified. The formula is given below

$K = 1 + 3.322 \log N.$

K = No. of class

N= Total no. of observations

Ex: If total number of observations are 100, then number of classes could be

$K = 1 + 3.322 \log 100$

$K = 1 + 3.322 \times 2$

$K = 1 + 6.644$

$K = 7.644 = 8$ (Rounded off)

<u>NOTE</u>: Under this formula number of class can't be less than 4 and not greater than 20.

Sturges" formula to find size of class interval

$$\text{Size of class interval (h)} = \frac{\text{Range}}{1 + 3.322 \log N}$$

Ex: In a 5 group of worker, highest wage is Rs. 250 and lowest wage is 100 per day. Find the size of interval.

$$h = \frac{\text{Range}}{1 + 3.322 \log N} = \frac{250 - 100}{1 + 3.322 \log 50} = 55.57 \ @ \ 56$$

10.3.4.2. Class Frequency

The number of observations falling within class-interval is called its class frequency.

Ex: The class frequency of the class 15-19 is 15. It means that there are 15 users between the age of 15 and 19.

If all the class frequencies are added, the total frequency represents total number of items studied. This is called total frequency. In Table... the total frequency is 150.

10.3.4.3. Class limits:

In the construction of grouped frequency distribution, the class intervals are delineated by pairs of numbers. The two numbers used to specify the limits of the class interval for the purpose of tallying the original observations into the various classes are called class limits. The smaller of the pair is called the lower class limit and the larger is upper class limits.

Ex: In the class 15-19, the lower value is 15 and the upper value is 19. These two edges of the class are called lower and upper limits of the class.

10.3.4.4. Class boundary: When measurements are taken on a continuous variable, all observed data are recorded to the nearest whole number of a certain unit. Thus for ages, any age between 14.5 years and 15.5 years is recorded as 15 years and it is similar for the age 19 years. So for getting the real class-limits, the most extreme values in a class intervals are accepted, and these extreme values are called class boundaries. The lower extreme value is the lower class boundary and the upper extreme value is the upper class boundary in a particular class interval. Class boundaries are true class limits.

Working out Class Boundaries:

Lower class boundary= lower class limit – ½ d

Upper class boundary= lower class limit + ½ d

(d= difference between upper class limit of a class and the lower class limit of the subsequent class)

10.3.4.5. Class midpoint or class marks

The mid value or central value of the class interval is called mid-point.

$$\text{Mid-point of a class} = \frac{(\text{lower limit of class} + \text{upper limit of class})}{2}$$

$$= \frac{(\text{lower boundary of class} + \text{upper boundary of class})}{2}$$

10.3.4.6. Class width:

The class width of a class in a frequency distribution is found by subtracting the lower class boundary from the upper class boundary (not the class limits) of a class.

width of a class= Upper class boundary – lower class boundary

For simplification of statistical computation, the classes of a frequency distribution should be of equal size.

For classes of equal width, this width can also be calculated by differentiating

a) The successive lower class limits/ boundaries

b) The successive upper class limits/ boundaries

c) The successive class marks

10.3.4.7. Frequency density: Frequency density of a class is its frequency per unit width. It shows the concentration of frequency in a class.

$$\text{Frequency density} = \frac{\text{Class frequency}}{\text{width of the class}}$$

10.3.4.8. Relative frequency: Relative frequency denotes the class frequency expressed as a fraction of the total frequency. When relative frequencies are shown against the corresponding classes, the table is known as Relative Frequency Distribution.

10.3.4.9. Tally Marks: Tally marks are a quick way of keeping track of numbers in groups of five. One vertical line is made for each of the first four numbers; the fifth number is represented by a diagonal line across the previous four.

Observation(Frequency)	Tally marks	Observation(Frequency)	Tally marks
1		6	
2		7	
3		8	
4		9	
5		10	

10.3.2.5. Constructing a frequency distribution

The following guidelines may be considered for the construction of frequency distribution.

a) The classes should be clearly defined and each observation must belong to one and to only one class interval. Interval classes must be inclusive and non-overlapping. Open end classes should be avoided since creates difficulty in analysis and interpretation.

b) Calculate the range of the observations

Range = H – L

H= Highest value

L=Lowest value

c) Find the number of class intervals applying Sturges' formula. The number of classes should be neither too large nor too small. Too small classes result greater interval width with loss of accuracy. Too many class interval results in complexity. Intervals would be continuous throughout the distribution. This is important for continuous distribution.

Sturges' formula

$K = 1 + 3.322 \log N.$

d) Find out the Width of class interval. All intervals should be of the same width. This is preferred for easy computations.

$$\text{Width of class interval} = \frac{\text{Range}}{\text{Number of classes}}$$

e) The number of observations falling in each class interval, i.e. class frequency, is determined by tally marks

Ex: A sample of 30 persons' weight of a particular class students are as follows. Construct a frequency distribution for the given data.

62	58	58	52	48	53	54	63	69	63
57	56	46	48	53	56	57	59	58	53
52	56	57	52	52	53	54	58	61	63

Steps of construction of frequency distribution

Step 1: Finding the range of data:

Range = H – L = 69 – 46 = 23

Step 2: Finding the number of class intervals/ classes.

K = 1 + 3.222 log 30

(applying Sturges' formula)

K = 5.90H" 6

\ No. of classes = 6

Step 3: Calculating the width of class interval

Width of class interval = = 4

Step 4: Determining the class frequencies by tally marks:

All frequencies belong to each class interval and assign this total frequency to corresponding class intervals as follows. The observations are shown one by one by tally marks in a tally sheet. The tally marks are counted in a group of five and every fifth one in a class interval is placed across the preceding four.

Class interval	Tally Marks	Frequency
46 – 50	III	3
50 – 54	IIII III	8
54 – 58	IIII III	8
58 – 62	IIII I	6
62 – 66	IIII	4
66 – 70	I	1

10.3.2.6. Cumulative frequency distribution

Cumulative frequency distribution is a statistical table which shows the values of the variable with their corresponding cumulative frequencies. It is the running total of frequencies and can also be described as the sum of all previous frequencies up to the current class interval. It indicates directly the number of observations that lie above or below the specified values of the class intervals.

Cumulative frequencies may be of two types – 'less/ smaller than' and 'more/ greater than' Cumulative frequency. When the interest of the investigator is on number of cases below the specified value, then the specified value represents the upper limit of the class interval. It is known as 'less than' cumulative frequency distribution. When the interest lies in finding the number of cases above specified value then this value is taken as lower limit of the specified class interval. Then, it is known as 'more than' cumulative frequency distribution.

The cumulative frequency simply means that summing up the consecutive frequency.

Table: 'Less than' cumulative frequency distribution

Marks	**No. of students**	**'Less than' cumulative frequency**
0 – 10	5	5
10 – 20	3	8
20 – 30	10	18
30 – 40	20	38
40 – 50	12	50

In the above 'less than' cumulative frequency distribution, there are 5 students less than 10, 3 less than 20 and 10 less than 30 and so on.

Similarly, following table shows 'greater than' cumulative frequency distribution.

Table: 'more than' cumulative frequency distribution

Marks	**No. of students**	**'more than' cumulative frequency**
0 – 10	5	50
10 – 20	3	45
20 – 30	10	42
30 – 40	20	32
40 – 50	12	12

In the above 'more than' cumulative frequency distribution, 50 students are scored more than 0, 45 more than 10, 42 more than 20 and so on.

10.3.3. Visual Representation of frequency distribution:

A frequency distribution of data can be presented in a table and graph/ diagram. A graphic presentation of frequency distribution is a visual form of presentation. Graphs are drawn on graph paper with the y-axis representing the frequency count, and the x-axis representing the variable to be measured. The Advantages of graphic presentation are

- It provides attractive and impressive view
- Simplifies complexity of data
- Helps for direct comparison
- It helps for further statistical analysis
- It is simplest method of presentation of data
- It shows trend and pattern of data

Some commonly used graphical methods of showing frequency distributions include

- Histograms,
- Frequency Polygon and
- Ogive/ Cumulative frequency polygon

10.3.3.1. Frequency Histogram

Histogram is the most common form of diagrammatic representation of grouped frequency distribution for a continuous variable. In this type of graphical presentation, the given data are plotted in the form of series of rectangles. Class intervals are marked along the x-axis and the frequencies are along the y-axis according to suitable scale. The width of rectangles, one for each class, extends over the class boundaries (not class limits).

In case of equal class width, the heights of rectangles indicate the corresponding frequency density. The height of rectangle is proportional to respective frequency and width represents the class interval.

For classes with unequal width, the rectangles will also be of unequal width, and then the heights will be proportional to frequency densities.

Although the vertical bar chart and histogram appear to be alike, they are quite different in nature. Bar chart is one-dimensional, and a histogram is two-dimensional in which the length and width are both important. Unlike bar chart, each rectangle in a histogram is joined with other and the blank space between the rectangles would mean that the category is empty and there are no values in that class interval. Above all, the histogram is an area diagram, in which the widths of the rectangles are considered, and thus it is two dimensional, but in a bar diagram only the height is all important, the spacing and with of the bars are arbitrary.

Ex: Construct a histogram for following data.

Marks obtained (x)	No. of students (f)	Mid-point
15 – 25	5	20
25 – 35	3	30
35 – 45	7	40
45 – 55	5	50
55 – 65	3	60
65 – 75	7	70
Total	30	

For convenience sake, we will present the frequency distribution along with mid-point of each class interval, where the mid-point is simply the average of value of lower and upper boundary of each class interval.

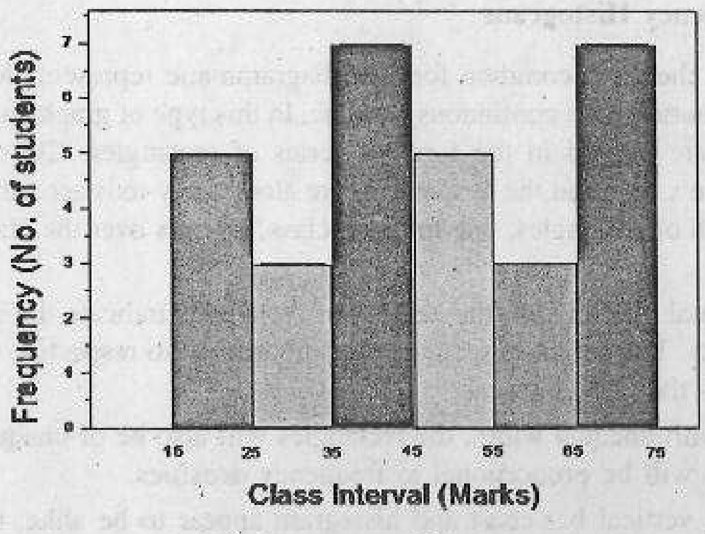


Fig. Frequency Histogram

10.3.3.2. Frequency polygon

A frequency polygon is a line chart of frequency distribution in which either the values of discrete variables or the mid-point of class intervals are plotted against the frequency and those plotted points are joined together by straight lines. Since, the frequencies do not start at zero or end at zero, this diagram as such would not touch horizontal axis. However, since the area under entire curve is the same as that of a histogram which is 100%. The curve must be 'enclosed', so that starting mid-point is jointed with 'fictitious' preceding mid-point whose value is zero. So that the beginning of curve touches the horizontal axis and the last mid-point is joined with a 'fictitious' succeeding mid-point, whose value is also zero, so that the curve will end at horizontal axis. This enclosed diagram is known as 'frequency polygon'.

Ex: For following data construct frequency polygon.

Marks (CI)	No. of frequencies (f)	Mid-point
15 – 25	5	20
25 – 35	3	30
35 – 45	7	40
45 – 55	5	50
55 – 65	3	60
65 – 75	7	70

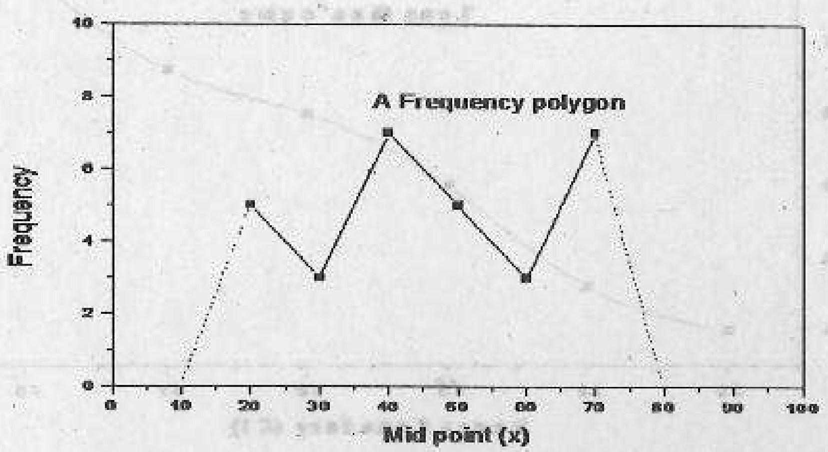


Fig. Frequency Polygon

10.3.3.3. Ogives

Ogives are the graphic representations of a cumulative frequency distribution, and therefore also known as cumulative frequency curve. These ogives are classified as 'less than' and 'more than' types. In case of 'less than', cumulative frequencies are plotted against upper boundaries of their respective class intervals. In case of 'more than' type, cumulative frequencies are plotted against upper boundaries of their respective class intervals. These ogives are used for comparison purposes. Several ogives can be compared on same grid with different colour for easier visualisation and differentiation.

Ex:

Marks (CI)	No. of frequencies (f)	Mid-point	Cum. Freq. Less than	Cum. Freq. More than
15 – 25	5	20	5	30
25 – 35	3	30	8	25
35 – 45	7	40	15	22
45 – 55	5	50	20	15
55 – 65	3	60	23	10
65 – 75	7	70	30	7

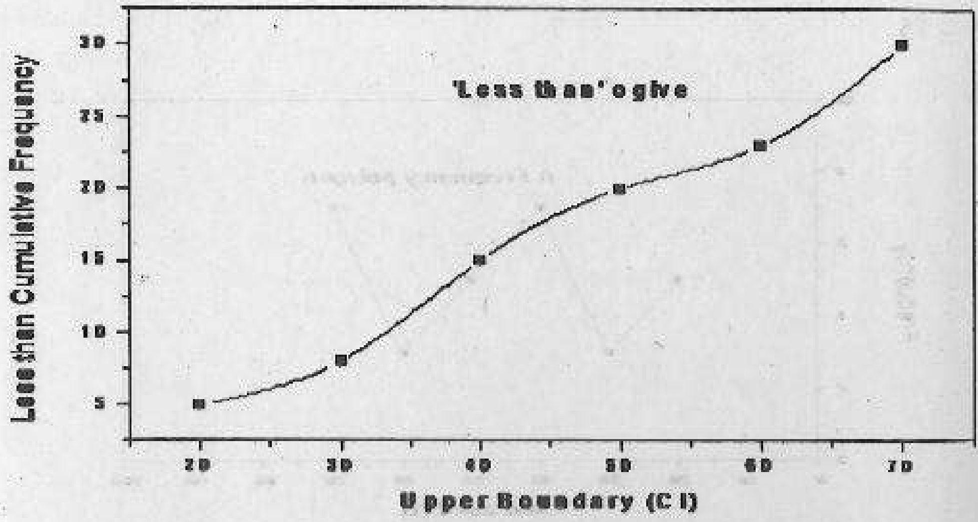


Fig. Less than Ogive diagram

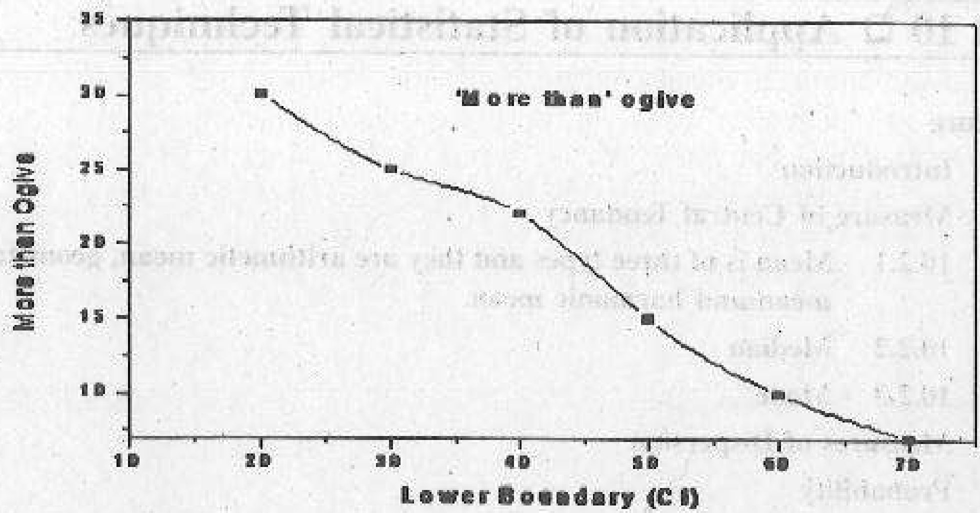


Fig. More than Ogive diagram

In this example, the y-axis is the number of children and the x-axis is the height. In general, the chart will show a normal distribution, which means that the majority of occurrences, or in this case children with a certain height, will fall in the middle column. In a histogram, the height of the column represents the range of values for that variable.

Unit 10 ❑ Application of Statistical Techniques

Structure

10.1 Introduction

This unit is the basic foundation of descriptive statistics which quantitatively summarises features of numerical information. It includes measure of central tendencies, measure of dispersion, probability theory, distributions, parametric and non-parametric statistical tests.

10.2 Measure of Central Tendancy

A measure of central tendency is a single value that attempts to describe a set of data by identifying the central position within that set of data. As such, measures of central tendency are sometimes called measures of central location. They are also classed as summary statistics. Central tendency works as reference point indicating the distribution's scatter around a central point. A **measure of central tendency (also referred to as measures of centre or central location) is a summary measure that attempts to describe a whole set of data with a single value that represents the middle or centre of its distribution.**

It is the standard way of describing a large set of data by calculating a representative standard number popularly known as Average. It can be also thought as the figures/ distributions is grouping around a central point. There are three types of averages- Mean, Median and Mode

10.2.1 Mean is of three types and they are arithmetic mean, geometric mean and harmonic mean.

Arithmetic mean: The mean, or more precisely **the arithmetic mean,** is simply the arithmetic average of a group of numbers (or data set) and is shown using a bar (‾) symbol. So the mean of the variable x is pronounced "x-bar" ($\bar{x}$).

Ungrouped data:- It is calculated by adding up all of the values in a data set and dividing by the number of values in that data set:

$$\bar{x} = \frac{\sum x}{n}$$

For example, the prices of 5 books in Rupees are 100.50, 125.50, 123, 145, 126.. The mean of this data would be

$$\bar{x} = \frac{\sum x}{n} = \frac{100.50 + 125.50 + 123 + 145 + 126}{5} = \frac{620}{5} = Rs.124/-$$

Grouped Data:- Let X be a variable and $x_1,x_2,x_3,x_4,x_5,x_6 \ldots x_n$ are the mid-values of n class intervals and $f_1,f_2,f_3,f_4,f_5 \ldots f_n$ are the frequencies of the n-class intervals. The mean is calculated as-

$$\overline{X} = \frac{x1f1 + x2f2 + x3f3 + \ldots\ldots + xnfn}{f1 + f2 + f3 + \ldots\ldots + fn} = \frac{\sum_{i=1}^{n} xi \ast fi}{\sum_{i=1}^{n} fi}$$

Here $\overline{X}$ is the weighted mean of x_1, $x_2,x_3 \ldots\ldots x_n$ and frequencies are the weights.

Example:- A departmental Library has 320 books and the prices of each books are tabulated in an interval for and the number of books are tallied along with each price group range. What is the mean value of price for the library collection?

Price of Books(INR) X	Frequency f_i	Mid-value x_i	Product $f_i * x_i$
50-99	6	74.5	447
100-149	10	124.5	1245
150-199	25	174.5	4362.5
200-249	36	224.5	8082
250-299	14	274.5	3843
300-349	18	324.5	5841
350-399	19	374.5	7115.5
400-449	56	424.5	23772
450-499	63	474.5	29893.5
500-549	12	524.5	6294
550-599	14	574.5	8043
600-649	19	624.5	11865.5
650-699	28	674.5	18886
Total	320		129690

The weighted mean is, $\frac{\sum_{i=1}^{n} xi*fi}{\sum_{i=1}^{n} fi} = \frac{129690}{320} = Rs.405.28\,/-$

Geometric Mean: The Geometric Mean is calculated by taking the nth root of the product of a set of data.

$$g = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

$$\log g = \frac{1}{n}\sum_{i=1}^{n} \log x_i$$

For example, if the set of data is: 1,2,3,4,5

The geometric mean would be calculated:

$$g = \sqrt[5]{1 \times 2 \times 3 \times 4 \times 5} = (1 \times 2 \times 3 \times 4 \times 5)^{1/5}$$

$$\log g = \frac{1}{5}\log(1 \times 2 \times 3 \times 4 \times 5) = \frac{1}{n}(\log 1 + \log 2 + \log 3 + \log 4 + \log 5)$$

Thus we get,

$$\log g = \frac{1}{n}\sum_{i=1}^{n} \log x_i$$

When to use the geometric mean

The arithmetic mean is relevant at any time several quantities add together to produce a total.

The arithmetic mean answers the question, "if all the quantities had the same value, what would that value have to be in order to achieve the same total?"

In the same way, the geometric mean is relevant any time several quantities multiply together to produce a product.

The geometric mean answers the question, "if all the quantities had the same value, what would that value have to be in order to achieve the same product?"

For example, suppose you have an investment which returns 10% the first year, 50% the second year, and 30% the third year. What is its average rate of return?

It is not the arithmetic mean, because what these numbers mean is that on the first year your investment was multiplied (not added to) by 1.10, on the second year it was multiplied by 1.50, and the third year it was multiplied by 1.30.

The relevant quantity is the geometric mean of these three numbers.

It is known that the geometric mean is always less than or equal to the arithmetic mean (equality holding only when A=B).

The proof of this is quite short and follows from the fact that ("A " "B)^2 is always a non-negative number.

Harmonic Mean

The arithmetic mean cannot be used when we want to average quantities such as speed.

Consider the example below:

Example 1: The distance from my house to town is 40 km. I drove to town at a speed of 40 km per hour and returned home at a speed of 80 km per hour. What was my average speed for the whole trip?

Solution: If we just took the arithmetic mean of the two speeds I drove at, we would get 60 km per hour.

This isn't the correct average speed, however: it ignores the fact that I drove at 40 km per hour for twice as long as I drove at 80 km per hour.

To find the correct average speed, we must instead calculate the harmonic mean.

For two quantities A and B, the harmonic mean is given by,

$$h = \frac{2}{\frac{1}{A}+\frac{1}{B}} = \frac{2}{\frac{B+A}{AB}} = \frac{2AB}{A+B}$$

For N quantities: A, B, C......

Harmonic mean (for ungrouped data), $h = \frac{N}{\frac{1}{A}+\frac{1}{B}+\frac{1}{C}+\cdots+\frac{1}{N}} = \frac{N}{\sum_{i=1}^{N} \frac{1}{x_i}}$

Let us try out the formula above on our example:

Harmonic mean = $\frac{2AB}{A+B}$

Our values are A = 40, B = 80.

Therefore, harmonic mean = $\frac{2 \times 40 \times 80}{40+80} = \frac{6400}{120}$ H" 53.333

Is this result correct? We can verify it.

In the example above, the distance between the two towns is 40 km. So the trip from A to B at a speed of 40 km will take 1 hour. The trip from B to A at a speed to 80 km will take 0.5 hours. The total time taken for the round distance (80 km) will be 1.5 hours.

The average speed will then be

$\frac{80}{1.5}$ km/ hr =53.33 km/ hr

Weighted harmonic mean(Grouped Data) H= $\frac{N}{\sum_{i=1}^{N} \frac{f_i}{x_i}}$ [f_i is frequency of occurrence of x_i]

The harmonic mean also has physical significance.

10.2.2 Median

The median is the "middle value" in a set. That is, the median is the number in the center of a data set that has been ordered sequentially.

For example, let's look at the data in a data set: {10,14,86,2,68,99,1}. What is its median?

- First, we sort our data set sequentially: {1,2,10,14,68,85,99}
- Next, we determine the total number of observations in our data set (in this case, 7.)
- Finally, we determine the central position of or data set (in this case, the 4th position),

So, the number in the central position is our median - {1,2,10,14,68,85,99}, making 14 our median.

An easy way to determine the central position or positions for any odd set is to take the total number of points, add 1, and then divide by 2.

If the number you get is a whole number, then that is the central position.

If the number you get is a fraction, take the two whole numbers on either side.

Because our data set had an odd number of points, determining the central position was easy - it will have the same number of points before it as after it.

But what if our data set has an even number of points?

Let's take the same data set, but add a new number to it: {1,2,10,14,68,85,99,100} What is the median of this set?

When you have an even number of points, you must determine the two central positions of the data set.

So for a set of 8 numbers, we get (8 + 1) / 2 = 9 / 2 = 4 1/2, which has 4 and 5 on either side.

Looking at our data set, we see that the 4th and 5th numbers are 14 and 68. From there, we return to our trusty friend the mean to determine the median.

(14 + 68) / 2 = 82 / 2 = 41.

- Find the median of 2,4, 6, 8

firstly, we must count the numbers to determine it is odd or even

Here we can see it is even so we can write:

M= (4+6)/2=10/2=5 [5 is the median of above sequential numbers]

10.2.3 Mode

The mode is the most common or "most frequent" value in a data set. It shows central tendency of a data set by the expression of the value(s) which occurs most frequently.

Example: the mode of the following data set (1, 2, 5, 5, 6, 3) is 5 since it appears twice.

This is the most common value of the data set.

Data sets having one mode are said to be unimodal, with two are said to be bimodal and with more than two are said to be multimodal.

An example of a unimodal dataset is {1, 2, 3, 4, 4, 4, 5, 6, 7, 8, 8, 9}. The mode for this data set is 4.

An example of a bimodal data set is {1, 2, 2, 3, 3}. This is because both 2 and 3 are modes.

Please note: If all points in a data set occur with equal frequency, it is equally accurate to describe the data set as having many modes or no mode.

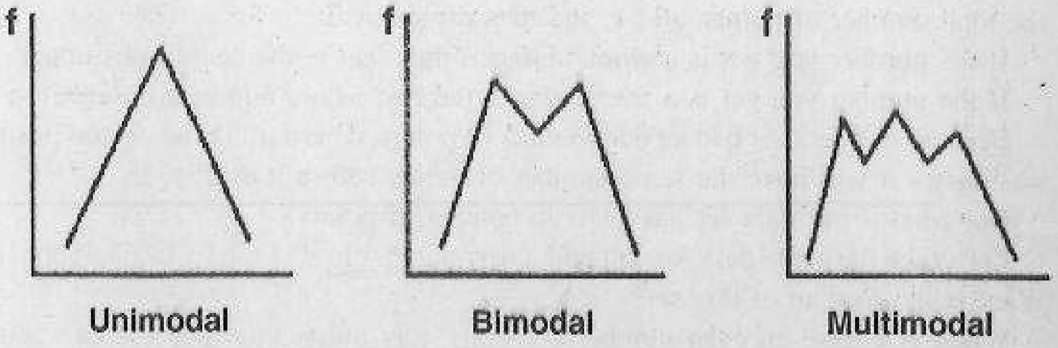


Figure:- Three types of modal distribution

It must also know that, some distributions do not contain any mode value such as reverse J shaped distribution, U-shaped distribution etc.

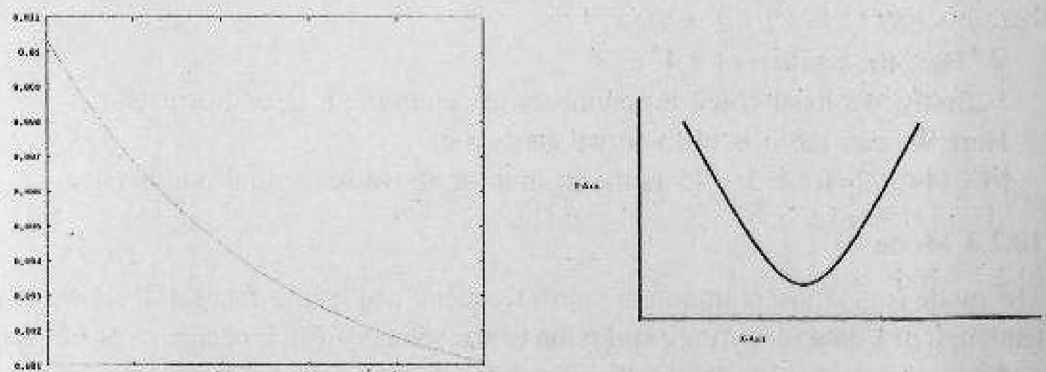


Figure:- No mode found for the distributions

Relationship of the Mean, Median, and Mode:

The relationship of the mean, median and mode to each other can provide some information about the relative shape of the data distribution. If the mean, median, and mode are approximately equal to each other, the distribution can be assumed to be approximately symmetrical. If the mean > median > mode, the distribution will be skewed to the left or negatively skewed. If the mean < median < mode, the distribution will be skewed to the right or positively skewed.

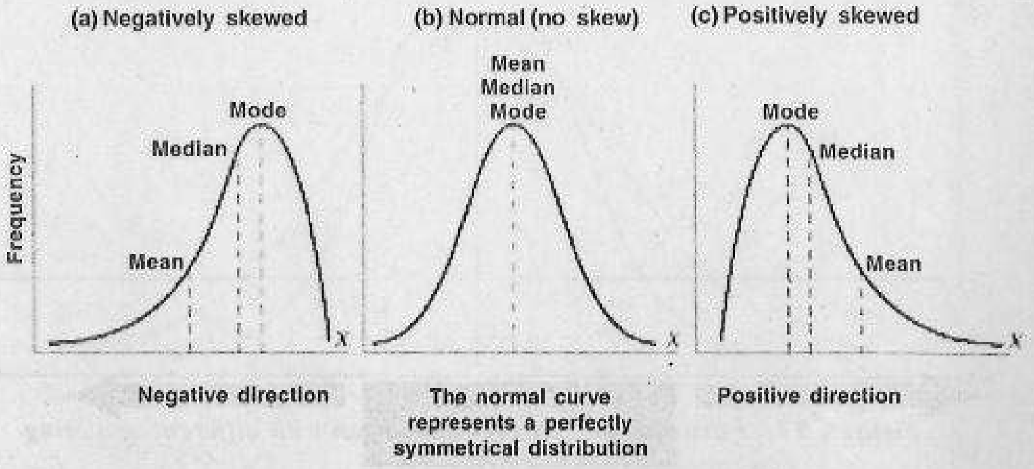


Figure:- Skewness

10.3 Measures of Dispersion

It is not always possible to understand the nature of data only by determining several central tendency measurements. Central tendency only shows the average value or a central point of a variable. Statisticians also study measure dispersion or spread of data from a central reference point.

Measures of dispersion are needed for four basic purposes:

(i) To determine the reliability of an average.

(ii) To serve as a basis for the control of the variability.

(iii) To compare two or more series with regard to their variability.

(iv) To facilitate the use of other statistical measures.

Measures of variation point out as to how far an average is representative of the mass. When dispersion is small, the average is a typical value in the sense that is closely represents the individual value and it is reliable in the sense that it is a good estimate of the average in the corresponding universe. On the other hand, when dispersion is large the average is not so typical, and unless the sample is very large, the average may be quite unreliable. It is also possible that distributions might have same mean values with different scatter over the interval space. In such cases, representativeness of central tendency about the populations from sample data sets might estimate very erroneous assumptions.

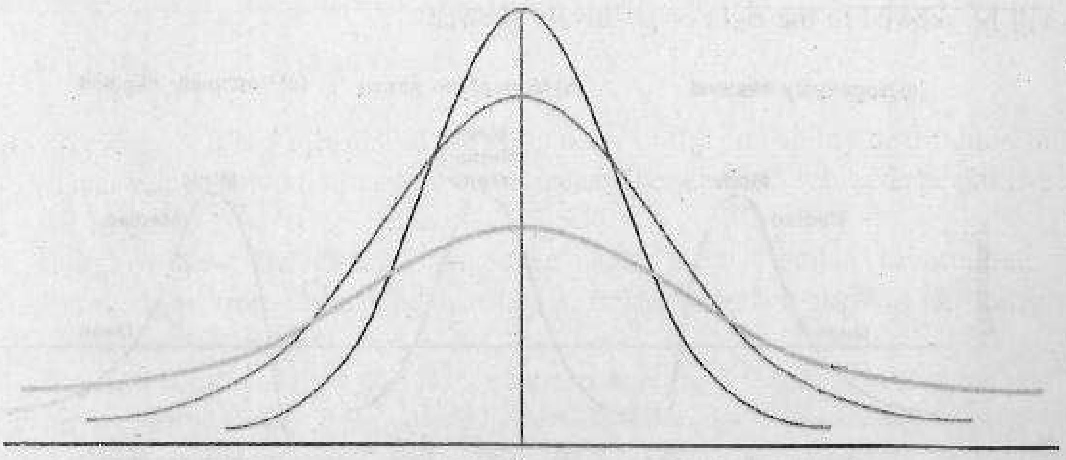


Figure:- Three distributions having same mean with different scattering

Another purpose of measuring dispersion is to determine nature and cause of variation in order to control the variation itself. In matter of health, variations in body temperature, pulse beat and blood pressure are the basic guides to guides to diagnosis. Prescribed treatment is designed to control their variation. In industrial production efficient operation requires control of quality variation, the cause of which are sought through inspection and quality control programmes. Thus measurement of dispersion is basic to the control of cause of variation. In engineering problems measures of dispersion are often especially important. In social sciences a special problem requiring the measurement of variability is the measurement of "inequality" of the distribution of income or wealth, etc.

Measures of dispersion enable a comparison to be made of two or more series with regard to their variability. The study of variation may also be looked upon as a means of determining uniformity or consistency. A high degree of variation would mean little uniformity or consistency whereas a low degree of variation would mean great uniformity or consistency. The measurements of dispersions are discussed as follows-

A. Range-It is the difference between the highest data point and the lowest data point. Suppose in a library the number of books issued are 12, 15, 48, 96, 36, 25, 41, 25, 36, 14, 25, 85, 76, 45, 21. So, the range of the distribution is (96-12)=84. Th given data is scattered to 84 data points from the lowest value. The range is expressed in terms of its lowest and highst value: 96 – 11.

Inter-Quartile Range:- In descriptive statistics, the interquartile range (IQR), The interquartile range (IQR) is a measure of variability, based on dividing a data set into quartiles. Quartiles divide a rank-ordered data set into four equal parts. The values that divide each part are called the first, second, and third quartiles; and they are denoted by Q1, Q2, and Q3, respectively. It may be stated as the midspread or middle 50%, or technically H-spread, is a measure of statistical dispersion, being equal to the difference between 75th and 25th percentiles, or between upper and lower quartiles, IQR = Q3 " Q1.

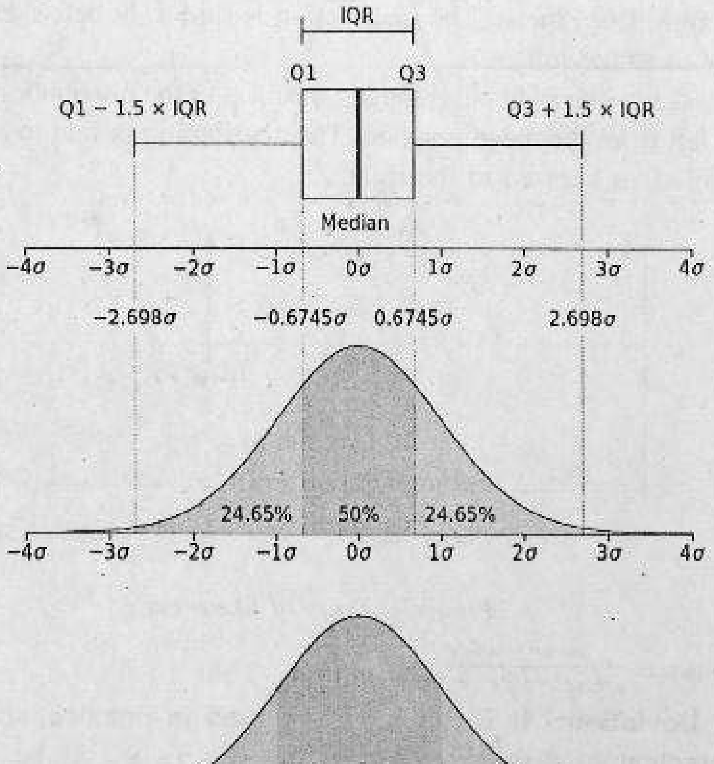


Figure:- Illustrations of IQR

Example:-

Serial No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Values	12	22	25	29	36	39	47	59	65	76	88	99	121
Quartiles				Q1			Median Q2				Q3		

So, IQR= Q3-Q1 = (88-29) = 59

B. Skewness:- It is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive or negative, or even undefined.

Negative skew: The left tail is longer; the mass of the distribution is concentrated on the right from mean. The distribution is said to be left-skewed, left-tailed, or skewed to the left.

Positive skew: The right tail is longer; the mass of the distribution is concentrated on the left from the mean position. The distribution is said to be right-skewed, right-tailed, or skewed to the right.

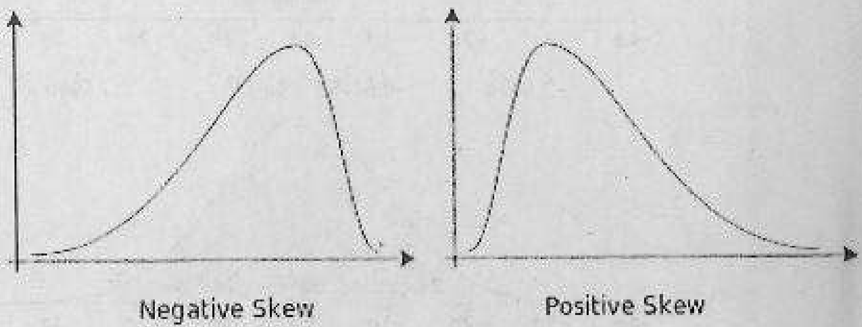


Figure:- Types of Skewness

Skewness = $\frac{mean - mode}{standard\ deviation}$

C. **Mean Deviation:-** It is though rarely used in practice, still it is a basic measurement of dispersion. Let x_1, x_2 ,x_3, x_4, x_5...x_n be n data points / observations and $\bar{x}$ is the mean. If $d_i = (x_i - \bar{x})$ is the individual difference from the mean (i=1,2,3,4,5,...n), then

Mean Deviation = $\frac{\sum_{i=1}^{n} |di|}{n} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$

Taking the previous example ,

mean $(\bar{x}) = \frac{12+15+48+96+36+25+36+14+25+85+76+46+21}{15} = 40$

Deviation is calculated as-

$|d_1| = |12-40| = |-28| = 28$, $|d_2| = |15-40| = |-25| = 25$ and so on.

Now, mean deviation is –

$$\text{M.D.} = \frac{28+25+8+56+4+15+1+15+4+26+15+45+36+5+19}{15}$$

$$= \frac{302}{15} = 20.133$$

C. **Root Mean Square Deviation –** This the most frequently used to measure dispersion of a particular variable but not among the variables.

Variance- It is denoted by and defined as = $\sigma^2 = \frac{\sum_{i=1}^{n} (x-\bar{x})^2}{n}$

Here, x_i are the value of the variable for i=1,2,3,....n. For groupd data, it may be

dfind as- $\sigma^2 = \frac{\sum_{i=1}^{n} (x-\bar{x})^2 fi}{\sum fi}$, here x_i are the mid-values and f_i are the frequncies in the i-th class.

D. **Standard Deviation**- The standard deviation of a distribution is the square root of variance and denoted by σ (sigma). In addition to expressing the variability of a population, the standard deviation is commonly used to measure confidence in statistical conclusions. For example, the margin of error in polling data is determined by calculating the expected standard deviation in the results if the same poll were to be conducted multiple times. This derivation of a standard deviation is often called the "standard error" of the estimate or "standard error of the mean" when referring to a mean. It is computed as the standard deviation of all the means that would be computed from that population if an infinite number of samples were drawn and a mean for each sample were computed. It is very important to note that the standard deviation of a population and the standard error of a statistic derived from that population (such as the mean) are quite different but related (related by the inverse of the square root of the number of observations).

Example:- Suppose in a library the number of books issued are 12, 15, 48, 96, 36, 25, 41, 25, 36, 14, 25, 85, 76, 45, 21 (for 15 days).Find the standard deviation.

$$\text{mean } (\bar{x}) = \frac{12+15+48+96+36+25+41+25+36+14+25+85+76+46+21}{15} = 40$$

So Variance is, $\sigma^2 = \dfrac{\sum_{i=1}^{n}(x-\bar{x})^2}{n}$ = **26.322**

Standard Deviation, σ = **√26.322 = 5.130515**

E. <u>Co-efficient of Variance(C.V.)</u> = 100 *

Mean deviation and standard deviation are expressed in terms of the variable. So these two can't be used to compare different series of data . C.V. is the percent of variation in its mean while considering standard deviation as the total variation in its mean. C.V. thus can be conveniently used to compare deviations among different series of data.

Example -1:- C.V. of the previous example = $\dfrac{5.130515}{50} * 100 = 12.82\%$

It is to be interpreted as 11.82% variability present in the data.

Example-2:- Two libraries have their own cataloguing section and the number of books catalogued are to be compared in order to measure their variability(for 7 days).

Library-1 :- 125,123,145,129,75,85,90 books in 7 days

Library-2 :- 139,168,145,125,147,99,102 books in 7 days

Answer: For Library 1, mean (μ) = 110.2857 , Standard Deviation(σ)= 26.55004

C.V. $\dfrac{\sigma}{\mu} * 100 = \dfrac{26.5504}{110.2857} * 100 = 24.07\%$

For Library 2, mean (μ) = 132.1429 , Standard Deviation (σ) = 25.08936

C.V. = $\dfrac{\sigma}{\mu} * 100 = \dfrac{25.08936}{1321429} * 100 = 18.986\% \approx 19\%$

It can be concluded as Library A has greater variability in their data (no. Of books catalogued).

C.V. can also be measured with respect to median value instead of mean value.

10.4 Probability

Probability is the measure of the likelihood that an event will occur. Probability is quantified as a number between 0 and 1 (where 0 indicates impossibility and 1 indicates certainty). The higher the probability of an event, the more certain that the event will occur. A simple example is the tossing of a fair (unbiased) coin. Since the coin is unbiased, the two outcomes ("head" and "tail") are both equally probable; the probability of "head" equals the probability of "tail." Since no other outcomes are possible, the probability is 1/2 (or 50%), of either "head" or "tail". In other words, the probability of "head" is 1 out of 2 outcomes and the probability of "tail" is also 1 out of 2 outcomes, expressed as 0.5 when converted to decimal, with the above-mentioned quantification system. This type of probability is also called a priori probability.

These concepts have been given an axiomatic mathematical formalization in probability theory, which is used widely in such areas of study as mathematics, statistics, finance, gambling, science (in particular physics), artificial intelligence/machine learning, computer science, game theory, and philosophy to, for example, draw inferences about the expected frequency of events. Probability theory is also used to describe the underlying mechanics and regularities of complex systems.

Random Variables and Expectations-

- **Random Experiment:** An experiment whose outcomes are determined only by chance factors is called a random experiment. Example:- Tossing coin(s), choosing an object from the lot without any bias, throwing two dices etc.
- **Sample Space:** The set of all possible outcomes of a random experiment is called a sample space. Example:- each result of two dices.
- **Event:** The collection of none, one, or more than one outcome from a sample space is called an event. Example:- Sum of results of two dices.
- **Random Variable:** A variable whose numerical values are determined by chance factors is called a random variable. Formally, it is a function from the sample space to a set of real numbers.
- **Discrete Random Variable:** If the set of all possible values of a random variable X is countable, then X is called a discrete random variable. Example:- number of students absent in a class, number of heads after flipping a coin 10 times, number of white marbles in a jar where red, blue, white and green marbles are present, student's mark in mathematics in final semester etc.
- **Probability of an Event:** If all the outcomes of a random experiment are

equally likely, then the probability of an event A is given by-

$$P(A) = \frac{\textit{Number of outcomes in the event A}}{\textit{Total number of outcomes in sample space}}$$

- **Probability Mass Function (pmf):** Let R be the set of all possible values of a discrete random variable X; and f(k) = P(X = k) for each k in R. Then f(k) is called the probability mass function of X. The expression P(X = k) means the probability that X assumes the value k.

  Example:- Suppose a fair coin is flipped 3 times. Let X denote the number of heads that can be observed out of these three flips. Then X is a discrete random variable with the set of possible values {0, 1, 2, 3}; this set is also called the support of X. Probable outcomes constitute the sample space and here it is $2^3 = 8$. They are-

  {HHH; HHT;HTH; THH;HTT; THT; TTH; TTT}

  All the possibilities are equally likely and its chance of occurrence is 1/8. Now, let's analyse the result to understand PMF.

  Chance of occurring No head: 1/8 i.e. TTT

  Chance of occurring 1 head: 3/8 i.e. HTT, THT and TTH

  Chance of occurring 2 head: 1/8 i.e. HHT, HTH and THH

  Chance of occurring 3 head: 1/8 i.e. HHH

  The probability distribution of *X* can be obtained similarly and is given below:

K :	0	1	2	3
P(X= k)	1/8	3/8	3/8	1/8

So, the PMF can be defined as-

$$P(X=k) = \binom{3}{k}\left(\frac{1}{2}\right)^{k}\left(1-\frac{1}{2}\right)^{3-k}, k = 0, 1, 2, 3$$

As the outcome of each flip is binomial type, so it is also known as binomial (3, 1/2) mass function.

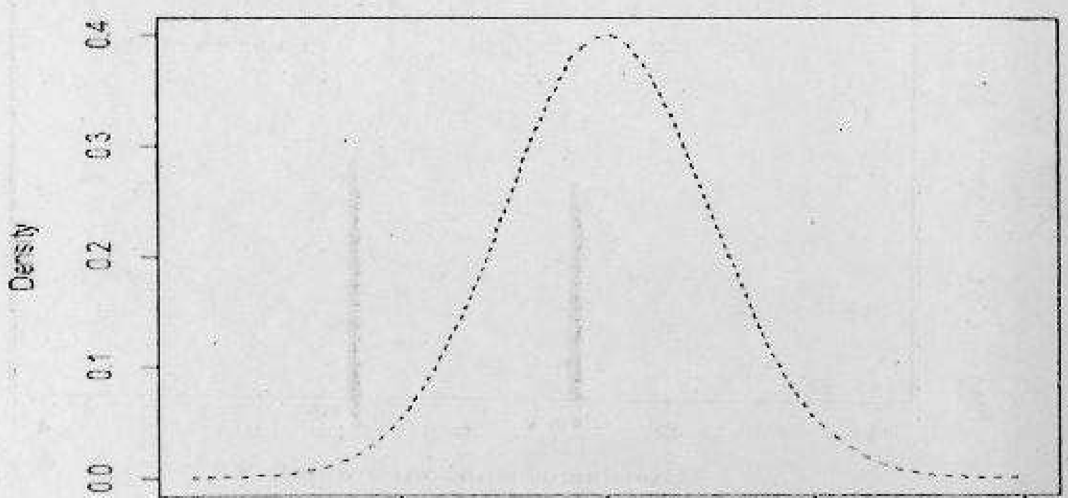
- **Continuous Random Variable:** If the set of all possible values of X is an interval or union of two or more non-overlapping intervals, then X is called a continuous random variable. Example:- Height of students in a class, weight of students in a class etc.

- **Probability Density Function:-** A continuous random variable takes on an uncountably infinite number of possible values. For a discrete random variable X that takes on a finite or countably infinite number of possible values, we determined P(X = x) for all of the possible values of X, and called it the probability mass function ("p.m.f."). For continuous random variables, the probability that X takes on any particular value x is 0. That is, finding P(X = x) for a continuous random variable X is not going to work. Instead, we'll need to find the probability that X falls in some interval (a, b), that is, we'll need to find P(a < X < b). We'll do that using a probability density function ("p.d.f."). For continuous random variable X, probability distribution function of X is a function f(x) such that any two numbers a and b with $a \leq b$,

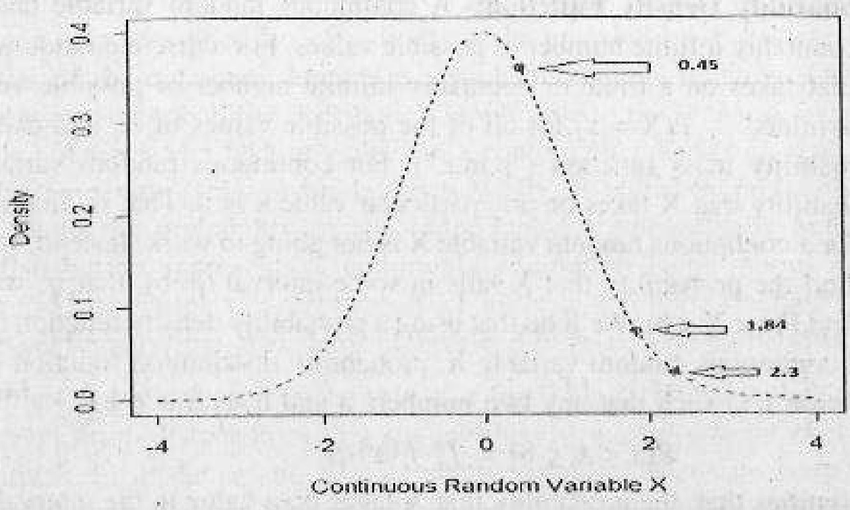
$$P(a \leq x \leq b) = \int_a^b f(x)\,dx$$

It signifies that, the probability that X takes on a value in the interval [a, b] is actually the area under the graph of the density function. That's why, this graph is also known as density graph.

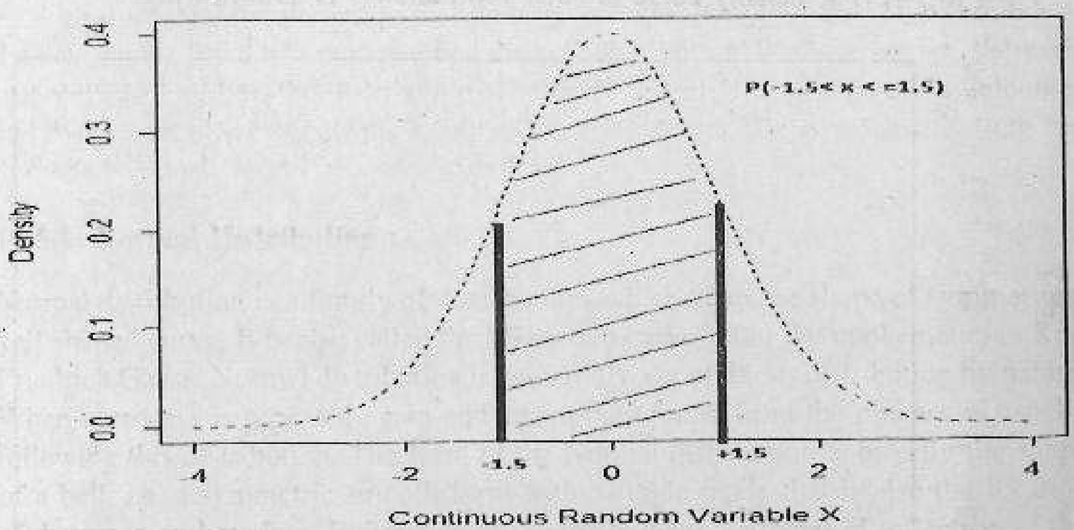
Example:- Suppose 100 temperature readings are taken within an interval [-4 to +4]. The density curve of such 100 numbers is as following-



Now, I want to see 3 points such as 0.45, 1.84 and 2.3 on the curve. These are the peak values lying on the curve.



Now the question is what is the probability that the temperature lies between -1.5 to 1.5? The answer can be found from the following graph-



The shaded region is actually area given within the window span. Here area denotes the probability.

Example-2:- From a library, issuing pattern of users are surveyed as against the age of books.

Age of book	0-1	1-2	2-3	3-4	4-5
Frequency of issue	35	63	56	45	21

Now, derive the probability of issuing books aged between 0-4 years?

Solution:- First investigate the data. Total number of issues is 220. Now, calculate the probability of each frequency group which is displayed here in the table-

Age of book	0-1	1-2	2-3	3-4	4-5
Frequency of issue	35	63	56	45	21
Probability	35/220= 0.15	63/220= 0.29	56/220= 0.2545	45/220= 0.2045	21/220= 0.09545

Probability of issuing books aged between 0-4 years –

P(0d" X d"4) = 0.15 + 0.29 + 0.2545+ 0.2045 = 0.899

10.5 Family of Distributions

Distribution can be defined as a stream of data or a data set which are continuous or discrete in nature. In another word, "*distribution* of a statistical data set (or a population) is a listing or function showing all the possible values (or intervals) of the data and how often they occur". In probability and statistics, a probability distribution is a mathematical function that, stated in simple terms, can be thought of as providing the probability of occurrence of different possible outcomes in an experiment. For instance, if the random variable X is used to denote the outcome of a coin toss ('the experiment'), then the probability distribution of X would take the value 0.5 for {X=Heads} and 0.5 for {X=Tails}.

In more technical terms, the probability distribution is a description of a random phenomenon in terms of the probabilities of events. Examples of random phenomena can include the results of an experiment or survey. A probability distribution is defined in terms of an underlying sample space, which is the set of all possible outcomes of the random phenomenon being observed. The sample space may be the set of real numbers or a higher-dimensional vector space, or it may be a list of non-numerical values; for example, the sample space of a coin flip would be {Heads, Tails}.

Probability distributions are generally divided into two classes. A discrete probability distribution (applicable to the scenario where the set of possible outcomes is discrete, such as in a coin toss or a flip of a dice) can be encoded by a discrete list of the probabilities of the outcomes, known as a probability mass function. On the other hand, a continuous probability distribution (applicable to the scenarios where the set of possible outcomes can take on values in a continuous range (e.g., real numbers), such as the temperature on a given day) is typically described by probability density functions (with the probability of any individual outcome actually being 0). The normal distribution represents a commonly encountered continuous probability distribution. More complex experiments, such as those involving stochastic processes defined in continuous time, may demand the use of more general probability measures.

Example:- A typical project analysis at a firm. Most estimates that go into the analysis come from distributions that are continuous; market size, market share and profit margins, for instance, are all continuous variables. There are some important risk factors, though, that can take on only discrete forms, including regulatory actions and the threat of a terrorist attack; in the first case, the regulatory authority may dispense one of two or more decisions which are specified up front and in the latter, you are subjected to a terrorist attack or you are not.

With discrete data, the entire distribution can either be developed from scratch or the data can be fitted to a pre-specified discrete distribution. With the former, there are two steps to building the distribution. The first is identifying the possible outcomes and the second is to estimate probabilities to each outcome. Here two distributions are discussed- Normal and Poisson distribution.

10.5.1 Normal Distribution :

Normal distribution is a family of distributions which takes the shape of symmetrical bell shaped curve. It is also called the "Gaussian curve" after the mathematician Karl Friedrich Gauss. Normal distribution is essentially a continuous distribution by nature. When a process is repeated again and again, data found from the process is usually following this distribution. The form of the Normal distribution is broadly the shape of a bell, i.e. a symmetric smooth form with a single mode that is also the location of the mean and median. Either side of the mode there is a point of inflection of the bell curve which is one unit (one standard deviation) from the mean (illustrated by the horizontal line in the graph above). Beyond this point the curve extends towards the x-axis asymptotically, with a theoretical extent to infinity in both directions. Example:- machine fills rice in sacks, IQ of population, temperature of normal human being, amount of water consumed by a human being on daily basis and so on.

As normal distribution is symmetric, f(x) = f(-x) and density function of the normal probability distribution is –

$$y = f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

The value of x lies between -" to +". Here, σ is the standard deviation and mean is μ.

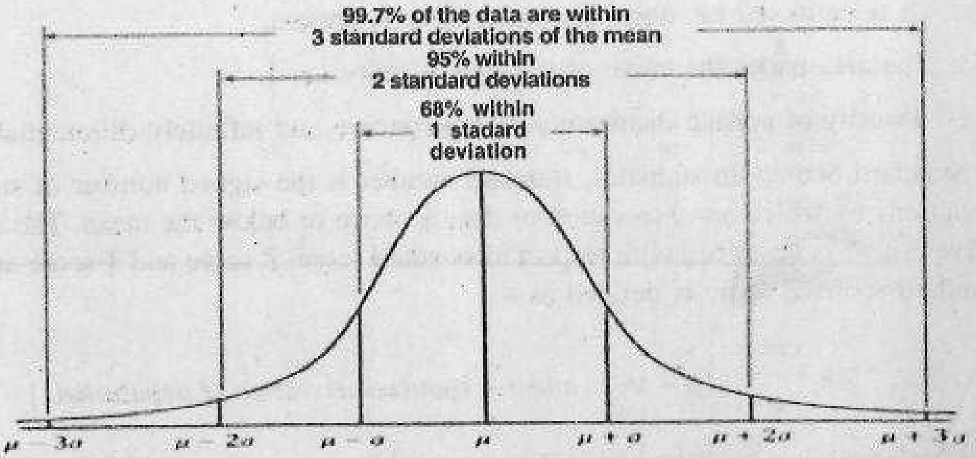


Figure:- Normal Distribution

A perfect normal curve usually has the same mean, median and mode in the population. 50% of the cases fall above the central point and the rest 50% fall below the central point. If the measures of central tendency are not equal, the distribution is distorted which are measured by skewness and curtosis. The central point of the curve is the mean.

$$\text{Let } t = \frac{x-\mu}{\sigma}$$

The variable t is called as standard normal variate. It can be easily verified that, mean(μ_t) =0 and standard deviation σ_t =1. The probability that an observation lies between t and t+dt is given by-

$$f(t) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2}t^2\right] dt$$

The probability that x lies between ($\mu-\sigma$) and ($\mu+\sigma$) is 68.27%

The probability that x lies between ($\mu-2\sigma$) and ($\mu+2\sigma$) is 95.45%

The probability that x lies between ($\mu-3\sigma$) and ($\mu+3\sigma$) is 99.73%

In practice, normal distribution is largely used to find the probabilities

Some of the properties of normal distribution are-

a. It is unimodal i.e. only one mode value is present.

b. The area under the curve over x-axis is unity i.e. 1.

c. Density of normal distribution is log-concave and infinitely differentiable.

Standard Score:- In statistics, standard number is the signed number of standard deviations by which an observation or data is above or below the mean. The normal curve is always described with respect to standard score. Z score and T score are such standard score. Z score is defined as –

$$Z = \frac{x-\mu}{\sigma} \quad [\mu = Mean\ and\ \sigma = standard\ deviation\ of\ population\ ]$$

The deviation is measured from the mean and the z score is scaled below or above the mean point. Always a predictive interval is set (L= lower limit and U= Upper limit) within limits. For Z of x is given by-

$$P\left(\frac{L-\mu}{\sigma} < Z < \frac{U-\mu}{\sigma}\right) = Y$$

Z-score allows to decide whether the sample mean and deviation is significantly different from population or not. For, a Z-score one has to know- population mean, population standard deviation, sample mean, sample deviation. For large value of n of a sample(if this follows normal distribution), with the help central limit theorem, standard deviation to be taken as σ D ("n) .

Another standard score in practice is T-scores. It is difficult to know mean and standard deviation of entire population before-hand. Statisticians try to estimate mean and variance from samples with those of population to an optimized mark. Knowing a value as reference point does not always comply with comparison while testing against population sub-set or unknown experiment. For an unknown distribution, even by making a normal curve, Z-score can intervene little with an limitation upto two sigma(σ)level. But if for this unknown distribution mean is shifted new value probability of existence with the help of normal curve might pose an erroneous result that's why T-score is used in such cases.

$$T = \frac{x - \bar{X}}{s} \quad [\bar{X} \text{ is sample mean and s is sample standard deviation}]$$

Confidence Interval and Level of Significance:-

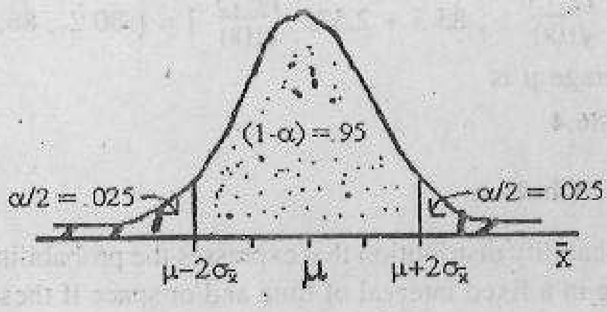
Confidence interval is a type of interval estimate of a population. If an experiment is repeated , the data as output are not same for each test. Confidence interval signifies a range of values that take a close approximate of unknown population parameters. A 90% confidence level means that we would expect 90% of the interval estimates to include the population parameter; A 95% confidence level means that 95% of the intervals would include the parameter; and so on. **A 95% confidence interval does not mean that for a given realised interval calculated from sample data there is a 95% probability the population parameter lies within the interval, nor that there is a 95% probability that the interval covers the population parameter.** For a variable asymptotically following a normal distribution with mean zero and one standard deviation and H_0 is true, for large n -

$$P\left\{\left|\frac{x-\mu}{\sigma/\sqrt{n}}\right| \le 1.96\right\} = 0.95,\ P\left\{\left|\frac{x-\mu}{\sigma/\sqrt{n}}\right| \le 1.65\right\} = 0.90 \text{ and } P\left\{\left|\frac{x-\mu}{\sigma/\sqrt{n}}\right| \le 2.58\right\} = 0.99$$

A general form of representing such Z value is-

$$P\left\{\left|\frac{x-\mu}{\sigma/\sqrt{n}}\right| \le z_{\alpha/2}\right\} = 1 - \alpha$$

The 95% confidence interval for μ



From the above picture it is understood that, if the value x under consideration is found to be within the dotted region within the curve it is found to be within 95% confidence interval. If the value z lies in the lined region then it is known as critical region. The limit of the interval it is found that, the lower limit is (μ-2σ) and upper limit is (μ+2σ). α is an arbitrary number which is selected to determine the critical region and also known as level significance of the test. If the computed value of z lies in the critical region with 95% confidence interval, it signifies null hypothesis is to be rejected [z: |z| > 1.96].

$$-1.96 \leq \frac{x-\mu}{\sigma/\sqrt{n}} \leq 1.96$$

It contains two algebraic inequalities which are equivalent to

$$\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96\frac{\sigma}{\sqrt{n-1}}$$

This states that it is with 95% confident the population average (μ) lies in the interval

$$\left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}};\ \bar{x} + 1.96\frac{\sigma}{\sqrt{n-1}}\right]$$

A 99% confidence interval is similarly given by

$$\left[\bar{x} - 2.575\frac{\sigma}{\sqrt{n}};\ \bar{x} + 2.575\frac{\sigma}{\sqrt{n-1}}\right]$$

Example-1:-

A sample of 101 books in library gives an average length of the reference list of 83.3 references (hen.e . The standard deviation is $s = 11.15$. Determine a 99% confidence interval for the real average length of a reference list over all books in the library.

Solution

The confidence interval we want is given by –

$$[\ 83.3 - 2.575\ \frac{12.15}{\sqrt{100}}\ ;\ 83.3 + 2.575\ \frac{12.15}{\sqrt{100}}\ ] = [\ 80.2\ ;\ 86.4]\ ,$$

So the real average μ is

80.2 d" μ d" 86.4

10.5.2 Poisson Distribution

"It is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event" . Here, the probability of occurring an event(p) is low and the event can occur 0,1,2,3,4.....n times within an time interval. The average number of events in an interval is called event rate and it's denoted as lambda (λ). So, the probability of occurring an event k in an interval is given by (Probability Mass Function)–

P(K event in an interval) = $\frac{n^{-1}\lambda^{k}}{k!}$ **k=0,1,2,3,4,5.........**

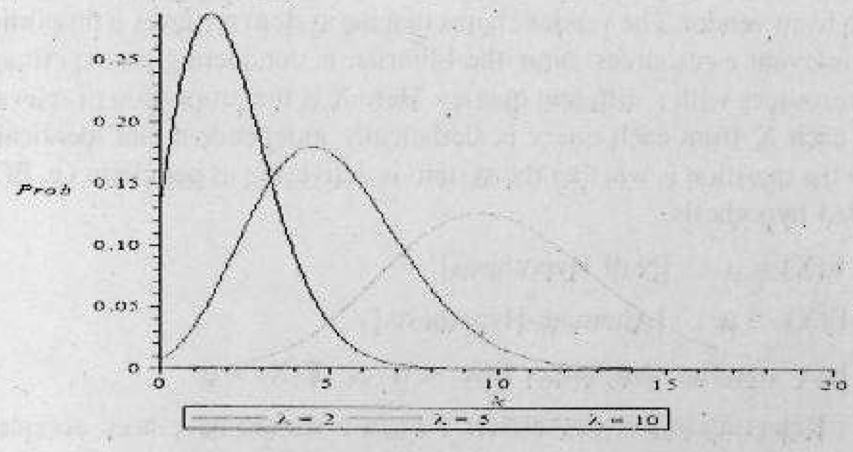
λ is the average number of events per interval (event rate)
e is euler's number, the base of natural logarithm
K! = k * (k-1) * (k-2) * (k-3)......3 * 2 * 1

Example

1. Number of meteors striking air layer with diameter greater than 5 metre.
2. Number of road accidents in X road in a year
3. Number of patients coming to hospital with burn injury
4. Decay rate of radioactive elements or half-life of radioactive elements
5. Number of time a book issued in a library in a particular year
6. Number of requests of a journal by researchers in May,2015
7. Number of books re-issued in X subject from Y department and so on.

In Poisson Distribution, though p is very small but np= λ is not negligible. More precisely,

Mean (μ) = $\sum_{k} kP(k) = \lambda$ (Event Rate)



Problem-1:- Average number of flats sold per month in a town is 40. What is the probability that, 50 flats would be sold in next month?

Solution: Here, $\lambda = 40$ and the event expected to occurred is k= 50. So the equation

is reduced to, $P(K=50) = \dfrac{e^{-\lambda}\lambda^{k}}{k!} = \dfrac{e^{-40}\ 40^{50}}{50!}$

$$= \frac{2.71828^{-40} * 40^{50}}{50!} = \mathbf{0.01770}$$

Ans. The probability of selling 50 (k=50) flats in next month is 0.01770

Thus the probability of selling 60 (k= 60) flats in next month is 0.000678

The probability of selling 30 (k= 30) flats in next month is 0.0184654

10.6 Hypothesis testing

Hypothesis is an assumption which may or may not be true. Statisticians use this to infer about a population with the information from the sample at hand. Methods for inferring involve statistical test about hypothesis with the laws of probability and theoretical distributions. The statement as a form of hypothesis to be put on statistical test is known as null hypothesis. If null hypothesis is found false, an alternate statement is to be stated to establish the outcome from the statistical test and this is known as alternate hypothesis. For example, suppose a librarian wants to buy a special resource discovery system from vendor. The vendor claims that the system retrieves μ proportion (on average) of relevant e-resources. Now the librarian is conducting an experiment by retrieving e- resources with n different queries. Here X is the proportion of relevant e-resources and each X_i from each query is statistically independent and identically distributed. Now the question is whether the system is delivering as per claim i.e. $E(X) = \mu$. This is called hypothesis.

H_0 : $E(X) = \mu$ [Null Hypothesis]

H_1 : E(X) "' μ [Alternate Hypothesis]

Not equal to here signifies that, either $E(X) < \mu$ or $E(X) > \mu$.

Type-1 Error : Rejecting null hypothesis(H_0) when it should have been accepted.

Type-2 Error : Accepting H_0 when it should have been accepted.

In hypothesis testing, null hypothesis is tested (E(X) = μ) against one of the following alternate hypothesis testing:

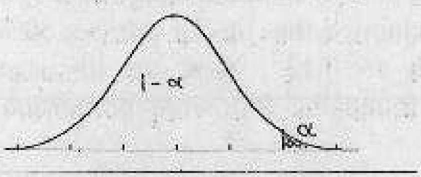
a. E(X) < μ

b. E(X) > μ

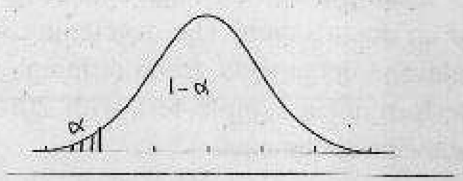
Under such conditions, critical values are given by-

a. $P\left\{\left|\frac{x-\mu}{\sigma/\sqrt{n}}\right| \geq -z_\alpha\right\} = 1-\alpha$

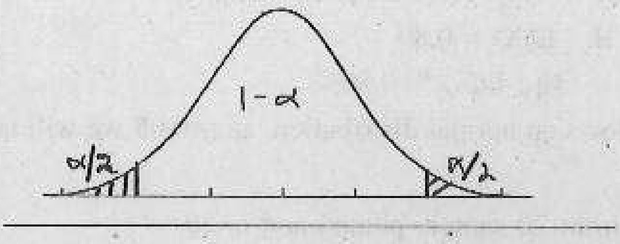
b. $P\left\{\left|\frac{x-\mu}{\sigma/\sqrt{n}}\right| \leq z_\alpha\right\} = 1-\alpha$



a) E(X) > μ



b) E(X) < μ



c) H_1 : E(X) ''' μ

If the critical region is chosen such that it lies either in left side or right side of the curve, then it is known as one-sided test. Otherwise it is called two-sided test(see figure-c).

Testing of significance with single known mean

Example-1:- In a library, average time taken to catalogue a book is 18 minutes with σ=5 minutes. Now, a sample of 36 books are taken and $\bar{x}$ = 20 minutes. Determine by statistical test whether the sample is too differing from the population mean? [Given at 95% CI, $|Z_{\alpha/2}| = 1.96$]

Solution:- Z-test can be used here to determine the difference.

$H_0 : \mu = 18$

$H_1 : \mu$ "' 18

$$Z = (\bar{x} - \mu)/(\sigma / \sqrt{n}) = \frac{20-18}{5/\sqrt{36}} = \frac{2*6}{5} = 2.4$$

Now, $|Z| > |Z_{\alpha/2}|$, calculated Z value is greater than theoretical Z value and therefore the value is lying at the critical region. In effect null hypothesis is rejected. It is concluded that the sample is showing considerable difference in its mean from the population mean.

Example-2:- A special search engine is inducted in information management system of an organisation. The developing vendor have claimed that, it can retrieve 80% of relevant documents from different widgets with $\sigma = 0.13$. Now, the librarian is performing a sample test with 20 queries and found the following proportion of relevant documents-

0.85, 0.89, 0.87, 0.96, 0.87, 0.81, 0.78, 0.83, 0.65, 0.78, 0.75, 0.89, 0.99, 0.69, 0.76, 0.87, 0.80, 0.80, 0.81, 0.78.

What the librarian would decide from the sample?

Solution:- $H_0 : E(X) = 0.80$

$H_1 : E(X)$ "' 0.80

Assuming X follows an normal distribution, at $\alpha=0.05$ we will tally $|Z|$ value with $|Z_{\alpha/2}| = 1.96$

Here = 0.82 (from 20 sample points) and n=20.

$$Z = (\bar{X} - \mu) / (\sigma / \sqrt{n}) = \frac{0.82 - 0.80}{0.13 / \sqrt{20}} = 0.688$$

Now $|Z| <$ ||, therefore null hypothesis is accepted i.e company's claim is not different from the sample mean and the system is performing satisfactorily.

Testing of significance between two means:-

Problem-1:- A new procedure has been implemented in a library to deal with reference services and training was given to a group of library staffs. Staffs are divided in two groups. One group works with old method and the second group works with new method and the efficiency is to be determined from the working efficiency from two methods.

1st Group(Old Method) processes 100 requests in a mean time of 6 minutes and s.d. 2 minutes

2nd Group(New Method) processes 150 requests in a mean time of 4.5 minutes and s.d. 1.7 minutes

Test for the significance of the difference between mean times (at α=0.05, || = 1.96).

Solution:- For old method: = 6 minutes and σ_1 = 2 minutes

For new method: = 4.5 minutes and σ_2 = 1.7 minutes

$$H_0 : = \quad \bar{x}_1 = \bar{x}_2$$

$$H_1 : \quad \bar{x}_1 \neq \bar{x}_2$$

For 2 sample mean test, $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}\right)}}$

$$Z = \frac{6-4.5}{\sqrt{\frac{4}{100} + \frac{2.39}{150}}} = \frac{1.5}{\sqrt{0.04+0.0192}} = \frac{1.5}{0.2433} = 6.1652$$

As |Z| > ||, null hypothesis is rejected and alternate hypothesis is accepted. There is significant difference between the means of two samples.

Testing of significance for unknown σ (Student t-test):-

In most of the situations the variance (σ^2) of the population(from where sapling is done) can't be known beforehand. In such cases σ^2 is estimated from the sample. Instead of using σ^2, sample variance is termed as s^2 . Though s^2 is not a reliable estimate of σ^2

$t = \dfrac{x-\mu}{s/\sqrt{n}}$ [This is known as student t-test]

Example-1:- In a library, the average price of civil engineering books is Rs. 1400/-.

Now, a sample of 40 civil engineering books are selected and found that average price is 1650/- and standard deviation(s) is 70. Now determine whether the claim of average price Rs. 1400/- is true or false.

Solution:- From the given problem s = 70 and =1650 and μ = 1400

$$H_0 : E(x) = 1400$$

$$H_1 : E(x) \text{ ``' } 1400$$

At α = 0.05 with 39 degrees of freedom, = 2.023 and | t | >

Therefore null hypothesis is rejected and alternate hypothesis is accepted.

Conclusion:- There is a significant difference between population mean and sample mean. It is quite unlikely that the average price of civil engineering books is Rs. 1400/- .

Non-Parametric Tests :

Nonparametric statistics are statistics not based on parameterized families of probability distributions. They include both descriptive and inferential statistics. The typical parameters are the mean, variance, etc. Unlike parametric statistics, nonparametric statistics make no assumptions about the probability distributions of the variables being assessed. The difference between parametric models and non-parametric models is that the former has a fixed number of parameters, while the latter grows the number of parameters with the amount of training data.

In statistics, the term "non-parametric statistics" has at least two different meanings:

The first meaning of non-parametric covers techniques that do not rely on data belonging to any particular distribution. These include, among others: distribution free methods, which do not rely on assumptions that the data are drawn from a given probability distribution. As such it is the opposite of parametric statistics. It includes non-parametric descriptive statistics, statistical models, inference and statistical tests. Non-parametric statistics (in the sense of a statistic over data, which is defined to be a function on a sample that has no dependency on a parameter), whose interpretation does not depend on the population fitting any parameterised distributions. Order statistics, which are based on the ranks of observations, are one example of such statistics and these play a central role in many non-parametric approaches.

The second meaning of non-parametric covers techniques that do not assume that the structure of a model is fixed. Typically, the model grows in size to accommodate the complexity of the data. In these techniques, individual variables are typically assumed to belong to parametric distributions, and assumptions about the types of connections among variables are also made. These techniques include, among others: non-parametric regression, which refers to modelling where the structure of the relationship between variables is treated non-parametrically, but where nevertheless there may be parametric assumptions about the distribution.

10.7 Correlation

The quantity called coefficient of correlation measures the strength and direction of relationship two variables X and Y. Concomitant variation expressed through Correlation is slanted to statistical reasoning. It tries to answer several queries such as – 1) The 2 variables have an relation which can be led to any predictive direction? 2) Do they vary in any direction?

The quantity r, called the linear correlation coefficient, measures the strength and the direction of a linear relationship between two variables. The linear correlation coefficient is sometimes referred to as the Pearson product moment correlation coefficient in honor of its developer Karl Pearson. Mathematical formula for Pearson Correlation is-

$$r = \frac{n\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{n\left(\sum x^2\right) - \left(\sum x\right)^2}\sqrt{n\left(\sum y^2\right) - \left(\sum y\right)^2}}$$

Here **n** is the number of data points of x and y variables

$\sum xy$ is sum of the product of paired score of x and y

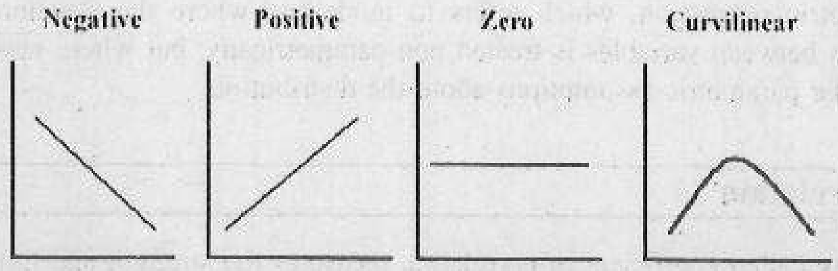
$\sum x$ sum of x score

$\sum y$ sum of y score

$\sum x^2$ sum of squared x scores

$\sum y^2$ sum of squared y scores

The value of correlation r varies between -1 to +1 (-1 d" r d" +1). Positive sign (+) indicates positive correlation and negative value (-) indicates negative correlation between two variables (x and y). Correlation measures the extent to which it confirming to a linear relation of the type y= a+bx.



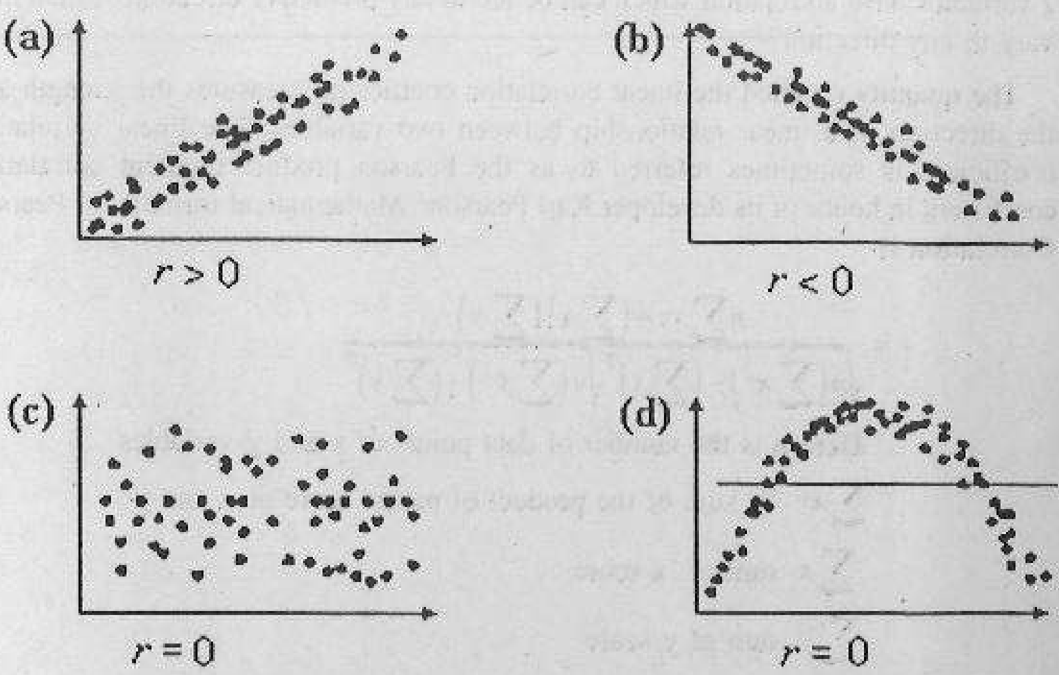


Figure:- Visualization of data with respect to values of r

Example:- In a library, every year hundreds of new books are added to existing stock. Number of issues of books to students can be also found after calculating the issue register. Library managers want to investigate whether addition of new books have any relationship with book borrowing mentality of students? Data is taken for a consecutive 10 year interval.

Addition of books (x)	**350**	**400**	**600**	**580**	**1225**	**2500**	**1455**	**2365**	**1489**	**1648**
Books i sued (y)	**1200**	**996**	**789**	**1489**	**2456**	**3658**	**2648**	**4632**	**1120**	**1789**

Solution:- To determine the existing relationship between two variables, pearson's correlation can be determined to understand the underlying nature.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2}\sqrt{n(\sum y^2) - (\sum y)^2}}$$

Here, **n = 10,** $\sum xy$ **= 33732492,** $\sum x$ **= 12612,** $\sum y$ **= 20777,** $\sum x^2$ **= 21372800,** $\sum y^2$ **= 57606807**

Putting the values in the equation, r = 0.8474111. Between two variables there exists a positive correlation and it can be easily concluded that addition of new books have a positive relationship or impact on readership among the students.

10.8 Non-Parametric Tests

Z-tests and t-test are done after an assumption that the sample data are from either normal or binomial population. For a large sample it tends to follow normal or binomial distribution. Null hypothesis is always formulated to test against population means or median or proportions or variance. But in many situations, it is essential to judge

1. Analysis of similarities
2. Statistical dependence between variables
3. tests for differences in scale between two groups
4. equality of two distributions by using ranks
5. treatments in randomized block designs with 0/1 outcomes have identical effects

There are some situations when it is clear that the outcome does not follow a normal distribution. These include situations:

- when the outcome is an ordinal variable or a rank,
- when there are definite outliers
- when the outcome has clear limits of detection.
- When categorical data are present

Advantages of Nonparametric Tests

Nonparametric tests have some distinct advantages. With outcomes such as those described above, nonparametric tests may be the only way to analyze these data. Outcomes that are ordinal, ranked, subject to outliers or measured imprecisely are difficult to analyze with parametric methods without making major assumptions about their distributions as well as decisions about coding some values (e.g., "not detected"). As described here, nonparametric tests can also be relatively simple to conduct.

There are several Nonparametric Tests such as –

- Chi-square Test
- Kolmogorov- Smirnov Test
- Wilcoxon signed-rank test
- Kendall's tau test
- Kendall's W test
- Mann-Whitney U test etc.

Chi-Square Test for independence:- The test is applied when you have two categorical variables from a single population. It is used to determine whether there is a significant association between the two variables. For example, in an election survey, voters might be classified by gender (male or female) and voting preference (Political Party1, Political Party2, Political Party3). Chi-square (χ^2) test for independence to determine whether gender is related to voting preference.

When to Use Chi-Square Test for Independence

The test procedure described in this lesson is appropriate when the following conditions are met:

- The sampling method is simple random sampling.
- The variables under study are each categorical.
- If sample data are displayed in a contingency table, the expected frequency count for each cell of the table is at least 5.

The null hypothesis states that knowing the level of Variable A does not help you predict the level of Variable B. That is, the variables are independent.

H_0: Variable A and Variable B are independent. H_a: Variable A and Variable B are not independent.

Example:- A survey has been conducted on 100 students of an university to see whether there is any relationship between gender and preference of library use. Out of 100 students, 50 are male and rest are female. The question is- is there any relationship between library use and gender. The data is given in the following table.

Library Use	User	Non-User	Total
Sex			
Male	25	25	50
Female	18	32	50
Total	**43**	**57**	**100**

Solution:- The solution to this problem takes four steps:

(1) state the hypotheses, (2) formulate an analysis plan, (3) analyze sample data and (4) interpret results.

Step-1:- Hypothesis statement

H_0 : Gender and library uses are independent

H_1: Gender and library uses are not independent

Step-2:- Analysis Plan

For this analysis, using sample data, chi-square test for independence would be conducted as categorical data is present and we can not assume whether they would follow binomial or normal distribution.

Step-3:- Analyze sample data

Library Use	User (c_1)	Non-User (c_2)	Total
Sex			
Male(r_1)	26	24	50
Female(r_2)	18	32	50
Total	**44**	**56**	**100 (N)**

Degree of Freedom- df = (r-1)* (c-1) [r= number of row, c=number of column]

Df = (r-1)*(c-1) = (2-1)*(2-1) =1

Expected Values - $E_{ij} = \frac{r_i * c_j}{N}$

r_i is the total number of observations in the *i*th row

c_j is the total number of observations in the *j*th row and N is the total observations

E_{ij} is the expected/theoretical frequencies in *ij*th cell (ith row and jth column).

$E_{11} = \frac{44 + 50}{100} = 22$ [expected number of male users]

$E_{12} = \frac{56 + 50}{100} = 28$ [expected number of male non-users]

$E_{21} = \frac{44 + 50}{100} = 22$ [expected number of female users]

$E_{22} = \frac{56 + 50}{100} = 28$ [expected number of female non-users]

Now reconstitute the table with expected values-

Library Use	User(c_1)	Non-User(c_2)	Total
Sex			
Male(r_1)	O_{11}= 26 E_{11} = 22	O_{12}= 24 E_{12} = 28	50
Female(r_2)	O_{21}= 18 E_{21} = 22	O_{22} = 32 E_{22} = 28	50
Total	44	56	100 (N)

Here for our understanding, observed values within cells are denoted likewise notations.

$$x^2 = \sum_i \sum_j \frac{\left(0_{ij} - E_{ij}\right)^2}{E_{ij}}$$

$$So,\ x^2 = \frac{(26-22)^2}{22} + \frac{(24-28)^2}{28} + \frac{(18-22)^2}{22} + \frac{(32-28)^2}{28}$$

$$= 0.72 + 0.57 + 0.72 + 0.57$$

$$x^2 = 2.58$$

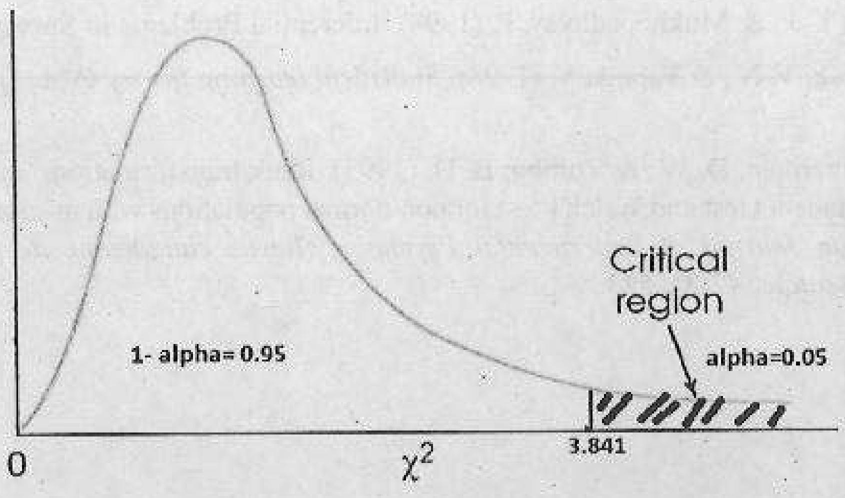


Figure:- Chi-square distribution with 1 degree of freedom

As the χ^2 value is less than critical value (i.e. 3.841) , value lies within the acceptance region. So, the null hypothesis is accepted. Gender and library uses are independent and there is no significant difference among male or female w.r.t. library uses.

10.9 References

Feller, W. (2008). *An introduction to probability theory and its applications* (Vol. 2). John Wiley & Sons.

Lancaster, H. O., & Seneta, E. (2005). *Chi Square Distribution*. John Wiley & Sons, Ltd.

Lawley, D. N. (1938). A generalization of Fisher's z test. *Biometrika*, *30*(1/2), 180-187.

Mukhopadhyay, P. (2008). *Theory and methods of survey sampling*. PHI Learning Pvt. Ltd..

Rao, IKR.(1983). *Quantitative Methods for Library and Information Science*. New Delhi: New Age.

Rao, T. J., & Mukhopadhyay, P. (1998). Inferential Problems in Survey Sampling.

Vapnik, V. N., & Vapnik, V. (1998). *Statistical learning theory* (Vol. 1). New York: Wiley.

Zimmerman, D. W., & Zumbo, B. D. (1993). Rank transformations and the power of the Student t test and Welch t'test for non-normal populations with unequal variances. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *47*(3), 523.

Unit 11 ❑ Statistical Analysis Using MS Excel

Structure

11. 1. Introduction

A spreadsheet is essentially a matrix of rows and columns. Consider a sheet of paper on which horizontal and vertical lines are drawn to yield a rectangular grid. The grid namely a cell, is the result of the intersection of a row with a column. Such a structure is called a Spreadsheet.

A spreadsheet package contains electronic equivalent of a pen, an eraser and large sheet of paper with vertical and horizontal lines to give rows and columns. The cursor position uniquely shown in dark mode indicates where the pen is currently pointing. We can enter text or numbers at any position on the worksheet. We can enter a formula in a cell where we want to perform a calculation and results are to be displayed. A powerful recalculation facility jumps into action each time we update the cell contents with new data.

Microsoft Excel is a big worksheet (it can take data rows in thousands across 256 columns). This worksheet can be used for data entry and for performing calculations by click of buttons. MS Excel can be used to create tables and graphs and perform statistical calculations. The work done in MS Excel can be easily copied and pasted to many window-based programs for further analysis. Excel is clearly not an adequate statistics package because many statistical methods are simply not available. This lack of functionality makes it difficult to use it for more than computing summary statistics and simple linear regression and hypothesis testing. The MS Excel worksheet is a collection of cells. There are 65,000 (rows) X 256 (columns) cells in an MS Excel worksheet. Each row or column can be used to enter data belonging to one category. Data entry in MS Excel is as simple as writing on a piece of paper. MS Excel assigns each column a field depending upon the type of data. It supports various data formats; one can choose a data format by formatting the cells.

11. 2. Toolbar and Icons

The Excel worksheet consist a standard menu bar and different tools as follows:

TOOLBARS AND THE ICONS

Standard Toolbar

New
Creates a new blank file.
Alternative: File → New
Shortcut → Ctrl + N

Open
Opens or finds a file.
Alternative: File → Open
Shortcut → Ctrl + O

Save
Saves the active file.
Alternative: File → Save
Shortcut → Ctrl + S

E-Mail
Sends contents of the current worksheet as body of the e-mail message

Print
Prints the active file using current defaults.
Alternative: File → Print
Shortcut → Ctrl + P

Print Preview
Displays full pages as they are printed.
Alternative: File → Print Preview
Shortcut → Ctrl + F2

Spelling
Checks the spelling in the active file.
Alternative: Tools → Spelling and Grammer
Shortcut → F7

Cut
Cuts the selection and puts it on the Clipboard.
Alternative: Edit → Cut
Shortcut → Ctrl + X

Past
Inserts the Clipboard contents at the insertion point.
Alternative: Edit → Past
Shortcut → Ctrl + V

Copy
Copies the selection and puts it on the clipboard.
Alternative: Edit → Copy
Shortcut → Ctrl + C

Format Painter
Copies the formatting of the selection to a specified location.
Shortcut → Ctrl + Shift + C

Undo
Reverses certain commands.
Alternative: Edit → Undo
Shortcut → Ctrl + Z

Redo
Reverses the action of the Undo command.
Alternative: Edit → Repeat
Shortcut → Ctrl + Y

Insert Hyperlink
Displays the destination object, document or page.
Alternative: Edit → Repeat
Shortcut → Ctrl + K

Auto Sum
Adds numbers automatically with the SUM function. Excel suggests the range of cells to be added. If the suggested range is incorrect, drag through the range you want, and then press ENTER

11.3. Working with Data

Excel can work with ***text***, ***numbers, Booleans, dates***, ***times***, and ***formulas*** into a cell.

Text

Text is any combination of numbers, spaces, and non-numeric characters. For example, Excel treats the following entries as text: 10AA109, 127AXY, 12-976, 208 4675. All text is left-aligned in a cell by default.

Number

A number can contain only the following characters: **0 1 2 3 4 5 6 7 8 9 + -** ( ) /. All numbers are right-aligned in a cell by default. Excel ignores leading plus signs (+) and treats a single period as a decimal.

You can use the following characters to format the number: **, $ %**., e.g. **1,000, $13.25**, and **57%**. All other combinations of numbers and nonnumeric characters are treated as text.

Boolean

The values **TRUE** and **FALSE** are special values called Boolean values. A Boolean expression in a formula will evaluate to one of these two values. Note that surrounding either TRUE or FALSE with quotes, e.g. "TRUE", does ***NOT*** indicate a Boolean value, rather a text value.

Date and Time

The way that a date or time is displayed on a worksheet depends on the format applied to the cell. All dates and times are right-aligned in a cell by default. A date may be a *short date*, e.g. **11/27/2008**, or a *long date*, e.g. **Thursday, November 27, 2008**. Excel only recognizes dates staring from January 1, 1900. Also, Excel treats dates as numbers where 1/1/1900 is the same as 1, 1/2/1900 is the same as 2, and so on.

A time is formatted as hours, minutes, seconds, and AM/PM, e.g. **1:12:04 PM**. Times are also numbers, where each time is a number between 0 and 1. For example, midnight is 0 and noon is 0.5, and 11:59:59 PM is 0.999988425925926. If Excel cannot recognize a date or time that you enter, then it is treated as text.

Formula

A formula is a ***mathematical expression*** that *evaluates to a single value*. It can be used to perform a useful calculation in a cell. You can construct a formula using any mix of the following components:

Constants

A constant is a number or text value that does not change.

Mathematical Operators

A mathematical operator specifies a calculation. Common operators are: + (add), - (subtract), * (multiplication), and / (division).

Cell References and Cell Ranges

A cell reference in a formula evaluates to the value contained with that cell. A range of cells can be specified using the colon : symbol. A range can contain a row, a column, or a block of cells. As shown below, the syntax **A1:C1**, **A1:A5**, and **B2:C5** specify ranges along a row, a column, and a block of cells, respectively.

Functions

A ***function*** is a pre-defined formula in Excel that you can insert into your formula. Each function is specified by its name and input arguments and ***evaluates to a single value***.

Creating Formulas

To create a basic formula, use the following procedure:

1) Select with your mouse the cell in which you would like to insert your formula, e.g. **A1** in the

figures below.

2) All formulas must start with an = sign.

3) After you type the = sign, type your syntactically correct formula.

4) Click **Enter** when you are done.

If you type your formula directly into the active cell then it will be simultaneously displayed in the ***Insert Function box***. Alternatively, you may type your formula into the Insert Function box. After you press **Enter**, the cell will display the evaluated value of your formula and the Insert Function box will display the formula (see figure on the right).

11.4. MS Excel Menu

Excel can be used with confidence to obtain basic descriptive statistics, such as mean, median, mode, maximum, and minimum. All of these functions can be accessed through Excel's formula function.

To enter a formula, choose an empty cell. In this cell, type the equal sign "=". Whatever you type after the "=" is considered the formula. For example, you can type **= A1 + A2** and then press the Enter button. The cell will now display the sum of cells A1 and A2.

Excel also provides a **SUM** function, which allows you to calculate a sum for a range of cells. For example, to use the SUM function on the first ten rows of column A, type in an empty cell: **=SUM(A1:A10)** . You can use the SUM function on a row the same way:

=SUM(A1:M1) or can also use the SUM function on a contiguous block of cells, for example, rows 1-5 of columns A-M: **=SUM(A1:M5)** in the formula cell.

The formula interface can be used in exactly the same way on the following functions:

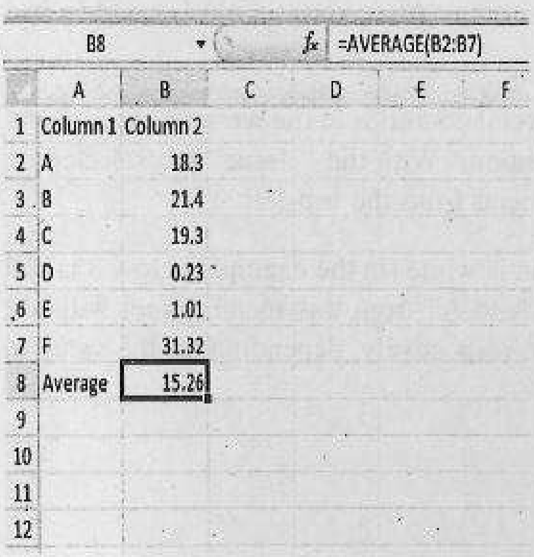
AVERAGE: the arithmetic mean of the selected data

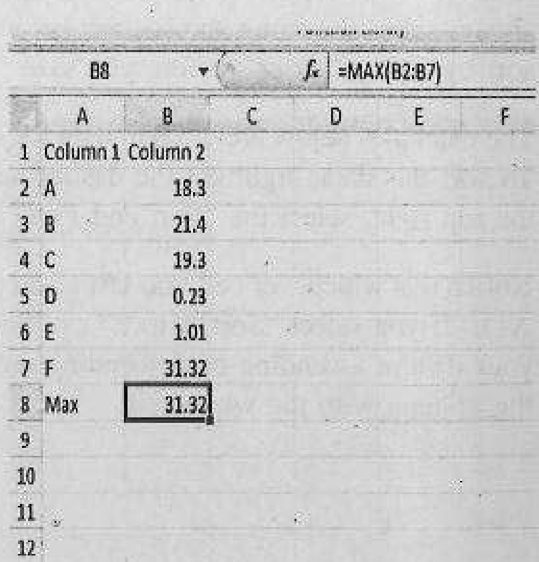
MEDIAN: the value at the 50th percentile of the selected data

MODE: the most commonly occurring value in the selected data

MIN: the smallest value in the selected data

MAX: the largest value in the selected data





B8 | f_x | =MIN(B2:B7)

	A	B	C	D	E	F
1	Column1	Column 2				
2	A	18.3				
3	B	21.4				
4	C	19.3				
5	D	0.23				
6	E	1.01				
7	F	31.32				
8	Max	0.23				
9						
10						
11						

B10 | f_x | =MODE(B2:B9)

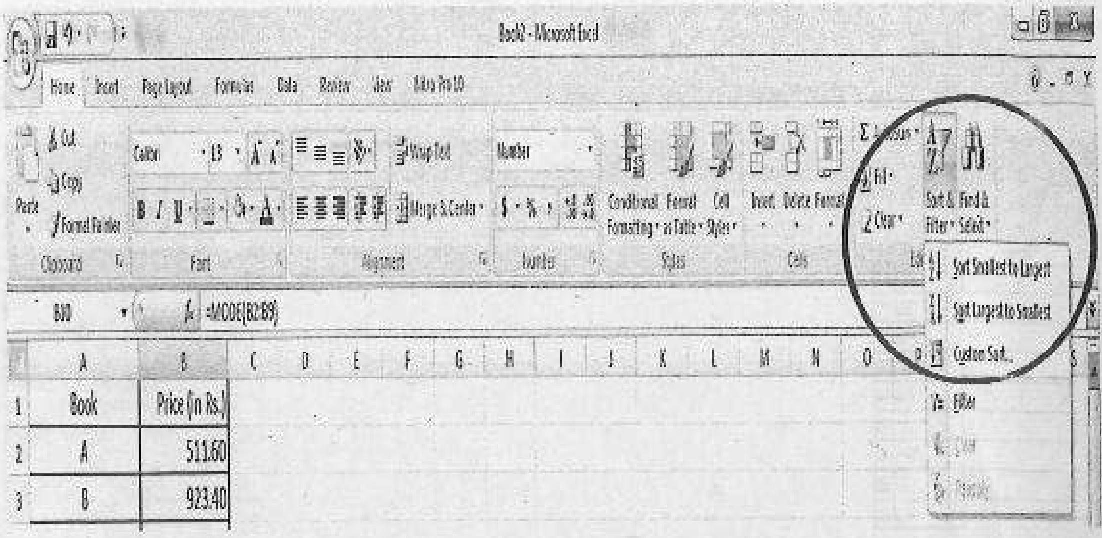
	A	B	C	D
1	Book	Price (in Rs.)		
2	A	511.60		
3	B	923.40		
4	C	918.30		
5	D	1143.60		
6	E	936.10		
7	F	1161.70		
8	G	923.40		
9	H	900.50		
10	MODE	923.40		
11				

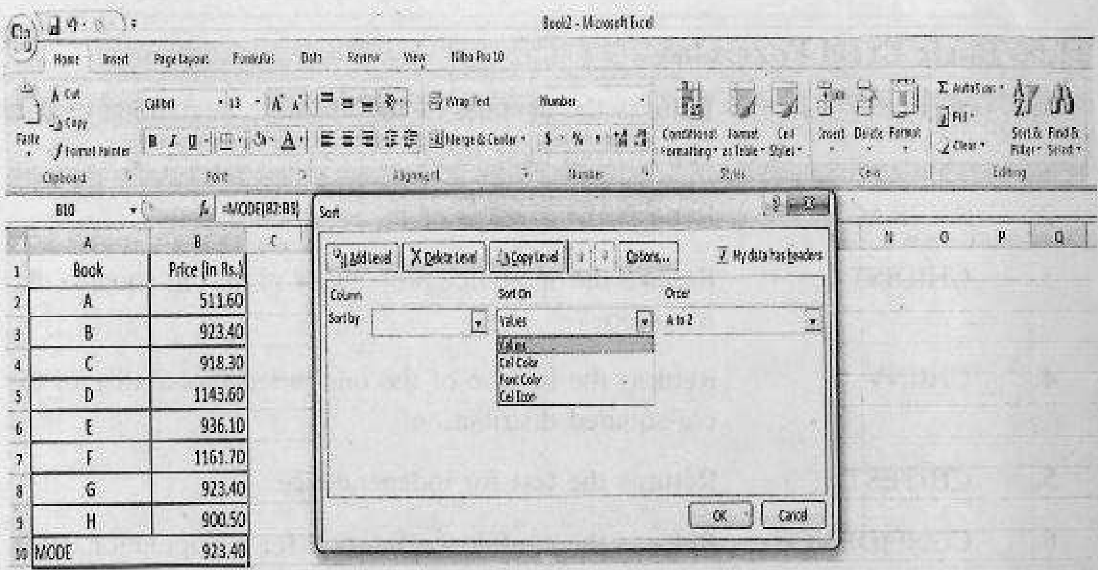
Sorting

The SORT function will arrange your data in increasing, decreasing, alphabetical, or reverse alphabetical order. Be careful when sorting. If you sort only one row or column, you will effectively "scramble" these data relative to the rest of the spreadsheet. If the relationship between data in different rows or columns must be preserved, always select the entire spreadsheet before sorting. Also, you can always undo a bad sort by typing "ctrl-Z" before you save.

The examples below are of GDP data for several countries in the western hemisphere. To sort this data, highlight the desired selection. With the "Home" tab selected on the top right, select the "Sort and Filter" menu from the top left.

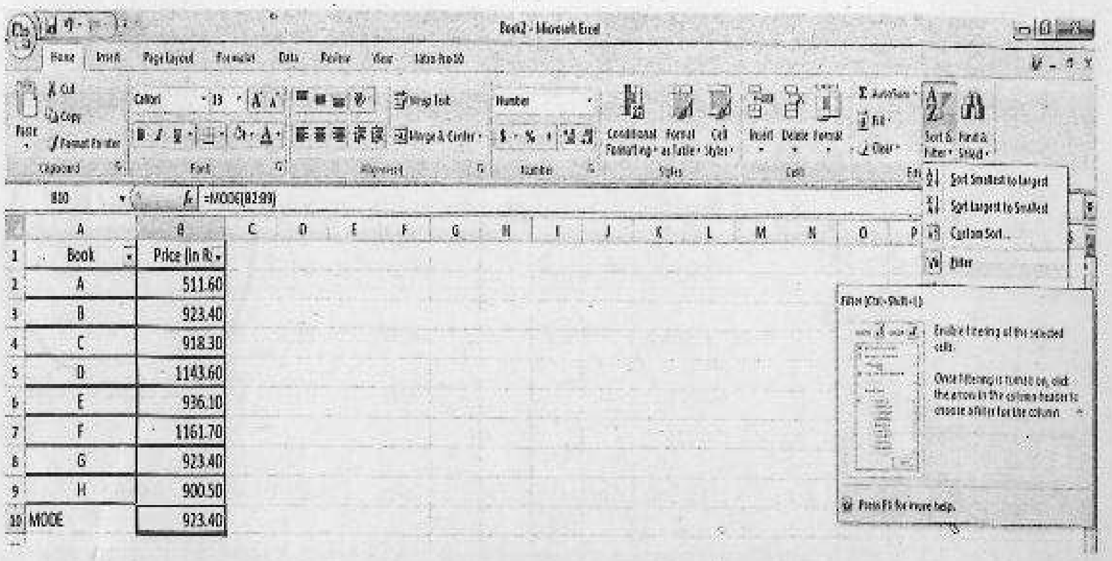
Notice that whichever cell you last clicked in is white (in the example below it is cell A1). If you select "Sort A to Z" or "Sort A to A" from this menu, Excel will sort your data in ascending or descending order, respectively, depending on the value in the column with the white cell.





Filtering

The FILTER function allows you to select a subset of your data to display. From the same "Sort and Filter" menu used above, choose "Filter". There will now be a small box with an arrow in the box on the first cell of each column:



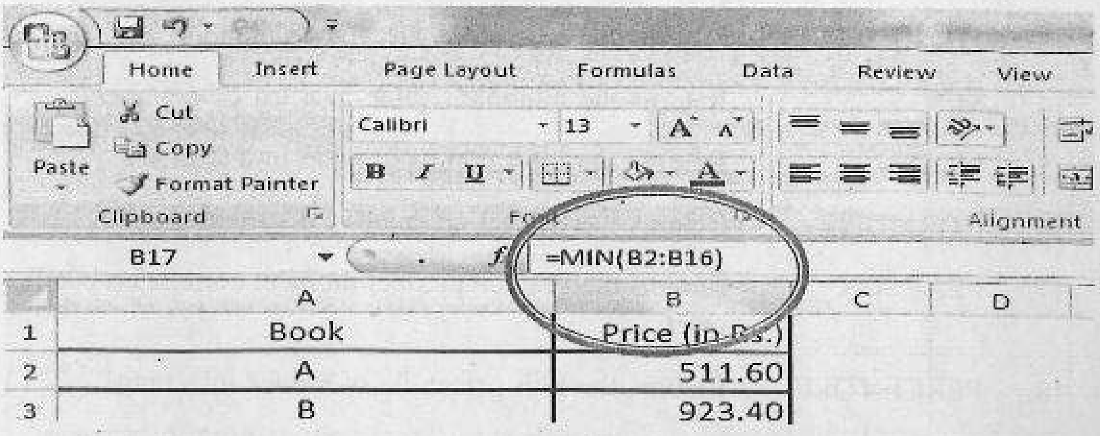
11.5. Basic Excel Formulas

1.	AVEDEV	Returns the average of the absolute deviations of data points, from their mean
2.	AVERAGE	Returns the average of its arguments
3.	CHIDIST	Returns the one-tailed probability of the chi-squared distribution
4.	CHIINV	Returns the inverse of the one-tailed probability of the chi-squared distribution
5.	CHITEST	Returns the test for independence
6.	CONFIDENCE	Returns the confidence interval for a population mean
7.	CORREL	Returns the correlation coefficient between two data sets
8.	COUNT	Counts how many numbers are in the list of arguments
9.	COVAR	Returns covariance, the average of the products of paired deviations
10.	FTEST	Returns the result of an F-test

11.	GAMMADIST	Returns the gamma distribution
12.	MAX	Returns the maximum value in a list of arguments
13.	MEDIAN	Returns the median of the given numbers
14.	MIN	Returns the minimum value in a list of arguments
15.	MODE	Returns the most common value in a data set
16.	NORMDIST	Returns the normal cumulative distribution
17.	PEARSON	Returns the Pearson product moment correlation coefficient
18.	PERCENTILE	Returns the k-th percentile of values in a range
19.	PERCENTRANK	Returns the percentage rank of a value in a data set
20.	POISSON	Returns the Poisson distribution
21.	PROB	Returns the probability that values in a range are between two limits
22.	RANK	Returns the rank of a number in a list of numbers
23.	SMALL	Returns the k-th smallest value in a data set
24.	STANDARDIZE	Returns a normalized value
25.	STDEV	Estimates standard deviation based on a sample
26.	TDIST	Returns the Student's t-distribution
27.	TREND	Returns values along a linear trend
28.	TTEST	Returns the probability associated with a Student's t-test
29.	VAR	Estimates variance based on a sample
30.	ZTEST	Returns the one-tailed probability-value of a z-test

11.6. Measures of Central Tendency

1. Arithmetic Mean



Book	Price (in Rs.)
A	511.6
B	923.4
C	918.3
D	1143.6
E	936.1
F	1161.7
G	848.0
H	900.5
I	1059.5
J	1053.2
K	1091.3
L	1314.7
M	1204.1
N	808.0
O	965.4
P	1068.3
Q	1293.1
R	880.9
S	1092.3
T	1001.5

Mean	=AVERAGE(B2:B16)	983.13
Minimum	=MIN(B2:B16)	511.60
Maximum	=MAX(B2:B16)	1314.70
Median	=MEDIAN(B2:B16)	936.10
Mode	=MODE(B2:B16)	923.40

11.6.2. Frequency Distribution

	A	B	C	D	E
1	Cost of Books				
2	511.6	977.7	600.2	1099.7	803.7
3	923.4	1108.3	906.7	759.6	1111.9
4	918.3	1051.1	992.5	817.2	665.3
5	1143.6	848.4	939.8	1163.0	715.2
6	936.1	750.5	991.2	1199.5	950.2
7	1161.7	1027.7	995.1	966.5	1146.5
8	848.0	956.8	1100.0	955.2	1023.0
9	900.5	982.3	699.2	1069.8	1245.3
10	1059.5	1091.0	850.7	1219.3	1012.6
11	1053.2	939.5	777.8	749.6	980.8
12	1091.3	1016.3	930.4	1242.2	1131.4
13	1314.7	1137.2	763.1	1294.4	917.3
14	1204.1	980.1	922.3	1057.7	907.2
15	808.0	857.7	1127.1	934.3	1262.3

16	965.4	873.4	955.1	806.5	1033.0
17	1068.3	950.3	930.6	1000.1	898.5
18	1293.1	940.9	1293.8	1035.2	706.0
19	880.9	912.2	803.5	922.6	846.1
20	1092.3	1182.0	985.2	945.3	835.0
21	1001.5	1048.8	895.1	1067.2	1062.8

CALCULATIONS USING FORMULAS

Calculation of	Result	Formula
MEAN	978.26	= AVERAGE(A2:E21)
SD	157.63	= STDEV(A2:E21)
MEDIAN	972.10	= MEDIAN(A2:E21)

11.6.3. Correlation

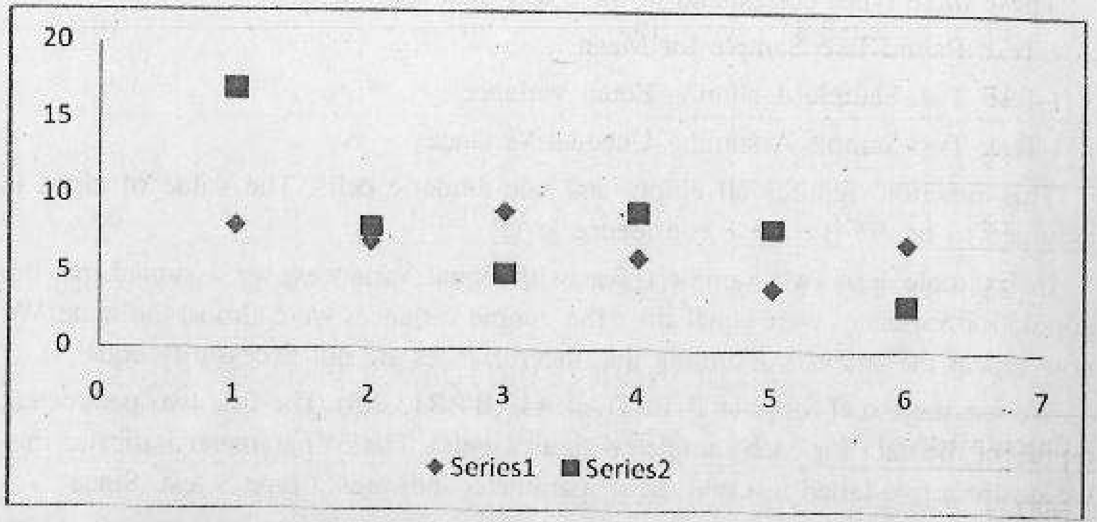
11.6.3.1. Pearson's Product Moment Correlation

Calculation of correlation between issue data of books in different years in a particular library

G6 fx

	A	B	C	D	E
1	Book Issued no of time	2015	2016	Correlation	Formula
2	A	8	17	0.056413103	CORREL(B2:B7,C2:C7)
3	B	7	8		
4	C	9	5		
5	D	6	9		
6	E	4	8		
7	F	7	3		
8	Total	41	50		
9					

11.6.3.2. Scatter Diagram



11.6.3.3. Spearman's Rank Correlation

	A	B	C	D	E	F	G
1	Book Issued no of time	2015	Rank 1	2016	Rank 2	Rank Correlation	
2	A	8	2	17	1	-0.269230769	CORREL(C2:C7,E2:E7)
3	B	7	3	8	3		
4	C	9	1	5	5		
5	D	6	5	9	2		
6	E	4	6	8	3		
7	F	7	3	3	6		
8							

11.6.4. Testing of Hypothesis

11.6.4.1. T Test

Excel provides the function TTEST to handle the various two sample t-tests.

TTEST(R1, R2, *tails, type*) = p-value of the t-test for the difference between the means of two samples R1 and R2, where *tails* = 1 (one-tailed) or 2 (two-tailed) and *type* takes the values:

1. the samples have paired values from the same population
2. the samples are from populations with the same variance

3. the samples are from populations with different variances

These three types correspond to the Excel data analysis tools

t-Test: Paired Two Sample for Mean

t-Test: Two-Sample Assuming Equal Variance

t-Test: Two-Sample Assuming Unequal Variance

This function ignores all empty and non-numeric cells. The value of alpha is assumed to be .05 (i.e. 95% confidence level).

In Example 1 of Two Sample t Test with Equal Variances, we assumed that the population variances were equal since the sample variances were almost the same. We now repeat the analysis assuming that the variances are not necessarily equal.

We use the Excel formula TTEST(A4:A13,B4:B13,2,3). The first two parameters represent the data for each sample (without labels). The 3rd parameter indicates that we desire a two-tailed test and the 4th parameter indicates a type 3 test. Since

TTEST(A4:A13,B4:B13,2,3) = 0.043456 < .05 = á

we reject the null hypothesis. Note that if we use the type 2 test, TTEST(R1, R2, 2, 2) = 0.043053, the result won't be very different, thus confirming our assumption that the population variances are almost equal.

11.6.5. Graphical Representations

11.6.5.1. Creating a Chart

Open a spreadsheet and add data as follows.

Select the cells A1 to G5 as shown below

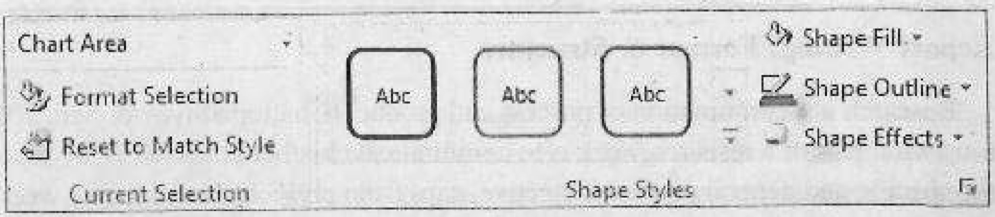
	A	B	C	D	E	F	G
1	Expenditure						
2	Month	Jan	Feb	Mar	Apr	May	Jun
3	Rent	200	200	200	250	300	250
4	Electricity	20	22	18	25	30	28
6	Household	150	145	150	130	150	140

Steps To create a chart:

- Select the data in the worksheet.
- Click on the **Insert** tab, in the **Charts** group, choose the **chart type** you want.
- By default the chart is embedded into the worksheet but you can change the location *(see chart location below)*. When you click on the embedded chart in

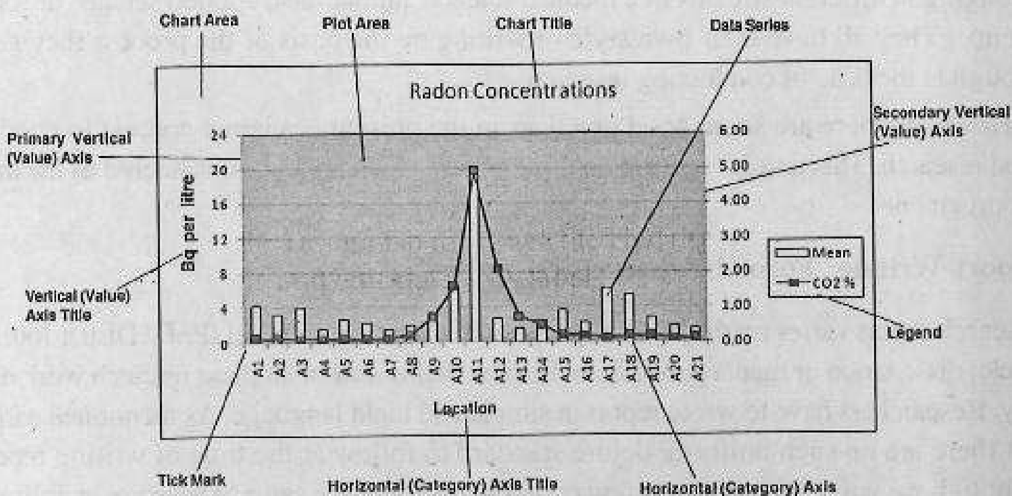
the worksheet the **Chart Tools** is displayed, at the top right of the ribbon, together with three additional tabs: **Design, Layout**, and **Format**.

- You can format different elements of the chart by clicking on the **Format** tab and select the required option from **Current Selection** and **Shape Styles** group.



- You can use the **Design** tab to change the **chart type, chart styles, Create a chart template and edit the x axis labels**

A chart will consist of various elements and it can be useful to know what Excel calls them, for example, the help system might tell us to select a certain part of a chart.



Unit 12 □ Report writing : Format and structure ; tudy of Style Manuals; Citation Standards; Plagiarism Detection

Report Writing: Format & Structure

Research is a combination of process and product (Chattopadhyay & Sen, 2013). The most vital part of a research work is to communicate his/her entire work by focusing on the specific and general purpose/objective, gaps from previous researchable work, analysis, conclusion, recommendations to make it easily understandable and further citable (Adams et al., 2014). The term "thesis" comes from the Greek 0eoic,, meaning "something put forth", and refers to an intellectual proposition. "Dissertation" comes from the Latin origin dissertatio, meaning "path". It is very much expected that researcher should communicate their research work in an appropriate manner so that right interpretation to be made to the scholarly communication. Otherwise inspite of all good works research may not be well accepted. According to Newman, *"A research rep'ort is a written document that communicates the methods and findings of a research project to others."* In nutshell, Research report is a detailed documentation of one's individual researcher observations, descriptions and recommendations after collecting, analysing and interpreting facts and findings of related research area.

Perhaps it is not feasible to write research report in a same flow for researchers belonging in different streams like medical science, humanities, applied science or social science. They all have their own style of writing on the basis of the process they gone through at the time of conducting research.

However, there are some good practices in the present academic context to conduct good research. This unit is focusing on those issues which are to be considered as the most important one.

Report Writing: Format & Structure

Research report varies on the basis of purpose like academic award (PhD/ DLit); Journal article; dissertation or media publication. In view of this user of all these research work may vary. Researchers have to wrote report in simple and lucid language. As mentioned earlier that there are no such uniform/ dejure standard to follow at the time of writing report. Although we will concentrate on few components to do the same. These are as follows:

Introduction:

Introduction of the research report should briefly state all about the research in nutshell. The importance or significance of the research or outline, scope of the respective research. This is the place from where users/readers can understand the outline of the entire work is trying to achieve. This section includes research problems, questions, design and findings.

Objectives of the study:

In this section one have to state in simply what researcher is trying to reveal/ achieve by conducting this research work. Researcher should state main objectives of the study pin pointedly in bullet form.

Literature Review:

Before starting their research work researchers are intended to review previous works on the concern topic in relation to existing literature. Researcher can take insight from published materials like reputed journals, articles, text books, non-governmental data sets etc. On the basis of previous literature, they are trying to find the actual research gap regarding his/her topic after going through related present literature. At the time of stating the process report consist Summary of major themes, concepts and/or trends as discussed by the previous researchers. Writing pattern of the literature review should critically analysed and focused on the specific problems. This process may effect whole research work by guiding and it simply pointed out the drawbacks of the previous work.

Statement of the research problems

Statement of research problem is focusing on,

To state/list actual and most important research problem(s) to be shorted out throughout the whole process. There may be one problem exists or may be more than one topic exists. What a researcher reveals after gone through the previous literature it should be clearly stated in this part of the report including how it will add to current knowledge, or address existing gap. A proper research work should be calculated on the basis of the problems and its efficient/potential results. These problems will help researchers to concentrates on the stated areas with the help of available resources.

Research Design:

In this section researchers are directed to write words regarding the whole process of the research which is to be followed by them at the time of execution of their work. This section actually means to states a frame work of the whole work. It clearly indicates the way to find out the problems relating to research work. It includes study types vis. (Descriptive, experimental, survey), research questions, hypothesis, variables, methods of data collection, design, plan for statistical interpretations.

Data Analysis & Interpretation

Facts and findings revealed from the collected data should be conveyed in this section so that research report carry a clear picture of the collected data and its interpretation for better execution of the research result. In this important section researchers have to state about the collected data and on the basis of those data he/she has to analyse the facts and should elaborately state the findings.

Summary & Conclusion:

Add your own conclusions and thoughts that will help the reader understand what you are trying to say at the end of the research report. This part will have conveyed by a researcher to express his/her own findings in his own words which he/she faced at the time of pursuing his/her research work. Even they also recommend few point for further study here.

APA Reference Style

Reference of Book and book chapterinclude four elements: (1) Author/Editor/ Producer (2) Date (3) Title of the work and (4) Publication Information

Author, A., & Author, B. (Year). *Title of the work.* Place name: Publisher.

Electronic books are available online, a retrieval statement or DOI is required after (3) Title. Exclude (4) Publication Information.

Author, A., & Author, B. (Year). *Title of the work.* Retrieved from http://... Author, A., & Author, B. (Year). *Title of the work.* http://dx.doi.org/xx-xxxxxxx
Book: Author, A., & Author, B. (Year). Chapter title. In A. Editor, B. Editor, & C. Editor (Eds.), *Title of the book* (pp.xx-xx). Place name: Publisher.
E Book: Author, A., & Author, B. (Year). Chapter title. In A. Editor, B. Editor, & C. Editor (Eds.), *Title of the book* (pp.xx-xx). Retrieved from http://...
E Book with DOI: Author, A., & Author, B. (Year). Chapter title. In A. Editor, B. Editor, & C. Editor (Eds.), *Title of the book* (pp.xx-xx). http://dx.doi.org/xx-xxxxxxx

Chapter in a book/e book:Several times a book consists of more than one chapters written by different author(s), You have to reference each chapter used.

Author, A., & Author, B. (Year). Article title. *Title of Periodical*, *x*(x), pp-pp

Reference of Periodicalsshould include the following elements: (1) Author (2) Date (3) Title of article (4) Title of Periodical (5) Volume, Issue and Page numbers

Journal article available online:

Author, A., & Author, B. (Year). Article title. *Title of Periodical*, *x*(x), pp-pp. http://dx.doi.org/xxx-xxxxx

Author, A., & Author, B. (Year). Article title. *Title of Periodical*, *x*(x), pp-pp. Retrieved from http://...

Online documents / Webpages

Author, A., & Author, B. (Year). *Title of the webpage*. Retrieved from http://...

Author, A., & Author, B. (Year). *Title of the webpage*. Retrieved from ...website: http://...

Author, A., & Author, B. (Year, Month Day). Title of the webpage [Description of form]. Retrieved Year, Month Day from http://...

Social Media chapter include four elements: (1) Author (2) Date (3) Title (Provide the name of the page or the content or caption of the post (up to the first 40 words) as the title) and (4) Source

Facebook : Gaiman, N. [Neil]. (2012, February 29). Please celebrate Leap Year Day in the traditional manner by taking a writer out for dinner. It's been four years since many authors had a good dinner. We are waiting. Many of us have our forks or chopsticks at the [Facebook status update]. Retrieved from https://www.facebook.com/neilgaiman/posts/10150574185041016

Twitter : Gates, B. [BillGates]. (2013, February 26). #Polio is 99% eradicated. Join me & @FCBarcelona as we work to finish the job and #EndPolio. VIDEO: http://b-gat.es/X75Lvy [Tweet]. Retrieved from https://twitter.com/BillGates/status/306195345845665792

Google+: Cornell University. (2012, October 11). Having a cup of coffee before closing your eyes is the most effective way to combat daytime drowsiness, according to research. Sounds counterintuitive, but it takes 20 minutes for the caffeine to get into your bloodstream. So if you take [Google+ post]. Retrieved from https://plus.google.com/116871314286286422580/posts/NqCFGr4eveT

Structure:

Parti

Title Page
Preface
Acknowledgment
List of Tables
List of Abbreviations
Content page

Part II

Introduction
Objectives of the study
Literature Review
Statement of Research Problem
Research Design
Data Analysis & Interpretation
Findings
Conclusion
Reco mmendatio ns

Part III

References (by following referencing style APA 6th edition)
Index
Glossary

Formatting Research Paper:

Margins: Every page of dissertation must have one-inch margin on all sides, top, left and right.

Fonts: Use Times New Roman font in 12 size. Bold face may be used within the text for reading/ sub headings/ highlighted terms.

Spacing: 1.5 spacing should be used except those materials having predefined spacing such as footnotes, quotations, tables, name of the figures etc.

Number(s) scheme

Chapters may be identified with Arabic numbers followed by preliminary pages in uppercase Roman numerals. Tables and figures should be numbered consecutively throughout the manuscript in Arabic numerals. In case of appendices use uppercase Roman letters. Use Arabic numerals to describe Headings and sub headings under principle chapter number. Ex: - Chapter 3: Methodology

3i: Data collection

311: Data collection: Tools

312: Data collection: Techniques

32: Data Analysis

321: Data Analysis: Statistical Interpretation So on as so forth

Pagination: All pages must be numbered except title page. Place page number at the bottom of the page in right side consecutively throughout the manuscript. Preliminary pages must be numbered sequentially in upper case Roman letter. Portrait orientation and all pages in A4 size is required. In case of wide chart, table & graphs, rotate the image 90 degree clockwise. Text should be oriented in the same way as the chart/ table/ graph/ graphics. In that case place page number as directed earlier.

Table of contents: Table of contents should include, the respective title of the chapter along with page number.

Abstract: Abstract of the educational research work should be within 2OO-25Owords which may include a brief statement of the problem, methods, procedure(s) used to study and a condensed summary of the finding(s).

Reference: All reference should be prepared according to APA 6TH ed standard style manual.

Physical Format: The length of the report maybe of approximately 25,000 words (excluding appendices and exhibits) with 1.5 space. However, 10% variation is permissible. Dissertation should be printed on A4 size papers and submitted in bound form.

Language: The language of the report of the dissertation should be in English. However, permission may be given by the faculty concerned to write the report in Bengali for a specific topic.

Citation Standards

At the time of writing research report it is mandatory to acknowledge authors from whom researchers got the relative information on the selective topic. As mentioned in earlier section that research is a "result of a certain accumulation" (Riders,2013). So they are compelling to cite those actual writers. Any university assignments/dissertation/project work that portrayed on the ideas, words or research of other previous writers must contain citations. In the process of writing a report it is most desirable that researchers follow one particular style of citation to maintain unity. In another terms it would be better to say that by citing the work of a particular scholar we are acknowledging and submitting our respect to the intellectual property rights of that researcher. Apart from this, citation is a document to provide evidence to strongly support the assertions of researchers own assignments. It is also a evidence that you are enough aware of this particular field and after going through thoroughly you are able to found a gap between the process and you are trying to filling the gap of the existing process/domain. Accurate references allowing future researchers to trace the actual and relevant sources of information which have been used throughout the research work.

In this section we are going to discuss about some most important citation style so that after going through this unit, researchers might be able to choose appropriate citation style. Citation style guides are generally frame the standards for writing with specific fields and creating consistence unity to writing. Determination of style guide depends on the stream of domain along with documents and sometimes on the organizations/publication house who create their own style/guidelines for submission in a specific/directed referencing style. Some most important style guides are discussed here in table no for having a stipulated idea of different available citation style of scholarly communication.

List of General reference Style Guides and Manuals

American Psychological Association 6th ed	APA is generally used by the scholars of social and behavioral sciences. It facilitates in text citation & Reference.
The Chicago manual of style	*The Chicago manual of style*. 16th ed. is a style guide for American English published since 1906 by the University of Chicago Press. It is also available

	online. Author-date style and Notes and bibliography style citations are there in CMS. Primarily used by writers and students in humanities particularly in history. Footnote /Endnote system for notation and use of "Ibid" is basically use for referencing of sources.
Modern Language Association	MLA Style Manual and Guide to Scholarly Publishing (3rd edition-2008) largely used for documentation in the humanities, specifically, languages and literature, including English, modern languages, and comparative literature. In text citation (Author, #) & "Works cited" are used as reference style.
American Medical Association	According to the MLA, the ***MLA style*** "has been widely adopted for classroom instruction and used worldwide by scholars, journal publishers, and academic and commercial presses". widely used in the United States, Canada, and other countries, directing guidelines for writing of research in the humanities, such as English studies, comparative literature; media studies; cultural studies and related domain.
American Sociological Association	**ASA style** is a widely accepted format for writing university research papers in the field of sociology. It specifies the arrangement and punctuation of footnotes and bibliographies.

Harvard Reference Format	**Harvard referencing, (Parenthetical referencing)** is a citation style in which partial citations are enclosed within parentheses and embedded in the text, either within or after a sentence, accompanied by a full, alphabetized list of citations in an end section, usually titled "references", "reference list", "works cited", or "end-text citations". Harvard referencing can be used in footnote citations.
IEEE	The Institute of Electrical and Electronics Engineers (IEEE) style which is based on Chicago Style, is generally used in technical fields, particularly in computer science. In IEEE style, citations are numbered, but citation numbers are included in the text in square brackets rather than as superscripts. All bibliographical information is exclusively included in the list of references at the end of the document, next to the respective citation number.

Sources are to be refer

Acknowledge every sources from which any words, ideas or information taken to frame/ develop new research work acknowledge every sources from which any words, ideas or information taken to frame/ develop new research work requires a reference. But generally accepted facts or information may not require any kind of references. Different forms of sources may have considered to references as follows:

- Books
- Journal articles;
- Newspapers and magazines;

- Pamphlets;
- Films, documentaries, television programs or advertisements;
- Websites or electronic resources;
- Letters, emails, online discussion forums;
- Personal interviews;
- Lecturers note

Reference when you reprint any diagrams, illustrations, charts or pictures.

Reference Management Software:

Objectives:

— To organize/store the metadata of collected articles/journals;

- To create a bibliographic database in a local system with global accessibility (Centralized reference collection);

- To cite maximum number of data at the time of writing and to make a proper bibliography with its appropriate style of submission;

- Import and export references/data; and
- Sharing and collaboration of bibliographic data
- To simplify writing process.

Introduction:

Managing references for scholarly communication is lengthy, laborious and hard task. Due to the vibrant and flexible electronic environment, the span of scholarly resources are increasing day by day. It is very much important to manage resources in a proper and systematic way at the time of creating a new scholarly communication. It is mandatory to acknowledge actual responsible author for using source(s) of each and every piece of information for a research work. In that case reference management tool/software help us to store, organise and cite the references at the time of writing. It also allows to create a list of bibliography automatically with its appropriate style of submission. According to Martin Fenner reference managers helps scholarly communication by 1) Searching relevant literature 2) Storing the search result in a personal database and 3)insert references at the time of writing. ("Martin_Fenner.pdf," 2010) Reference management software allows to access any type of documents regarding any sorts of article/journal, conference proceedings from web pages. Each record may have its additional information consisting of title, keywords, URLs, abstract etc. Scholarly communication can enrich their database by adding a huge number of records.

References can be imported from online databases such as PubMed, SciFinder etc. with an abstract and with other bibliographic information. On the same way reference management software can export data from various online databases in a localized systems. Every reference management software can export and import datasets on the basis of reference trancation standards like BibTex, Endnote, RIS (Research Information Systems). As a result, standards-enabled reference management softwares can snatch bibliographic datasets from compatible databases such as SciFinder, ISI (Information Sciences Institute) database, PubMed, Medline. Similarly, library management softwares that support above standards, can export/import/transfer datasets from catalogue database to reference management software databases. After creating database/library of references creator can cite these references relavant to his/her document. In such a way, a complex and time consuming task have done by author easily.

Methods of Selections and integration of Reference management software:

There are many reference management software like Aigaion, Bibus, Connotea, CiteULike, Endnote, Jabref, Refbase, Wikindx, Zotero are ready to facilitate effective and efficient management of references. We choose four open source software as competitors to focus on which are the most comprehensive reference management software in view of operating among these. A comparative discussion is given below to show which one is the best reference management software.

Sl. No	Parameters for comparison	Mandley	Connotea	JabRef	Zotero
1	Developer	Elsevier	Nature Publishing Group	JabRef developers	Zotero Center for History and New Media at GMU
2	Open Source	Yes	Yes	Yes	Yes
3	Operating system Support	Windows, Mac OS X, Linux, Unix, BSD	Windows; Mac OS X, Linux, Unix, BSD	Windows, Mac OS X, Linux, Unix, BSD	Windows, Mac OS X, Linux, Unix, BSD
4	Export file formats	BibTeX, Endnote, BibIX, Refer, Medline, RIS, SQ Lite	BibTeX, Endnote, BibIX, Refer, RIS, RDF	BibTeX, Endnote, BibIX, Refer, MODXML, RIS(depends), BibTeXML, DocBook, OpenDocument for OO.o, user-customizable	BibTeX, Endnote, BibIX, Refer, MODXML, RIS, RDF, Wikipedia citation templates
5	Import file formats	BibTeX, Endnote, BibIX, Refer, ISI Medline, RIS,	BibTeX, Endnote, BibIX, Refer, ISI, RIS, Medline,	BibTeX, CSA, Endnote, BibIX, Refer, ISI, Medline, Ovid, PubMed, RIS, SciFinder	BibTeX, Endnote, BibIX, Refer, RIS, COinS, TEI, RefWork, CSV, CSL JSON,
6	Citation Style	APA APA, MLA, Chicago/Turabian, Harvard, any BibTex style		APA, MLA, Chicago/Turabian, Harvard, any BibTex style	APA, MLA, Chicago/Turabian, Harvard, IEEE, MHRA, Natur, Vancouver
7	Word processor integration	MicroSoft word, Open Office	No	MicroSoft word(depends), Open Office	MicroSoft word, Open Office
8	Database connectivity	Web of Science, PubMed, Scopus, CAB Abstracts	ArXiv, CiteSeer, Pubmed	ArXiv, CiteSeer, IEEE Explore, Pubmed	ArXiv, CiteSeer, IEEE Explore, Pubmed

Table 1: Comparison of four reference management software

From the above table we see that Zotero and Mendeley are the most comprehensive reference management software in view of its functionality, usability and accessibility.

Zotero:

Zotero is free/ open source reference management software product which helps scholarly communication to collect, organise and manage references from primary, secondary and tertiary document through the webpages. Zotero is an extended Firefox Plug-in which integrates with online resources at the time of browsing. Zotero extracts the metadata of documents from sources which are to be included and to be cited. By adding each and every references, creator can enrich their database which is his/her personal library. ("Bibus - Wikipedia, the free encyclopedia").

Installation:

At first we have to install Zotero as Firefox Plug-in from www.zotero.org. Then Zotero will come on the left side of the task panel. Download Openoffice.org or word processor plug-in for insert intext citation and bibliography in manuscript. Additional plug-in required extracting metadata from pdf files. Zotero can't directly/ automatically extract the metadata from pdf files.

Figure 1: Zotero Installation

Configuration:

User have to register themselves on Zotero website for creating an account on Zotero platform and save their required data on the web so that it could be accessible from anywhere, at any time by researcher.

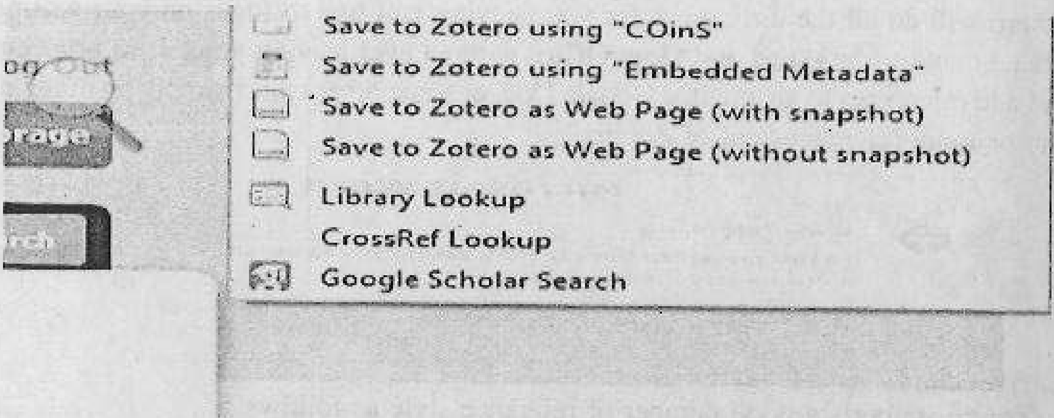
Zotero allow researchers to:

1.Collect:

"Zotero is the only research tool that automatically senses content in your web browser, allowing you to add it to your personal library with a single click. Whether you're searching for a preprint on arXiv.org, a journal article from JSTOR, a news story from the *New York Times*, or a book from your university library catalog, Zotero has you covered with support for thousands of sites."

Figure 2: Creating new collection

Documents could be saved by clicking on any of these options as per requirement.



2.Organise:

"Zotero organizes your research into collections that act like iTunes playlists. Research items can be added to any number of named collections and subcollections, which in turn can be organized however you like. With saved searches, you can create smart collections that automatically fill with relevant materials as you add them to your library."

This is an example of **Zotero My Library** which includes different kinds of resources in different collection.

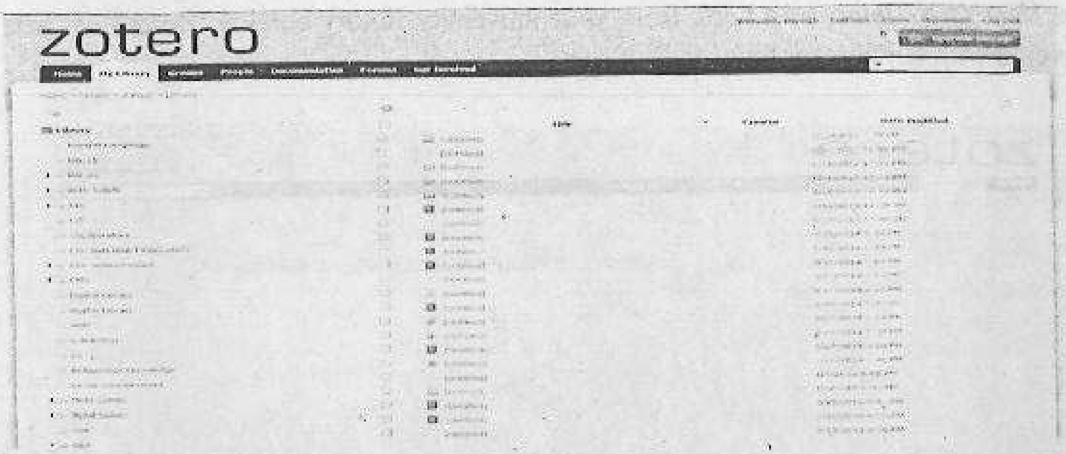


Figure 3 : Zotero My Library

3.Cite:

"Whether you need to create footnotes, endnotes, in-text citations, or bibliographies, Zotero will do all the dirty work for you, leaving you free to focus on your writing. Create citations in **Word** and **OpenOffice** without ever leaving your word processor and add references to an email, a Google Doc, or some other editor simply by dragging one or more references out of Zotero."

Zotero supports a good number of reference style as follows:

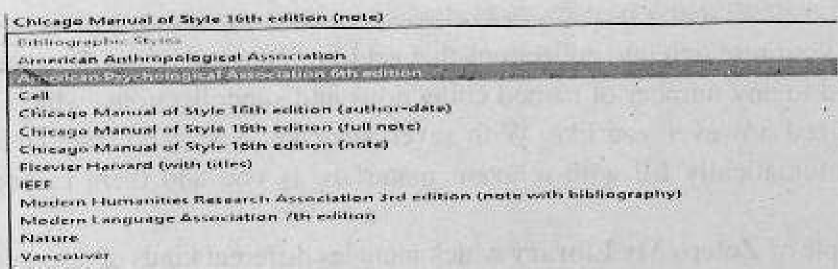


Figure 5 supported reference styles

Export Formats

References might be exported in following formats:

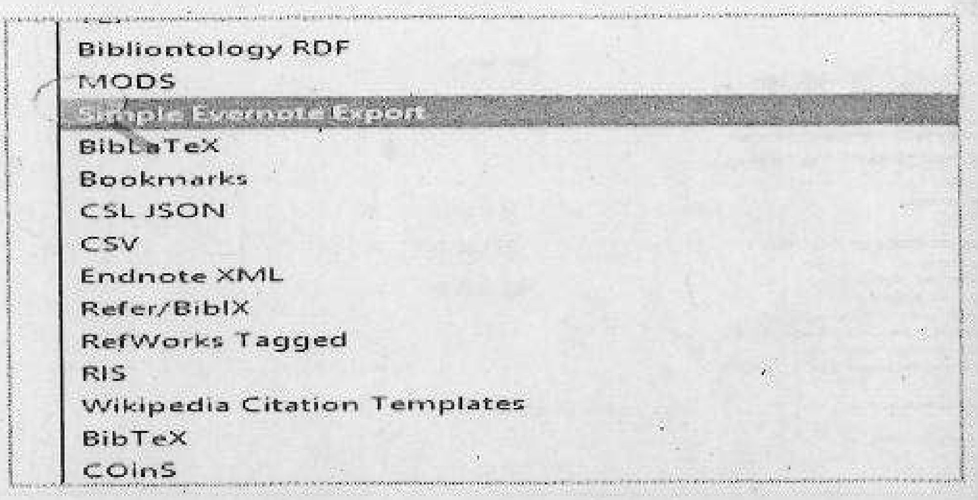
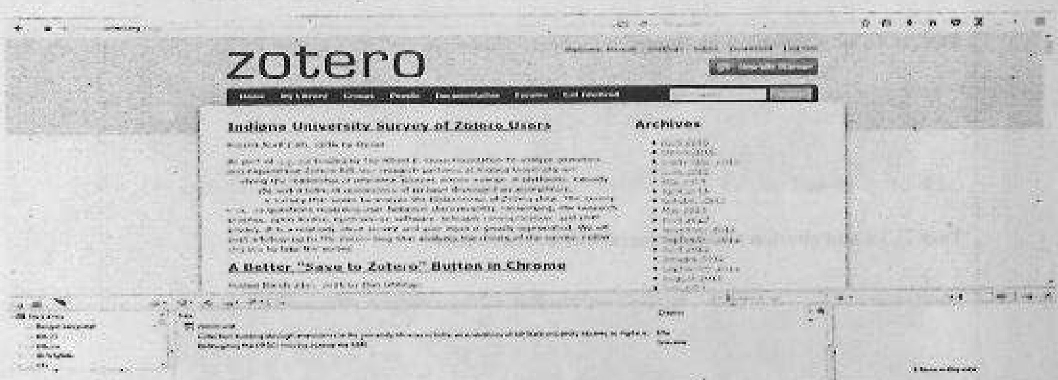


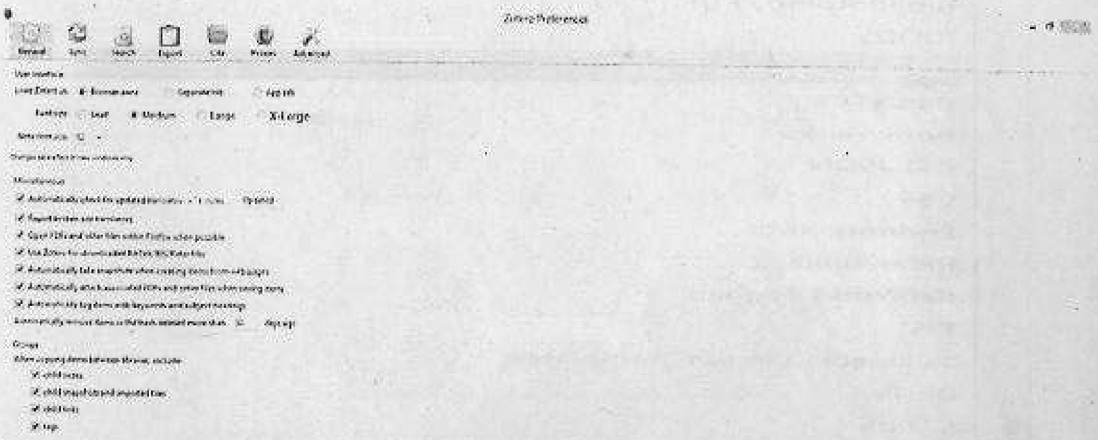
Figure 6: Supported Export Formats

4.Sync:

"Zotero automatically synchronizes your data across as many devices as you choose. Add to your research library on your work PC, and organize your collections on your home laptop. All of your notes, files, and bibliographic data remain seamlessly and silently up to date. Returning from field work? Your data will be waiting for you when you get home."



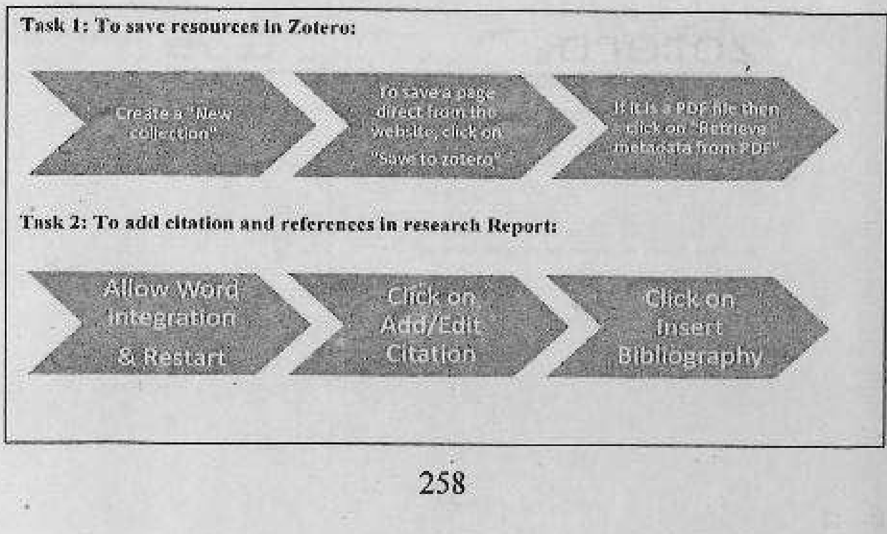
Upgrading is also possible while you are using even on a new computer. Zotero will automatically fetch a complete copy of research's saved library from Zotero server network. Researcher might be able to use this service from any web browser Even without installing Zotero.



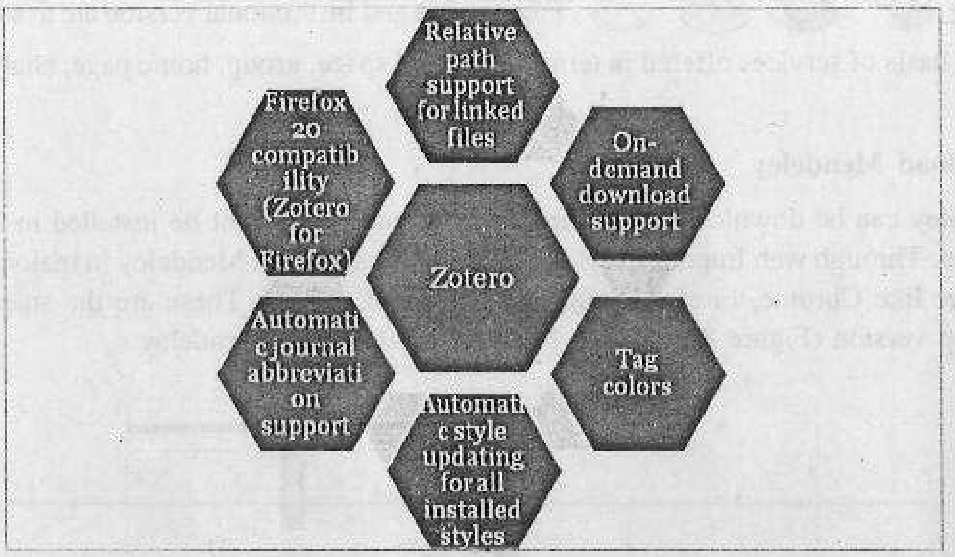
5. Collaborate

"Create and join research groups to focus on any topic you choose. Each group can share its own research library, complete with files, bibliographic data, notes, and discussion threads. Tag and analyze your research together with others. Work with a single colleague or an entire class: Zotero groups can include as many members as you please."

Work of Action in nutshell to work with Zotero



New version of Zotero i.e. Zotero 4.0 has been announced with mentioned updated features as follows:



Mendeley

Mendeley is a crossed-platform based free reference management software as well as an academic social network which facilitates the provision to organize works related to research, and discover the latest scholarly communication in an online and collaborative platform. It enhances the possibilities to store reference data centrally in cross platform in order to simply the writing procedure and make it globally accessible.

Figure 7: Mendeley Home

Mendeley exists both as a desktop application (slightly more advanced) and as a browser application. It is available on all platforms and in all popular browsers like . Free version and Institutional version are available on the basis of services offered in terms of allotted space, group, home page, analytics etc.

Download Mendeley

Mendeley can be downloaded by using browser and/or it might be installed in local desktop. Through web Importer version researchers can install Mendeley in major web browser like Chrome, Firefox, Safari and Internet Explorer. These are the snaps of desktop version (Figure 8) and web browser (Figure 9) of Mendeley.

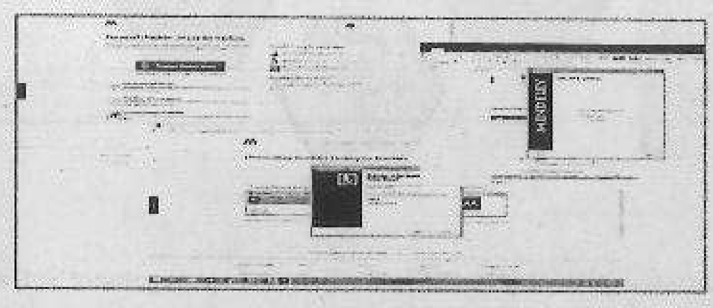


Figure 8: Desktop Version

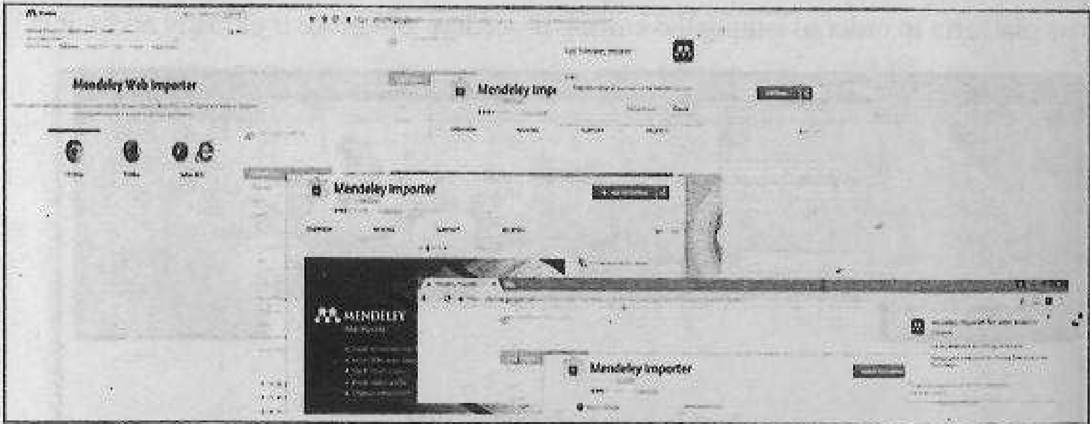


Figure 9: Web Importer

Getting started with Mendeley

Before beginning researcher have to have an mendeley account on the web. Then only researcher can use this reference management software.

Figure 10 Mendeley register

MENDELEY **Work of Action**

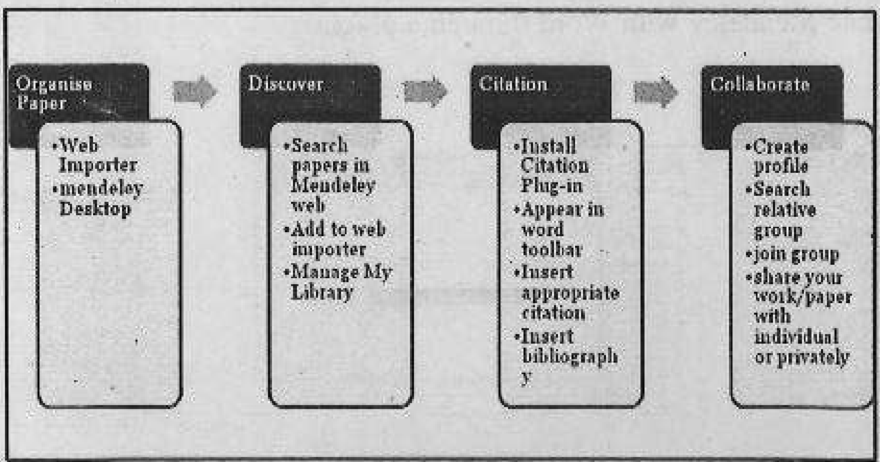


Figure 11 Mendeley Work of Action

Researcher could have specialized their search by subject, title or author and add documents to their own library. Even they can also filter their search. They are able to add Select a file or folder to add from computer, import from another reference manager, or BibTeX or simply Drag and Drop the pdf into Mendeley Desktop.

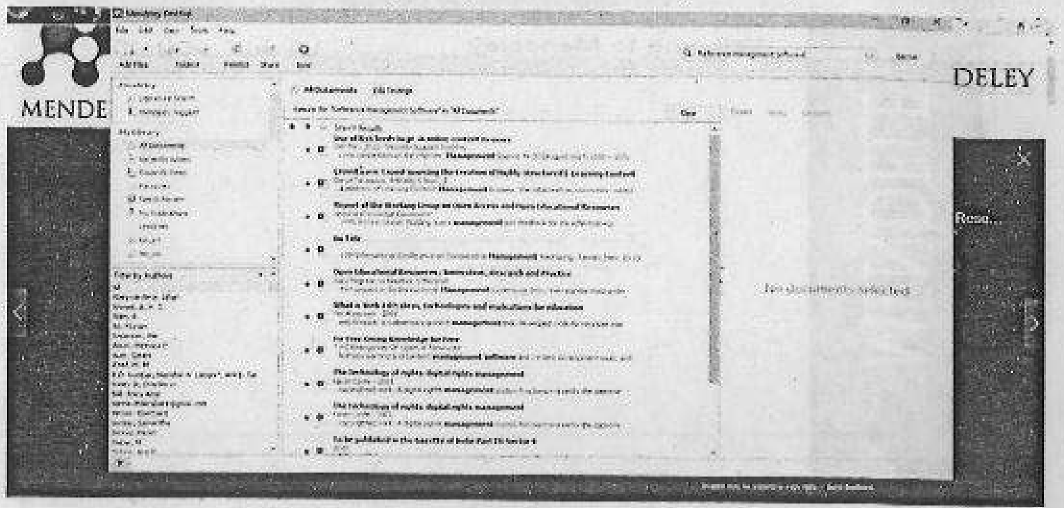


Figure 12 Mendeley Desktop

Integration of Mendeley in MS Word Plugin

By integration of Mendeley researchers will be able to add intext citation and relative bibliography at the time of writing research report easily. It's a matter of a single click to integrate Mendeley with Word through a plugin.

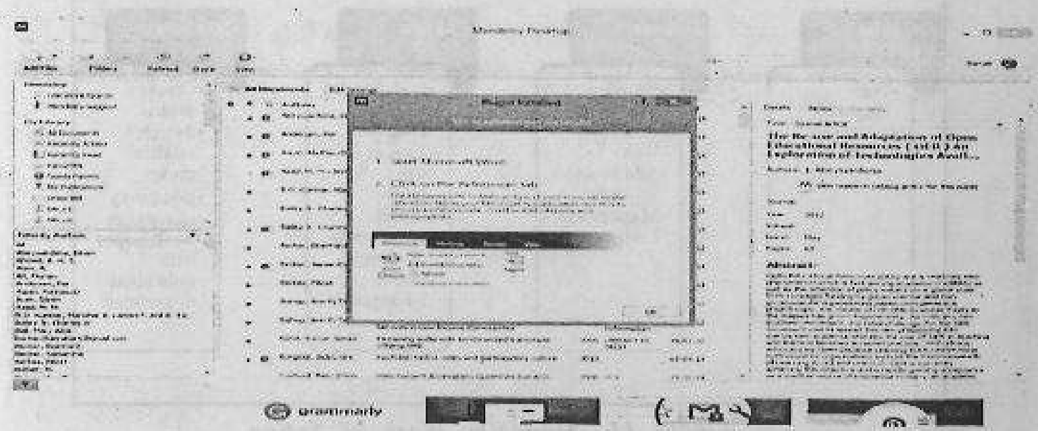


Figure 13 Word integration

My Library

Mendeley My library is a place where researcher can store their searched documents and add new references. Folder wise or group wise even author wise retrieval of references is possible here.

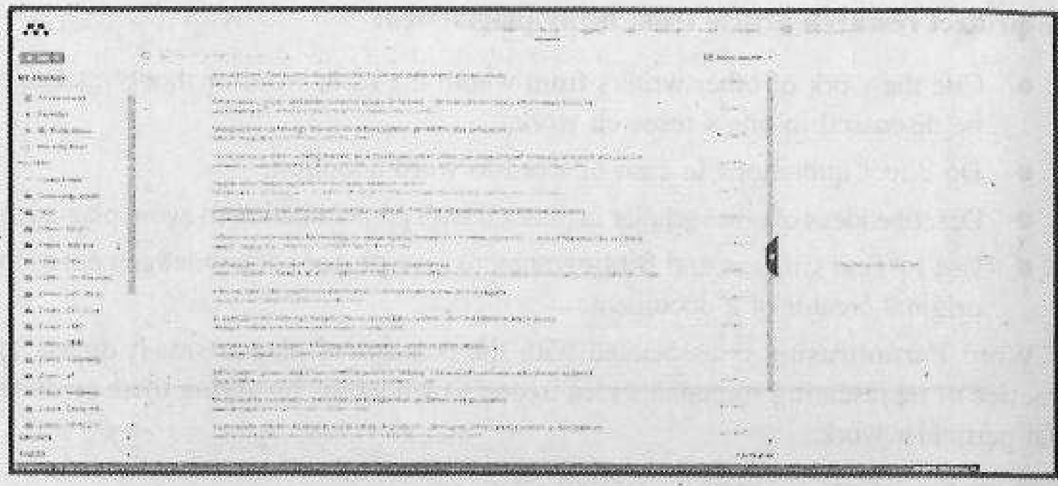


Figure 14 : Mendeley My Library

Plagiarism Detection :

Plagiarism is representing someone's work, word, idea as one's own work without giving proper acknowledgement. It is necessary to protect author intellectual property right for shake of academic integrity. Claes Sandgren, Professor of Civil Law in a report to the Swedish Association of University Teachers', SULF, congress in November 2002, presented various copyright issues within universities and colleges. In his deliberations he stated that, "Internet is both a legal problem and a control problem. Cheating and plagiarism are not more tolerable just because work is stolen from the Internet rather than from a book or magazine. However, the ability to screen has become more difficult, and it is has become easier for those wishing to cheat to steal from each other" (Universitetsläraren magazine, 17/2002)

Reasons behind Plagiarism:

- Some scholars don't have clear concept of plagiarize;
- They may have language problem
- They don't have in-depth knowledge in said topic

- Scholars have Don't care attitude.
- Action taken being caught as plagiarized document:
- Alleged person may have caught legally May loss degree, job, grunts; Loss of social dignity and unexpected situation.

To protect research article from being plagiarized:

- Cite the work of other writers from whom the idea, word or thought taken to be discussed in one's research work;
- Do direct quotations in case of word to word adoption;
- Describe ideas of other scholar in nutshell with proper citation to avoid plagiarism.
- Use In-Text Citation and Bibliography to give proper acknowledgement to the original creator of a document.

Word **Paraphrasing** is associated with the concept of plagiarism. It directs the practice of representing someone's idea in one's own words by giving them credit for that particular work.

Free and Open Source Plagiarism Checker:

1. Open Software Plagiarism Checker: A open source software plagiarism checker. It checks the similarity of individual files and generates various types of reports.

2. Anti-Plagiarism (Check on plagiarism): Anti-Plagiarism Software is free Anti-Plagiarism - software designed to detect and prevent plagiarism to support academic integrity. This software checking documents in *.rtf, *.doc, *.docx, *.pdf format.

To get rid of from this copyright issues researcher can use CC licensed documents with proper license. Creative Commons' public copyright licenses incorporate a unique and innovative "three-layer" design which includes cc to legal code, human readable contents and cc to machine readable contents. The main conditions of creative commons are stated below:

Attribution:the rights to use a content of an original author by giving him/her credit the way author requested for by given kind of license. If people want to use it without giving author's credit or for endorsement purposes, they must get author's permission first.

Non-Commercial: With this licensing tool people can copy, distribute, display, perform, and (unless author declared it as No Derivatives) modify and use a original work for any purpose other than commercially unless they get author's permission first.

Share Alike: Through this CC tool people can copy, distribute, display, perform, and modify a work, and may distribute any modified work on the same terms. Permission is needed, if they want to distribute modified works under other terms.

No Derivatives: People can copy, distribute, display and perform only original copies of a work in this case. If they want to modify any work having this license, people must get author's permission first. A symbolic representation is given in figure 2.

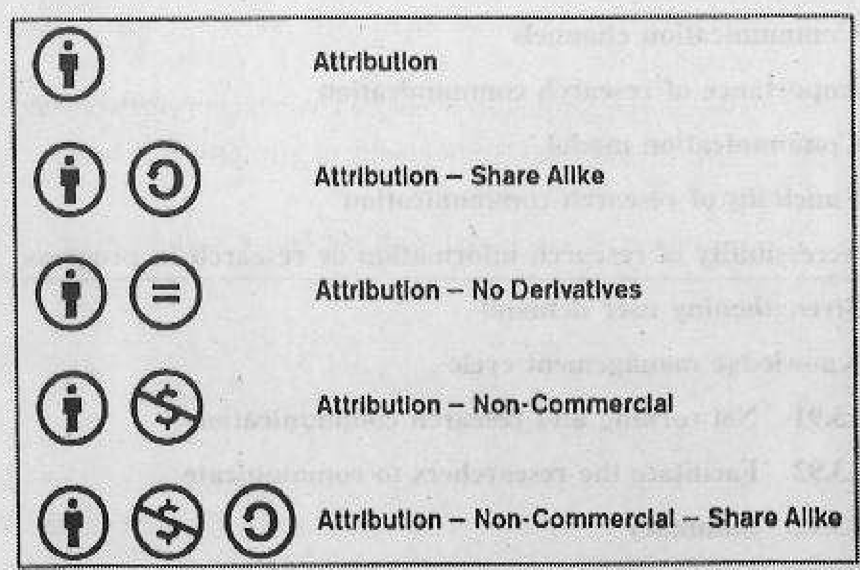


Fig:2 Creative commons attributes

Source: creative commons.org

Use of Creative Commons licensed references in research may reduce the risk of plagiarism by catering open license documents in scholarly communication. But it is very much expected that in time of report writing researcher should acknowledge original author(s).

Unit 13 ❑ Research Communication: Process and Channels

Structure

13.1 Introduction

The word communication is derived from the Latin word, *communis,* which means *common.*

It is the process of transmitting information and common understanding from one person to another (Keyton, 2011). So the meaning highlights the fact that when a common understanding derives from the exchange of information then only it will be considered as communication. Sender and the receiver are the two common elements in every communication exchange. However, the communication is usually initiated by the sender. In communication the sender is the individual who has a need or desire to convey an idea or concept to others. The *receiver* is the individual who receives the message from the sender. The sender can select words, symbols or gestures to compose a message.

The *message can be in* verbal, nonverbal, or written language. The *medium* or channel of message is called the carrier of the communication. The medium can be a face-to-face conversation, telephone call, e-mail, or written report. Message is very often distorted by *noise.* Different perceptions of the message, language barriers, interruptions, emotions, and attitudes are examples of noise. Feedback is the process in the communication of message and it occurs when the receiver responds to the sender's message and returns the message to the sender. Feedback allows the sender to determine whether the message has been received and understood.

The elements in the communication process determine the quality of communication. A problem in any one of these elements can reduce communication effectiveness (Keyton, 2011). For example, information must be encoded into a message that can be understood as the sender intended. Selection of the particular medium for transmitting the message can be critical, because there are many choices.

For written media, a researcher may choose from books, magazines, newspapers, manuscripts, reports, bulletin boards, handbooks, newsletters, and the like. For verbal media, choices include face-to-face conversations, telephone, public address systems, closed-circuit television, tape-recorded messages, sound/slide shows, for written communication e-mail, lettersare the best mode of communication. Nonverbal gestures, facial expressions, body language, and even clothing can transmit messages. People decode information selectively as per their choice.

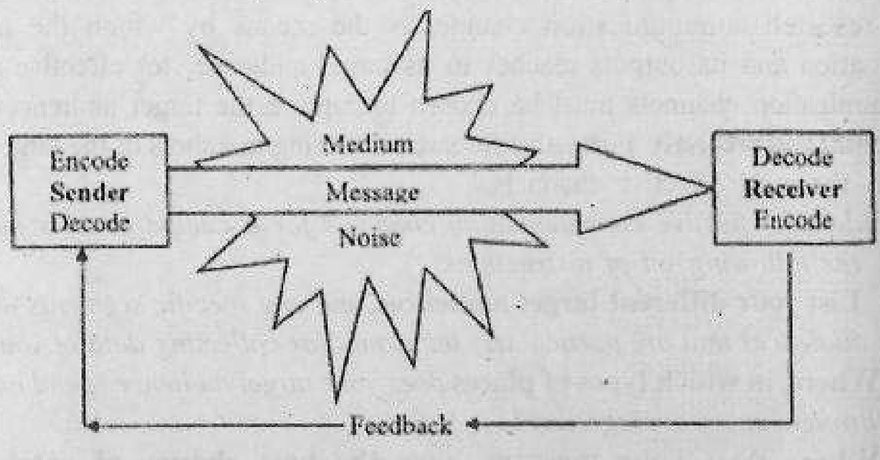


Figure 1. The communication process.

13.2 Definition

Research communication has been defined as 'the system through which research and other scholarly writings are created, evaluated for quality, disseminated to the scholarly community,and preserved for future use'.

It interprets or translates complex research findings into a language, format and context that non-experts can understand. It is far beyond the mere dissemination of research results. It is a network of participants and beneficiaries. Researchers themselves, journalists, editors and their media, intermediaries who provide links between stakeholders form an interdependent network, linking their differing roles in the communication process. Policy makers, government, user organizations and the individual beneficiaries are all the potential users of research whose information needs have to be addressed in very different ways and within very differing contexts. They also require opportunities to articulate their own needs so that communication is driven by demand rather than from the top down.

Communicating research is unlike marketing and promoting a product or service; it is rather a process that transforms raw research outputs into something that addresses the expressed needs of beneficiaries. It can have a vital advocacy role: relevant and timely information can result in positive interventions by policy makers and governments.

13.3 Communication Channels

A research communication channel is the means by which the message of information and its outputs reaches to its target audience, for effective messaging, communication channels must be chosen to capture the target audiences' attention frequently and precisely. Formative research involving members of the target audiences reveals the most effective channels.

To identify effective communication channels for a campaign it is important to follow the following set of instructions:

a) **List your different target audiences,** *and any specific segments within those target audiences that are particularly important for collecting data of your research.*

b) **Where, in which types of places** *does your target audience spend most of their time throughout an average day?*

c) **Where does your message have the best chance of catching their**

attention? *Consider both "real life" venues, such as public transport, markets and shopping malls, and "virtual" spaces such as radio channels, TV programs and web-based social networks.*

d) Based on this analysis, determine **which of the communication channels identified you can afford, or gain external support for.**

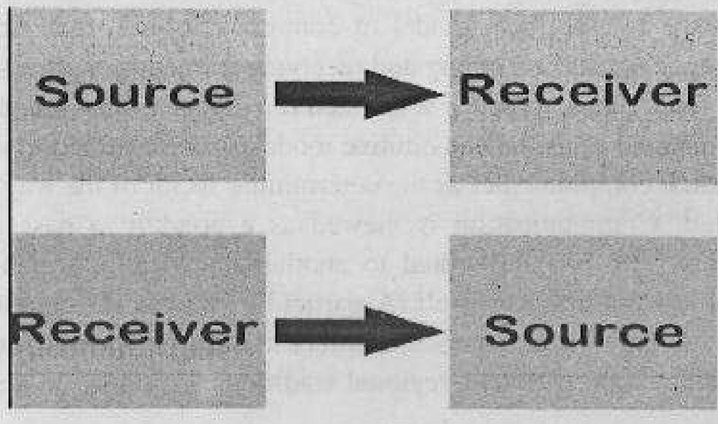
13.4 Importance of research communication

Research can provide the knowledge to empower the common public to take greater control and betterment of their lives. Research communication communicates the relevant issues and brought to the attention of both policy makers and people who will be benefitted by the outcome of the research. Effective and innovative research communication is a vital element in ensuring that research makes a difference. Without it, a lot of research effort is wasted. So research communication as a two way process benefits both the research scholar and the end users.

13.5 Communication Models :

Models of communication are conceptual models used to explain the human communication process.

The first major model for communication came in 1948 by Claude Elwood Shannon and published with an introduction by Warren Weaver for Bell Laboratories. Following the basic concept, communication is the process of sending and receiving messages or transferring information from one part (sender) to another (receiver).



13.51 Berlo Model:

In 1960, David Berlo created the Sender-Message-Channel-Receiver (SMCR) Model of research communication. The SMCR Model of Communication separated the model into clear parts and has been expanded upon by other scholars.

13.52 Schramm Model:

Schramm described research communication along a few major dimensions: Message; what type of things are communicated, source; sender / encoderby whom, form;in which form, channel; through whichmedium, destination; receiver / target / decoder;to whom, and receiver.Wilbur Schramm also indicated the value of examining the impact that a message has (both desired and undesired) on the target audience. Between parties, communication includes acts that confer knowledge and experiences, and ask questions. These acts may take many forms, in one of the various manners of communication. The form depends on the abilities of the group communicating. Together, communication content and form make messages that are sent towards a destination. The target can be oneself, another person or being, another entity.

Research communication can be governed by three levels of semiotic rules:

1. Syntactic (formal properties of signs and symbols),
2. Pragmatic (concerned with the relations between signs/expressions and their users)
3. Semantic (study of relationships between signs and symbols and what they represent).

13.53 Barnlund Communication Model

In 1970 Barnlund proposed a transactional model of research communication.The basic premise of the transactional model of communication is that individuals are simultaneously engaging in the sending and receiving of messages. In a slightly more complex form, a sender and a receiver are linked reciprocally. This second attitude of communication, referred to as the constitutive model or constructionist view, focuses on how an individual communicates as the determining factor of the way the message will be interpreted. Communication is viewed as a conduit; a passage in which information travels from one individual to another and this information becomes separate from the communication itself. A particular instance of communication is called a speech act. The sender's personal filters and the receiver's personal filters may vary depending upon different regional traditions, cultures, or gender; which may alter the intended meaning of message contents. In the presence of "communication noise" on the transmission channel reception and decoding of content may be faulty, and thus the result may not have the desired effect.

Theories ofco-regulation describe communication as a creative and dynamic continuous process, rather than a discrete exchange of information. Canadian media scholar Harold Innis had the theory that people use different types of media toolsas research communicate and which one they choose to use will offer different possibilities for the shape and durability of society.

13.6 Functions of research communication:

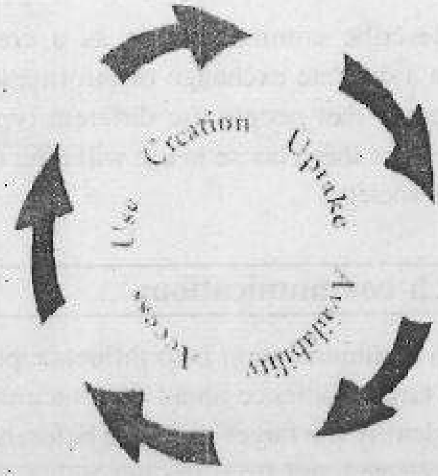
The sole purpose of research communication is to influence people to get data for the research and to inform the target audience about the outcome of the research. It is therefore important that we identify the target audience beforehand from whom the data for the research will be collected and from the beginning make them a part of the process so they are aware about the day to day progress of the research.

13.7 Accessibility of research information or research in progress

The quality of the science may not be the only thing that influences decision making, even if a robust culture of science and evidence exists. There is a need to make existing information more accessible and to analyze and synthesize research to provide tailored information services. There is also a need for more harmonized and effective communication of research acrossresearch scholars, institutions using agreed language, tools and standards.

13.8 Strengthening user demand

The traditional approach to the research process considers the users of the research outcome as passive recipients of information, but effective research communication includes them from the very beginning in shaping the research process and responding throughout the research cycle. Since the users are engaged at an early stage of the research process and not as passive and appreciative recipients of knowledge they feel their involvement as stakeholders, with a sense of involvement and ownership of the research process. Research communication regards the involvement and response of users of research information as an essential componentin dissemination of research information.



13.9 Knowledge management cycle

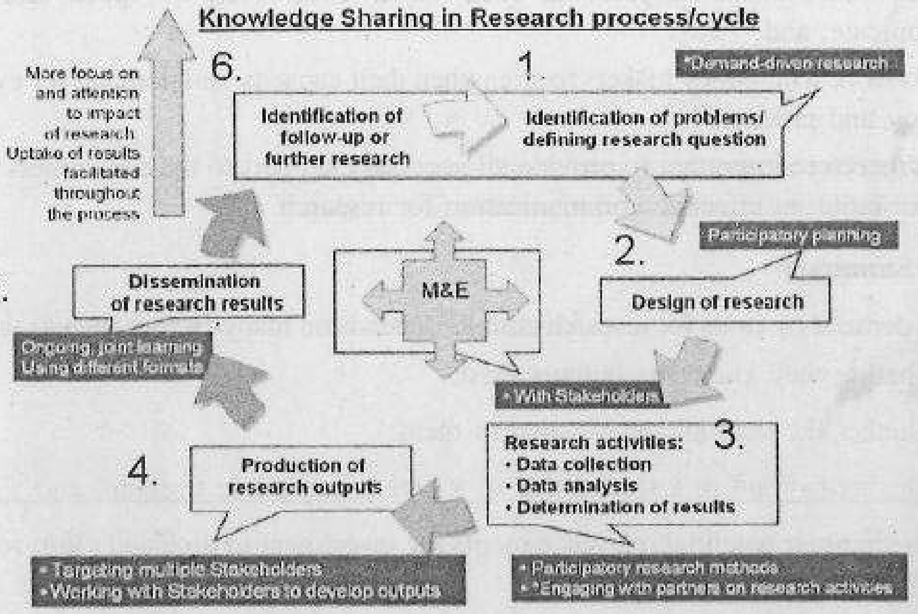
For any research Identification of the problem is the first step in communication process. It is important to identify the problem and define the research questions. In the communication chain the next process is design of research. In this stage research design is made and goes to the stake holders. Here the most crucial part of research communication is data collection, data analysis and determination of results. Once the data analysis is done then it goes to multiple stakeholders to production of research after this stage the result of the research is disseminated, after making necessary followups and getting necessary feedback further research is taken into consideration.

The five perspectives do give coherence to the richness and vitality of the research communication experience. The themes are addressed under the following headings:

- Strengthening user demand;
- The role of intermediaries in research uptake;
- Monitoring and evaluation tools;
- The role of the media in research uptake;
- Implications for resourcing successful research communication;

• Engagement with the private sector; and

• The potential of information and communication technologies.



13.9.1 Networking and research communication

When research takes place in an organizational setup then a Network of people is established to conduct the research. For speedy and effective communication the network does not have to be large. Within large research organizations the team is set up in house involving people of various departments.The organizations encourage research teams to set aside time regularly to share learning, progress, problems and feedback. This would prove a simple and valuable exercise for the team involved in research. They do not even have to be just within one country or institution. The communication of research can be of face-to-face meetings, exchange visits, peer reviews, exchange of report summaries, or email, telephone or Skype conferences with peers in other countries. More formal initiatives, staff time permitting, could include an organizational intranet or an internal e-newsletter or update.

13.9.2 Facilitate the researchers to communicate

It was noted that support to help researchers to communicate more effectively is needed in three areas;

1. To improve incentives for researchers to communicate beyond their normal audience of other researchers;

2. To build skills at personal and institutional levels to more effectively communicate; and

3. To work with policy makers to strengthen their capacity and demand for evidence in policy and practice.

It is therefore important to provide all necessary support to the researchers so that they can build an effective communication for research.

13.9.3 Summary

The demand by users for research outputs depends on many factors such as whether:

• Whether they know the outputs exist;

• Whether the outputs are of value to them;

• Whetherthey are in a format that they feel is accessible to them; and

• Whethertheir potential relevance merits the investment in time and effort required to absorb and apply them.

Building capacity for research-based evidence requires effort at individual, organizational and institutional levels through a complex mix of real and virtual partnerships, networks and support tailored to the circumstances. With user demand and capacity both strengthened, whether or not research evidence actually feeds into improved decisions, policies and practices also depends on the specific context, political pressures, individual personalities and timing.

Researchers who make the challenging decision to engage with users early in the research cycle through proper communication, strategically and imaginatively, can achieve very positive results. So research communication is one of the most important components in fulfilling the demand of the target audience. It is therefore necessary to consider the research communication and its various components from the beginning of the research.

The research communication should have stimulated demand and understanding within their user groups. There are many opportunities to extend these processes with end user groups that form the ultimate beneficiaries of research and policy initiatives, such as communities, educational and civil society organizations. Relatively simple and low-cost approaches such as exchange visits and ICT techniques offer considerable potential.

13.9.4 Know yourself:

1. Define research communication. What is the purpose of research communication
2. What are the instructions that you follow to identify the channels of research communication
3. What are the models of research communication? Describe any one of them.
4. Describe knowledge management cycle in research communication
5. What is communication barrier? Name few barriers that slow down the process of research.

13.9.5 Further reading

1. Barnlund, D. C. (2008). A transactional model of communication. In. C. D. Mortensen (Eds.), *Communication theory* (2nd ed., pp47-57). New Brunswick, New Jersey: Transaction.
2. Berlo, D. K. (1960). *The process of communication*. New York, New York: Holt, Rinehart, & Winston.
3. *Crag, Robert T. (1999). Communication Theory as a Field. International Communication Association*
4. *Craig, Robert T. (May 1999). "Communication Theory as a Field" (PDF). Communication Theory. Blackwell Publishing Ltd.; International Communication Association. 9 (2): 119–161.*
5. *Chandler, Daniel (1994). The Transmission Model of Communication. University of Western Australia. Retrieved 11.06.2011.)*
6. Littlejohn, S.W. and Foss, K.A. (2008). Theories of human communication, 9th edition. Belmont, CA: Thomson Wadsworth.
7. *Miller, Katherine (2005). Communication Theories: Perspectives, Processes, and Contexts 2nd Edition. Boston, MA: McGraw-Hill Higher Education. ISBN 0-07-293794-7.*
8. Schramm, W. (1954). How communication works. In W. Schramm (Ed.), *The process and effects of communication* (pp. 3-26). Urbana, Illinois: University of Illinois Press.

Unit 14 ❑ Research Promotional Agencies

Structure

The UNESCO defines R&D as - 'R&D is any creative systematic activity undertaken in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this knowledge to devise new applications'. The United Nations statistics division defines R&D as - 'Research and Development by a market producer is an activity undertaken for the purpose of discovering or developing new products, including improved versions or qualities of existing products, or discovering or developing new or more efficient processes of production'.

India is one of the oldest civilisations in the world, has-one of the longest traditions of research and writing. It has a very rich history dating back several millenniums. Knowledge was preserved and propagated through an oral tradition. In this context, the teachers set up 'residential schools' in their own homes. Students were to live with the teacher and his family and were expected to share the daily chores of the family. Sanskrit was the language of the educated and the texts were composed in this language. Most of the major modern languages in India are derived from Sanskrit. During the rules of Buddhist kings belonging to the Mauryan dynasty in the third and second century BC India flourished with the establishment of institutions of learning. Taxila, now in Pakistan, became the seat of learning where scholars journeyed to learn and to be educated. Nalanda in eastern India became famous for the Buddhist University where several religious conclaves were held. In the 10th century, India was invaded from the northwest and many founded their dynastic rule in India. Persian became the court language and the educated elites became conversant in Farsi and Arabic. The dual traditions of Sanskrit and Farsi education were kept alive till the colonization of India by the British. The British established schools to teach English and the sciences. In 1857 three universities were established in three metropolitan cities, Bombay (now Mumbai), Calcutta (now Kolkata) and Madras (now Chennai) following Oxford or Cambridge as models. Another university was established in 1887 in Allahabad. These universities imparted education in the liberal arts and sciences. The main objective was to prepare people for careers in the civil service, legal profession and in medicine. The need for technical education was also felt by the British, who established the first industrial school attached to the Gun Carriage Factory in Guindy, Chennai, in 1842.

With this varied history of the higher education system, the current system is primarily modeled after the British system. However, some the technical institutions in engineering and management are modeled on the US system. The higher education system remains primarily the responsibility of the state governments, although the central government has taken the initiative in establishing and funding a few central universities and other institutions of national repute. India has 14 major languages. Institutions of higher education use English as the medium of instruction for most courses, particularly in the technical fields, though the regional language remains a major cultural artifact that provides the cultural context. The institutional framework of higher education in India is complex. There are several types of institutions: universities, colleges, institutions of national importance, post-graduate institutions and polytechnics. Only the universities are generally authorized to grant degrees. By special acts of Parliament, the institutions of national importance have been authorized to grant degrees. Post-graduate institutions and polytechnics can grant diplomas and are to be recognized by the All India Council of Technical Education. Universities are of four types: state universities, central universities, deemed universities (aided and unaided), and private universities.

According to twelfth five year plan report "Research and innovation are now vital functions of higher education worldwide. The value of interdisciplinary research is recognised globally, as innovation is now happening at the intersections of disciplines. Collaboration is now central to innovation. Entrepreneurship that leverages innovation is also an increasingly integral part of higher education systems. While all (Higher Education Institutes) HEIs cannot be expected to become research-based institutions, it is vital that the country promote a research culture across all institutions while ensuring special support for those able to engage in state-of-the-art research. The HEIs should contribute to the nationalinnovation agenda, even when they are not research intensive—albeit in different ways. Teaching-focused institutions must train their students in the techniques of research so that the doors to research-based graduate education and employment are opened to them. Vocational institutions must enable the future workforce to engage at least in the 'development'component of R&D. It is essential that all institutions equip their graduates with core skills of critical thinking, communication, collaboration and creativity to enable the country to continuously innovate to adapt to new environments".

India's research performance turned around in the last two decades, after over a decade of stagnation. An improvement in scientific output is evident both in absolute terms and relative to the comparison group. During the past 10 years, India's overall share of publications in the world has risen from 2.8 per cent to 3.4 per cent, with a significant improvementin researcher productivity since 1999. India produces over twice as many scientific publicationsa year than it did a decade ago. Though dwarfed by China's achievements, India's output of publications has grown faster than that of Brazil and Russia.

There are indications that research quality has improved as well. India's publications have accumulated 16,10,511 citations with 5.77 citations per paper, better than China, but still low compared to the world average of 10.81 citations per paper. The relative impact rose from 0.48 to 0.66 (world average being one). In 2009, India stood eleventh in terms of the number of papers published, seventeenth in terms of the number of citations, and thirty-fourth in terms of number of citations per paper as per the ISI Web of Science.

Notwithstandingsuch achievements, Indian higher education continues to have limited research capacity. Low levels of funding and segregation of the country's R&D institutions from universities and colleges have been responsible for the weak research capacity of Indian universities. It is disappointing to note that even the country's top universities remain largely teaching-focused with limited research and doctoral education.

This lack of research orientation, even in the best of the Indian institutions, is reflected intheir standing in global rankings, most of which rely heavily on measurable indices of research performance. No Indian university figured amongst the top 200 universities in the Times Higher Education (THE) Rankings or the Academic Ranking of WorldUniversities (ARWU) for the year 2011. While it is neither necessary nor realistic to expect all institutions to achieve high levels of research excellence, a natural pyramid of quality excellence suggests that, if the average quality improves, then the best will enter the top leagues of research-intensive universities. India's output in PhDs was small at 10,781in 2008–09, when compared against international peers. The total number of PhDs in science and engineering at 4,500 is miniscule as compared to the approximately 30,000 and 25,000 for China and the USA, respectively. In terms of innovation and the creation of intellectual property, Indians file and receive only a small number of worldwide patent applications (merely 11,937 applications filed by Indians compared to 2,41,546 by Chinese in 2009)and no Indian academic institution figures in the list of top applicants for patent filing.

Output measures related to publications,patents/licensing and spinoffs can provide some indications of research and innovation performance for research intensive institutions though even for them, these would be too narrow for gauging overall research performance. For less research intensive institutions, their contributions to innovation and economic development could derive from much less visible activities such as faculty consulting or development projects or education to instil students with creativity and entrepreneurship.

14.1 Objectives

Objective of this unit is to aware a researcher

- The initiatives which have been taken for enhancement of R&D in India.
- The agencies such as various Councils under different Ministries which are acting as advisory body as well as funding R&D in India.
- Guidelines provided by the agencies for writing research proposals as well as reports.
- Also acting as the repositories of theses and dissertations.
- Policies taken by Government of India for R&D.

14.2 Introduction

The current national innovation system in India is a vast and complex system comprised of knowledge producers such as science and technology institutions, social sciences institutions,humanities institutions,academia, and innovating individuals and knowledge users (e.g., industry-production/services in the public and private sectors).

Thus India has evolved a large publicly funded R&Dstructure. There are various regulatory bodies and research structures under various Ministries, which cater to different research areas and which are distributed around the country. Examples include:

- Council of Scientific and Industrial Research (CSIR; csir.res.in): established in 1941; 39 laboratories
- Indian Council of Agricultural Research (ICAR; www.icar.org.in): established in 1929; 99 institutes and 17 research centres
- Indian Council of Medical Research (ICMR; icmr.nic.in): established in (1911); 30 laboratories
- Defence Research & Development Organisation (DRDO; drdo.gov.in): established in (1958); 48 laboratories
- Indian Council of Social Science Research (1969)
- Indian Council of Historical Research(1972)
- National Council of Educational Research and Training (1993)
- All India Council for Technical Education (1987)
- University Grants Commission (1956)

The Government of India declared 2010-2020 as the "Decade of Innovation", for which the roadmap would be prepared by the newly established National Innovation Council (NInC; innovationcouncil.gov.in). The National Innovation Council is"the first step in creating a crosscutting system which will provide mutually reinforcing policies, recommendations and methodologies to implement and boost innovation performance in the country" (NationInnovation Council, 2010). *The Science, Technology and Innovation Policy 2013* outlines the major policy initiativesto strengthen the innovation ecosystem and give aboost to the development of innovation-led entrepreneurshipin India:

"The guiding vision of aspiring Indian STI [Science,Technology, and Innovation] enterprise is to accelerate the pace of discovery and delivery of science-led solutions for faster, sustainable and inclusive growth. A strong and viable Science, Research and Innovation System for High Technology-led path for India (SRISHTI) is the goal of the new STI policy." (Ministry of Science and Technology, 2013).

Fund released state-wise by Government of India(2013-14 to2015-16) for research and development is given in the following table:

State-wise Funds Released for Improvement and Enhancementof Capabilities of Universities/Institution in Researchand Development in India (2013-2014 to 2015-2016)

(Rs. in Lakh)

States/UTs	2013-2014	2014-2015	2015-2016
Andaman and Nicobar Islands	209.44	83.62	124.59
Andhra Pradesh	1449.18	2490.83	-
Arunachal Pradesh	14.28	44.00	-
Assam	657.74	542.20	535.96
Bihar	55.51	464.17	6.60
Chhattisgarh	111.22	355.50	21.64
Delhi	4951.96	5193.64	6366.37
Goa	206.36	343.65	-
Gujarat	771.50	767.89	482.82
Haryana	1016.71	1644.38	1218.57
Himachal Pradesh	538.25	93.60	219.74
Jammu and Kashmir	455.91	1134.34	739.66
Jharkhand	329.21	127.52	5.75
Karnataka	15391.40	7580.66	5990.47
Kerala	3290.48	1717.42	927.79
Madhya Pradesh	477.63	1068.12	83.15
Maharashtra	5548.51	4085.13	2845.17
Manipur	352.12	379.63	118.24
Meghalaya	141.75	453.11	16.75
Mizoram	11.40	65.16	-
Nagaland	-	34.59	-
Odisha	985.21	893.20	653.83
Puducherry	285.25	240.40	254.36
Punjab	2136.15	1689.56	1268.05

States/UTs	2013-2014	2014-2015	2015-2016
Rajasthan	220.41	1305.79	134.07
Sikkim	57.57	-	-
Tamil Nadu	5974.52	6294.51	4159.52
Telangana	1678.66	1379.15	1266.87
Tripura	60.90	149.00	-
Uttar Pradesh	4384.52	2374.67	2107.30
Uttarakhand	603.06	494.27	187.67
West Bengal	4656.31	3474.73	1806.02
India	**57024.12**	**46964.44**	**31540.96**

Source : Lok Sabha Unstarred Question No. 1798, dated on 04.05.2016.

14.3 Research Promotional Agencies: At A Glance

Different regulatory authorities under various ministries of Government of India mentioned below strive to promote teaching and research in emerging areas in Humanities, Social Sciences, Languages, Literature, Pure Sciences, Engineering & Technology, Pharmacy, Medical, Agricultural, Sciences etc.

1. Central Council for Research in Siddha

http://www.siddhacouncil.com/

For the development of Siddha system of Medicine, Govt. of India, by bifurcating the erstwhile CCRAS, formed CCRS (Central Council for Research in Siddha) with its headquarters in Chennai and five Research Institutes/Units in three states namely, Tamil Nadu (Chennai, Mettur&Palayamkottai), Puducherry (Puducherry) and Kerala (Thiruvananthapuram). Siddha is a science of holistic health emphasising both drug and diet for human health care.

The Council has the vision of preservation and transmission of Knowledge and enhancement of the quality of research for developing drugs with quality, safety and efficacy through well-established preclinical and clinical research facilities — to prevent / manage /cure the diseases of varied aetiology.

To undertake scientific research works in Siddha in a time-bound and cost-effective manner, to coordinate, aid, promote and collaborate research with different units of sister Councils and Research Organizations.

To publish research articles/research journals, to exhibit achievements and to propagate research outcomes for all the stakeholders. To provide consultancy services for research projects and drug development (adopting both classical and modern techniques/equipments for Diagnosis, evolving evidence based Siddha drug treatment/ therapy and promoting Siddha drug manufacture in collaboration with the other technical organizations)

Publication :

1. **E book:Research in Siddha system of Medicine: science of holistic health/ Central council for Research in Siddha**
2. **E book: Yoga in Siddha/Central council for Research in Siddha**
3. **CCRS guide lines for Siddha Practioners for Clinical Management of Dengue fever**
4. **Prnited books in Tamil, Hindi and English languages.**
5. **Annual reports.**

2. National Council for Science and Technology Communication (NCSTC)

http://www.dst.gov.in/scientific-programmes/st-and-socio-economic-development/ national-council-science-technology-communication-ncstc

The National Council for Science and Technology Communication (NCSTC) is mandated to communicate Science and Technology to masses, stimulate scientific and technological temper and coordinate and orchestrate such efforts throughout the country. The programmes of NCSTC aims at building capacity for informed decision making in the community and promote scientific thinking. It is devoted towards societal upliftment through the dissemination of scientific knowledge in an informed manner and builds programmes with the help of different media which percolate down to every nook and corner of the society.

The NCSTC focuses on outreach activities, training in Science and Technology communication, development, production & dissemination of S & T software, incentive programmes, and field based Sci-Com projects, research in S&T communication, international co-operation, motivating students and teachers, environment awareness and programmes with a special component exclusively for women.

Some of its important successful initiatives, over the years include the campaigns over the Year of Scientific Awareness, Year of Physics, Year of Astronomy, Year of

Mathematics, observation of the National Science Day and National Mathematics Day, the National Children's Science Congress, National Teacher's Science Congress, and Science Express etc.

A multi- pronged effort has been developed by the NCSTC including:

- Communicating science using folk media;
- Use of mass & digital media for science communication and popularization;
- Use of Social media in science and Technology Popularization

Some important initiatives of the division are:

- Science Media Research Initiatives
- Eco Media & Eco Next initiatives
- Science Express Climate Action Special (15th Oct - 07th May 2016)
- National Children's Science Congress
- National Teacher's Science Congress
- Initiative for Research and Innovation in Science (IRIS)
- India Innovation Initiative-i3
- Regional innovation Science Hubs for Innovators (RISHI)
- Health and Nutrition through Community Radio
- Academic/Training courses in S& T communication
- Science and Technology Communication Programmes for Students
- Explaining Science behind Miracles

Major Programmes:

Science Express is a unique mobile science exhibition mounted on a 16-coach, AC train, travelling across India since 2007. Science Express showcased the fascinating world of science, groundbreaking discoveries, cutting-edge science & future-oriented technologies for initial 4 years. In 2012, Science Express was redesigned on the theme 'Biodiversity' and for 3 years it was running as Science Express Biodiversity Special (SEBS), aimed at creating awareness on Biodiversity of India. Science Express has covered over 122,000 km across the country, receiving more than 1.33 crore (13.3 million) visitors at its 391 halts, over 1,404 days. It has thus become the largest, longest running and most visited mobile science exhibition, probably in the world and has created several records in its wake. The next phase of Science Express is soon coming up as Climate Change Special. More information on ***www.sciencexpress.in***

• National Children's Science Congress

This is an opportunity for brilliant young scientists (10 -17 years of age group), started since 1993 to popularize the method of science with following steps:

- work in teams under a guide on an identified theme
- select a problem from the neighborhood
- develop a hypothesis and conduct field research
- see patterns in data and prepare a report
- present findings before peer group in one's own language.

Innovative projects are selected for district congress, state congress and national convention. Each state also sends two delegates to the annual session of the Indian Science Congress.

The NCSC has been organized in almost all districts every year with more and more schools joining the activity. Developing an activity book, training workshops for guide teachers, evaluating the research projects and coordinating the district and state level conventions are enriching experiences for a number of organizations.

• National Science Day

Every year on February 28, is celebrated as the National Science Day. The programme was initiated by department to trigger science popularization activities throughout the country and to disseminate scientific education about the current issues of science and technology amongst the citizens of the country. The programmes starts around the national science day and activities like lectures, quiz, radio, television shows, open houses and debate etc. are organized around a central theme. These programmes are organized nationwide through the state S&T Councils, science and Technology Departments.

• Eco Media & Eco Next initiatives

Effective science communication can change ways of community learning, peoples' access to knowledge & resources, and societal paradigms. The science behind management of ecological resources, habitat & ecosystems services is central to our well being and ability to work. It thus addresses critical life cycle needs and complexities of social, economic and ecological nature. In today's fast-paced world, role of media, both in printed and electronic forms, is critical in leveraging furthering these goals. The initiatives would support a direct and purposeful communicative paradigm, as Science Media & Eco Next Youth leadership and promoting informed choices and decision making. The Eco Media & Eco Next initiatives focus upon promoting real-time responsiveness and eco-media leverage for promoting integrated efforts for conservation of natural resources, specifically aimed at –

- Developing models in public guidance systems based on science communication, like location specific innovative initiatives for actionable learning and building field capacity for adopting scientific & best practices in knowledge critical domains.
- Knowledge led motivation of youth for leadership and improvement of quality of life of specific target groups based on scientific approaches of 'Being- on-their -Own' and 'Collective response' to challenges and location specific problems.
- Serving national priorities of Government of India like Swachh Bharat, Access to Safe Water, Ecological Health of Rivers, etc.
- S&T Awareness for conservation of Resources &Sustainable development

STARS Nationwide training cum awareness programmes on understanding weather, climate change and disaster preparedness for sustainable development; **Grameen Vigyan Jagriti.**

A programme on rational use of pesticides, fertilizers, demystification and awareness generation about malpractices like female feticides, food safety, genetically modified crops, Biodiversity awareness related programmes.

India Innovation Initiative - i3 is a Public-Private Partnership (PPP) initiative of NCSTC. Each year hundreds of innovations are received from across the country contributors include students, industry professionals, individual and grassroots innovators in various fields such as Life Sciences, Electronics & Communication, IT, Energy and Engineering Technologies. It is a national level competition for promoting the spirit of science & technology in the country. Students enrolled in AICTE approved institutions and others are eligible.

- **National Teachers' Science Congress (NTSC)**

Initiated by NCSTC, the NTSC provides a forum to teachers of the country to enhance their level of scientific awareness. The first national event was held in 2003. Presently it has become a biennial activity of NCSTC. Besides teachers educators of vocational/ open schools/teachers educators/DIET faculty/B.Ed and university researchers, scientists, technologists and activists are also eligible to submit their papers.

- **Mathematics Awareness Resources & Initiatives (MARI)**

Various initiatives are encouraged to enthuse, motivate and inculcate, a positive attitude to learn mathematics in the younger generation ranging from identification and introduction of best teaching practices and resource persons; Capacity building of in-service teachers & students through training/ camps in Mathematics; Development, production & Dissemination of teaching learning materials (TLMs) like posters, Mathematics kits, resource manuals, booklets, models, etc. for Mathematics and research in related areas.

Every year since 2012, birthday of SrinivasaRamanujan, a great mathematician is celebrated as "National Mathematics Day" with the help of state S&T Councils, throughout the country.

- **Explaining Science behind Miracles**

There is exposure on the training module aiming at providing basic skills in performing & investigating such miracles and based on science behind it. Miracles are performed by so called Godmen who want to mislead the gullible, but fact is that all miracles are based on various principles of science.

- **Motivational Programme**

To encourage bright students to select careers in science it is imperative to sensitize them about science & technology related work going on various laboratories, universities and scientific institutions in the country.

To encourage bright students to select careers in science, the National Council for Science & Technology Communication, New Delhi invites proposals from research laboratories, universities, Aided institutions of the Department of Science & Technology, GOI, New Delhi, State S&T Councils and other interested scientific institutions for hosting motivational programmes in their institutions.

Focus: The main objective is to kindle interest of the students in taking up pure sciences as academic and career option, students are sensitized about the science & technology related work going on with an exposure of various sophisticated instruments, facilities in S&T laboratories;

Duration: Duration of the programme is one week; which can also include a one day open house for the parents and other target groups, the programme will be highly interactive and preferably residential.

Target: Students of X & XI th standard with science background (for one week program), teachers and parents (for a day long interaction only)

Proposals are invited as per the timeline for NCSTC proposals i.e, in April and October months each year from universities, laboratories, coordinating field agencies. The same should be submitted as per NCSTC guidelines and format.

Many significant initiatives have given rich dividends to the nation. Some of these include:

Audio/Radio Programmes

Science based radio serials on various topics are produced and broadcasted by119 stations of All India radio (AIR) and through community radio. Some of the popular serials are: "JeevanEk-RoopAnek", Rahiyematwaley, "AnkonkeKhiladi", Radio Mathematics (on community radio).

Video Programmes

Video programmes were made and telecast on DD. They are Challenge Chatani, "Utthan", Ignited Minds I-II, Hum HongeKamyab, Fossil Memoirs etc. Video CD's on achievements of S&T in India were also developed and telecast.

Web based programmes

- **Interactive Web Portals on Mathematics**
- **Web Portal on Planet Earth**

GanitYatra in the state of Maharashtra

A state level campaign in the tribal districts of Maharashtra was undertaken as "GanitYatra", a regional jatha Supported by NCSTC. Resource material in terms of dramas, CDs, ganitgeet, booklets, calander etc. were brought out.

Training Manuals & Kits Developed

Mathematics kit, Weather Kit, Simple Tasks - Great Concepts, LASER kit for Students, Parakhi, Module on Microorganisms, Vermi-Composting, Astronomy kit, Biodiversity kit etc. are made and disseminated among the children

Publications:

- General Guidelines for NCSTC
- Guidelines for filling post sanction details for NCSTC Programmes
- GFR 19a
- Revised guidelines for Utilization Certificate Submission
- Timelines for receipt of proposals in NCSTC
- Programme Not Supported
- Theme for the National Science Day 2016
- Projects

Library : Library has a collection of 22000 books . Library is subscribing 110 scientific and technical journals. Library provides OPAC service.

3. National Innovation Council:

http://innovationcouncilarchive.nic.in/

The former Prime Minister set up the National Innovation Council (NInC) under the Chairmanship of Mr. Sam Pitroda to discuss, analyse and help implement strategies for inclusive innovation in India and prepare a Roadmap for Innovation 2010-2020. NInC would be the first step in creating a crosscutting system which will provide mutually reinforcing policies, recommendations and methodologies to implement and boost innovation performance in the country.

The National Innovation Council is working with the Ministry of Human Resource Development on a proposal for the creation of 20 Design Innovation Centres, an Open Design School and a National Design Innovation Network that will connect these new design schools together, along with a wide range of stakeholders. The goal is to increase the reach of design education and promote wide-ranging design innovation.

Initiatives:

1.Industrial Initiatives: The Innovation Clusters initiative seeks to bring out the needs of both industry and academia and provide a means of addressing the needs through pro-active regional ecosystems of collaboration. NInC will catalyse and facilitate creation of such innovative clusters through Cluster Innovation Centers (CICs) which will act as hubs for connecting various regional/national actors and stakeholders in symbiotic relationships. These CICs will then drive need based innovation in the clusters connecting demand to supply and sharing information/knowledge among the stakeholders.

Publication: Book-National Innovation initiatives2010-2014.

Resources Library: Library has a collection of presentations,reports etc.

4. Indian Council of Agricultural Research

(http://www.icar.org.in/)

The Indian Council of Agricultural Research (ICAR) is an autonomous organisation under the Department of Agricultural Research and Education (DARE), Ministry of Agriculture and Farmers Welfare , Government of India. Formerly known as Imperial Council of Agricultural Research, it was **established on 16 July 1929** as a registered society under the Societies Registration Act, 1860 in pursuance of the report of the Royal Commission on Agriculture. The ICAR has its headquarters at New Delhi.

The Council is the apex body for co-ordinating, guiding and managing research and education in agriculture including horticulture, fisheries and animal sciences in the entire country. With 101 ICAR institutes and **71 agricultural universities** spread across the country this is one of the largest national agricultural systems in the world.

It has played a major role in promoting excellence in higher education in agriculture. It is engaged in cutting edge areas of science and technology development and its scientists are internationally acknowledged in their fields.

Publication: Annual reports, Project reports, Abstracting journals, Magazines etc.

Resource Centre/ Library: Library has a collection of 85000 documents including books, reports, journals. Library provides Web OPAC service.

5. Council of Scientific and Industrial Research

http://www.csir.res.in/

The Council of Scientific & Industrial Research (CSIR), known for its cutting edge R&D knowledgebase in diverse S&T areas, is a contemporary R&D organization. Having pan-India presence, CSIR has a dynamic network of 38 national laboratories, 39 outreach centres, 3 Innovation Complexes and 5 units. CSIR's R&D expertise and experience is embodied in about 4600 active scientists supported by about 8000 scientific and technical personnel.

CSIR covers a wide spectrum of science and technology – from radio and space physics, oceanography, geophysics, chemicals, drugs, genomics, biotechnology and nanotechnology to mining, aeronautics, instrumentation, environmental engineering and information technology. It provides significant technological intervention in many areas with regard to societal efforts which include environment, health, drinking water, food, housing, energy, farm and non-farm sectors. Further, CSIR's role in S&T human resource development is noteworthy.

Pioneer of India's intellectual property movement, CSIR today is strengthening its patent portfolio to carve out global niches for the country in select technology domains. CSIR is granted 90% of US patents granted to any Indian publicly funded R&D organization. On an average CSIR files about 200 Indian patents and 250 foreign patents per year. About 13.86% of CSIR patents are licensed - a number which is above the global average. Amongst its peers in publicly funded research organizations in the world, CSIR is a leader in terms of filing and securing patents worldwide.

CSIR has pursued cutting edge science and advanced knowledge frontiers. The scientific staff of CSIR only constitute about 3-4% of India's scientific manpower but they contribute to 10% of India's scientific outputs. In 2012, CSIR published 5007 papers in SCI Journals with an average impact factor per paper as 2.673. In 2013, CSIR published 5086 papers in SCI journals with an average impact factor per paper as 2.868.

Publication : Annualreport,Directory,Scientific Papers.

Library(Knowledge Resource Centre): The library of CSIR-HQ was established in 1964 and has gradually become the Knowledge Centre with the emergence of new technologies. It was, therefore, renamed as '*Knowledge Resource Centre (KRC)*' in 2008. The aim of KRC is to meet the growing information needs of the staff.

The KRC has large collection of information resources on Science, Technology, and Management. The collection mainly comprises of print books and journals, electronic resources, and online resources. The KRC provides information services to the users from its print and non-print resources.Library provides Online Public Access Service.

6. Indian Nursing Council

http://www.indiannursingcouncil.org/

Indian Nursing Council was established for following purpose

- To establish and monitor a uniform standard of nursing education for nurses midwife, Auxiliary Nurse- Midwives and health visitors by doing inspection of the institutions.

- To establish and monitor a uniform standard of nursing education for nurses midwife, Auxiliary Nurse- Midwives and health visitors by doing inspection of the institutions.
- To recognize the qualifications under section 10(2)(4) of the Indian Nursing Council Act, 1947 for the purpose of registration and employment in India and abroad.
- To give approval for registration of Indian and Foreign Nurses possessing foreign qualification under section 11(2)(a) of the Indian Nursing Council Act, 1947.
- To prescribe minimum standards of education and training in various nursing programmes and prescribe the syllabus & regulations for Nursing programms.
- Power to withdraw the recognition of qualification under section 14 of the Act in case the institution fails to maintain its standards under Section 14 (1)(b) that an institution recognised by a State Council for the training of nurses, midwives, Auxiliary Nurse Midwives or health visitors does not satisfy the requirements of the Council.
- To advise the State Nursing Councils, Examining Boards, State Governments and Central Government in various important items regarding Nursing Education in the Country.
- To regulate the training policies and programmes in the field of Nursing.
- To recognise Institutions/Organisations/Universities imparting Master's Degree/ Bachelor's Degree/P.G. Diploma/ Diploma/Certificate Courses in the field of Nursing.
- To Recognise Degree/Diploma/Certificate awarded by Foreign Universities/ Institutions on reciprocal basis.
- To promote research in Nursing.
- To maintain Indian Nurses Register for registration of Nursing Personnel.
- Prescribe code of ethics and professional conduct.
- To improve the quality of nursing education.
- Conducts Ph.D. in nursing.

Publication : Brochures, Mannuals, Syllabus, Annual reports, News Letter etc.

7. Dental Council of India

http://www.dciindia.org.in/

Dental Council of India is a Statutory Body incorporated under an Act of Parliament viz. The Dentists Act, 1948 (XVI of 1948) to regulate the Dental Education and the profession of Dentistry throughout India and it is financed by the Govt. of India in the Ministry of Health & Family Welfare (Department of Health) through Grant-in-aid. The General Body of the Dental Council of India representing various State Governments, Universities, Dental Colleges, Central Government, etc

Publication: Rules, Regulations, Acts

8. Indian Council of Medical Research

http://www.icmr.nic.in/

The Indian Council of Medical Research (ICMR), New Delhi, the apex body in India for the formulation, coordination and promotion of biomedical research, is one of the oldest medical research bodies in the world.

The ICMR has always attempted to address itself to the growing demands of scientific advances in biomedical research on the one hand, and to the need of finding practical solutions to the health problems of the country, on the other. The ICMR has come a long way from the days when it was known as the IRFA, but the Council is conscious of the fact that it still has miles to go in pursuit of scientific achievements as well as health targets.

- 1911

  -First meeting of the Governing Body of the Indian Research Fund Association (IRFA) was held on November 15, 1911 (at the Plague Laboratory, Bombay, under the Chairmanship of Sir Harcourt Butler).

  -Articles of the Association were considered and a Scientific Advisory Board was constituted at the same meeting.
- 1912

  -At the 2nd meeting of the Governing Body, a historic decision was taken to start a journal for Indian Medical research.
- 1913-1914

  -The Indian Journal of Medical Research was started in 1913-14 (under the authority of the Director-General, Indian Medical Services).
- 1918-1920

-The 'Beri-Beri Enquiry' was started at Coonoor (under the guidance of Sir Robert McCarrison).

-'Quinine and Malaria Enquiry' was initiated (under Major Sinton at Kasauli).

-Kala-azar Ancillary Enquiry was started (with Major Knowles and Dr. Napier).

-Research on Indigenous Drugs was initiated (under Col. R.N. Chopra at the Calcutta School of Tropical Medicine, Calcutta).

- 1923

  -The first All India Conference of Medical Research Workers was convened at the Calcutta School of Tropical Medicine and Hygiene, Calcutta. (This became an annual event subsequently).

- 1925

  -Research on Nutritional diseases was started at Coonoor (by Col. McCarrison under 'Deficiency Diseases Enquiry').

- 1926

  -IRFA received the first munificient public contribution of Rs.1 lakh from the Maharaja of Parlakimedi.

- 1927

  -Fructification of the plans of Lt.Col. S.R. Christophers for creation of a Central Malaria Organization as "Malaria Survey of India" (by absorbing the Central Malaria Bureau at Kasauli and the Enquiries on Quinine and Malaria and Indian Culicidae).

  -An Experimental Malaria Station was set up at Karnal as a part of Malaria Survey of India.

- 1929

  -The 'Deficiency Diseases Enquiry' was converted into a Centre of Nutrition Research (with Col. McCarrison as its first Director).

  -The Publication of "Records of Malaria Survey of India" was started.

- 1932

  -The Governing Body of IRFA completed the task of setting up the Institute of Hygiene and Public Health at Calcutta.

- 1937

  -A course of training in Nutrition was started at the Nutrition Research Laboratories at Coonoor.

  -"The Nutritive Value of Indian Foods and Planning of Satisfactory Diets" was prepared (which has now been reprinted repeatedly).

- 1938

-IRFA was registered as a local body not administered by the Government on March 22, 1938 under the Government of India Act No. XXI of 1860.

-In tune with the recommendation of the Conference of Far Eastern Countries on Rural Hygiene held in Java in 1937, the Government of India decided that the Nutrition Advisory Committee of the IRFA should also function as the National Nutrition Committee for India.

-The "Malaria Survey of India" was re designated as the "Malaria Institute of India".

-"The Records of the Malaria Survey of India" was re designated as the "Journal of the Malaria Institute of India" (which subsequently became the Indian Journal of Malariology in 1947).

- 1941

  -A Research Fellowship Scheme was started by IRFA.

- 1942

  -Transmission cycle of the parasite of Kala-azar was elucidated by Swaminath, Smith, Shortt and Anderson.

- 1945

  -A Clinical Research Advisory Committee was appointed as a first step to enable greater attention being paid to clinical research and the development of research in medical colleges.

  -A Clinical Research Unit (the first research unit of IRFA attached to a medical institution) was established at the Indian Cancer Research Centre, Bombay.

- 1948

  -Dr. C.G. Pandit was appointed as the first full time secretary of IRFA in July 1948.

- 1949

  -IRFA was re designated as the Indian Council of Medical Research (with Dr. C.G. Pandit as its first Director).

The ICMR is funded by the Government of India through the Department of Health Research, Ministry of Health & Family Welfare.

The Council's research priorities coincide with the National health priorities such as control and management of communicable diseases, fertility control, maternal and child health, control of nutritional disorders, developing alternative strategies for health care delivery, containment within safety limits of environmental and occupational health problems; research on major non-communicable diseases like cancer, cardiovascular diseases, blindness, diabetes and other metabolic and haematological disorders; mental health research and drug research (including traditional remedies). All these efforts are undertaken with a view to reduce the total burden of disease and to promote health and well-being of the population.

The Governing Body of the Council is presided over by the Union Health Minister. It is assisted in scientific and technical matters by a Scientific Advisory Board comprising eminent experts in different biomedical disciplines. The Board, in its turn, is assisted by a series of Scientific Advisory Groups, Scientific Advisory Committees, Expert Groups, Task Forces, Steering Committees **etc.**which evaluate and monitor different research activities of the Council .

The Council promotes biomedical research in the country through intramural as well as extramural research. Over the decades, the base of extramural research and also its strategies have been expanded by the Council.

Publication: Indian Journal of Medical Research,Bulletins, library bulletins, Annual Reports,Other Research Reports,Health Research Policy,Directories, Guidelinesetc.

9 Indian Council of Social Science Research:

http://icssr.org/

Indian Council of Social Science Research (ICSSR) was established in the year of 1969 by the Government of India to promote research in social sciences in the country. The Council was meant to:

- Review the progress of social science research and give advice to its users;· Sponsor social science research programmes and projects and administer grants to institutions and individuals for research in social sciences;
- Institute and administer scholarships and fellowships for research in social sciences;
- Indicate areas in which social science research is to be promoted and adopt special measures for development of research in neglected or new areas;
- Give financial support to institutions, associations, and journals engaged in social science research;
- Arrange for technical training in research methodology and to provide guidance for research;
- Co-ordinate research activities and encourage progrmmes for interdisciplinary research;

- Develop and support centers for documentation services and supply of data;
- Organize, sponsor, and finance seminars, workshops and study groups;
- Undertake publication and assist publication of journals and books in social sciences;

Advise the Government of India on all matters pertaining to social science research as may be referred to it from time to time; and take such measures generally as may be necessary from time to time to promote social science research and its utilization. At present ICSSR is supporting 30 Research Institutes in India.**Publication**: National Fellowship Guidelines, Senior Fellowship Guideline,Post Doctoral Fellowship Guideline, Doctoral Fellowship Guideline, Newsletters, monographs, research reports,ICSSR Journal of Abstracts and Reviews. **Library:** ICSSR Library has a reach collection of books journals, reports, research reports etc.Library provides OPAC service.

10 Indian Council of Historical Research

http://ichr.ac.in

Indian Council of Historical Research is an autonomous organization which was established under Societies Registration Act (Act XXI of 1860) in 1972. The objectives of the Indian Council of Historical Research (hereafter referred to as the 'Council') as laid down in the Memorandum of Association are as follows:

- "to bring historians together and provide a forum for exchange of views between them;
- to give a national direction to an objective and scientific writing of history and to have rational presentation and interpretation of history;
- to promote, accelerate and coordinate research in history with special emphasis on areas which have not received adequate attention so far;
- to promote and coordinated a balanced distribution of research effort over different areas;
- to elicit support and recognition for historical research from all concerned and ensure the necessary dissemination and use of results."

In pursuance of these objectives (a) the Council provides fellowships and financial assistance to the young teachers in colleges, universities and registered research organizations, as well as to senior scholars who might need financial support, (b) brings historians together by providing financial assistance for holding symposia, seminars, workshops, etc for exchanging views related to history, (c) provides publication subsidy to the seminars, congress proceedings and journals so that these publications may reach to researchers and scholars, (d) publishes a biannual Journal - the Indian Historical Review, and another journal Itihas in Hindi, (e) maintains a large and expanding Library-cum-Documentation Centre exclusively for researchers and scholars, (f) maintains two regional centres namely ICHR North-East Regional Centre (Guwahati) and ICHR Southern Regional Centre (Bangalore), which provide assistance to researchers / scholars, (g) and takes such other measures as the Council considers appropriate in order to implement the stated objectives of the Indian Council of Historical Research.

Publication: The Indian Historical Review,Itihas Manuscript subscription guidelines, 80 books,

Library: Library has 70000 printed materialswhich include books, reports, conference proceedings, theses and dissertations.

11 Indian Council of Philosophical Research

http://www.icpr.in

The Indian Council of Philosophical Research set up by the Ministry of Education, Government of India, was registered as a society in March 1977 under the Societies Act, 1860, but it actually started functioning in July 1981 under the Chairmanship of Professor D.P. Chattopadhyaya.

The Council has been set up by the Government of India to achieve the following aims and objectives:

- To review the progress of research in Philosophy from time to time;
- To sponsor or assist projects or programmes of research in Philosophy;
- To give financial support to institutions and organizations engaged in the conduct of research in Philosophy;
- To provide technical assistance or guidance for the formulation of research projects and programmes in Philosophy, by individuals or institutions, and/or organize and support institutional or other arrangements for training in research methodology;
- To indicate periodically areas in and topics on which research in Philosophy should be promoted and to adopt special measures for the development of research in neglected or developing areas in Philosophy;

- To co-ordinate research activities in Philosophy and to encourage programme of inter-disciplinary research;
- To organize, sponsor and assist seminars, special courses, study circles, working groups/parties, and conferences for promoting research in Philosophy, and to establish institutes for the same purpose;
- To give grants for publication of digests, journals, periodicals and scholarly works devoted to research in Philosophy and also to undertake their publication;
- To institute and administer fellowships, scholarships and awards for research in Philosophy by students, teachers and others;
- To develop and support documentation services, including maintenance and supply of data, preparation of an inventory of current research in Philosophy and compilation of a national register of philosophers;
- To promote collaboration in research between Indian philosophers and philosophical institutions and those from other countries;
- To take special steps to develop a group of talented young philosophers and to encourage research by young philosophers working in universities and other institutions;
- To advise the Government of India on all such matters pertaining to teaching and research in philosophy as may be referred to it by the Government of India from time to time;
- To enter into collaboration on mutually agreed terms, with other institutions, organizations and agencies for the promotion of research in Philosophy;
- To promote teaching and research in Philosophy;
- Generally to take all such measures as may be found necessary from time to time to promote research in Philosophy; and
- To create academic, administrative, technical, ministerial and other posts in the Council and to make appointments, thereto in accordance with the provisions of the Rules and Regulations

Library: Library is having 30000 books,107 journals.

Publications: ICPR journal,books,monograph,projectreport,survey reports

12 Central Council for Research in Homeopathy:

http://ccrhindia.org/

The Central Council for Research in Homoeopathy (CCRH) was formally constituted on 30th March 1978 , as an autonomous organization and was registered under the Societies Registration Act XXI of 1860. It was, however, in January 1979, that the Council started functioning as an independent organization.

Activities:

- To formulate the aims and patterns of research on scientific lines in Homoeopathy
- To initiate, develop, undertake and coordinate scientific research in fundamental and applied aspects of Homoeopathy
- To exchange information with other institutions, associations and societies interested in the objectives similar to those of the Council
- To collaborate research studies with other Institutes of Excellence towards promotion of Homoeopathy
- To propagate research findings through monographs, journals, newsletters, I.E.&C. materials, seminars/workshops and develop audio-visual aids for dissemination of information to the profession and public

The council is supporting and providing guidelines to 29 units in different states and union territories in India.

Publication: Guidelines, manuals, books, monographs, Indian Journal of Research in Homeopathy, newsletter, Annual Report.

Library: Library holds 10923 books, subscribing 16 journals,2165 back volumes.

13. National council of teachers Education

http://ncte-india.org/

The National Council for Teacher Education, in its previous status since 1973, was an advisory body for the Central and State Governments on all matters pertaining to teacher education, with its Secretariat in the Department of Teacher Education of the National Council of Educational Research and Training (NCERT). Despite its commendable work in the academic fields, it could not perform essential regulatory functions, to ensure maintenance of standards in teacher education and preventing proliferation of substandard teacher education institutions. The National Policy on Education (NPE), 1986 and the Programme of Action thereunder, envisaged a National

Council for Teacher Education with statutory status and necessary resources as a first step for overhauling the system of teacher education. The National Council for Teacher Education as a statutory body came into existence in pursuance of the National Council for Teacher Education Act, 1993 (No. 73 of 1993) on the 17th August,1995.

Activities : Activities of the Council is described below:

- Undertake surveys and studies relating to various aspects of teacher education and publish the result thereof;
- Make recommendations to the Central and State Government, Universities, University Grants Commission and recognised institutions in the matter of preparation of suitable plans and programmes in the field of teacher education;
- Co-ordinate and monitor teacher education and its development in the country;
- Lay down guidelines in respect of minimum qualifications for a person to be employed as a teacher in schools or in recognised institutions;
- Lay down norms for any specified category of courses or trainings in teacher education, including the minimum eligibility criteria for admission thereof, and the method of selection of candidates, duration of the course, course contents and mode of curriculum;
- Lay down guidelines for compliance by recognised institutions, for starting new courses or training, and for providing physical and instructional facilities, staffing pattern and staff qualification;
- Lay down standards in respect of examinations leading to teacher education qualifications, criteria for admission to such examinations and schemes of courses or training;
- Lay down guidelines regarding tuition fees and other fees chargeable by recognised institutions;
- Promote and conduct innovation and research in various areas of teacher education and disseminate the results thereof;
- Examine and review periodically the implementation of the norms, guidelines and standards laid down by the Council, and to suitably advise the recognised institution;
- Evolve suitable performance appraisal system, norms and mechanism for enforcing accountability on recognised institutions;
- Formulate schemes for various levels of teacher education and identify

recognised institutions and set up new institutions for teacher development programmes;

- Take all necessary steps to prevent commercialisation of teacher education; and
- Perform such other functions as may be entrusted to it by the Central Government.

NCTE has four regional bodies.

Publication : Indian Journal of Teacher Education, Journal of Teacher Support, Anweshika (Hindi Journal), Regulations, Curriculum framework, brochure, books, statute, etc.

14.4 Few Facts on Research and Development in India

In this unit activities of few Research Promoting Agencies are described in 14.3

Following table shows the Councils which are at present active and promoting research and development in India:

Serial No	Name	Year	Web Address/email
1	Central council for Research in Siddha		http://www.siddhacouncil.com/
2	National Council for Science and Technology Communication (NCSTC)		http://www.dst.gov.in
3.	National Innovation Council (NInC)		http://innovationcouncilarchive.nic.in/
4	Indian Council of Agricultural Research(ICAR)		http://www.icar.org.in/

Serial No	Name	Year	Web Address/email
5	Council of Scientific and Industrial Research (CSIR)	1942	http://www.csir.res.in/
6	Indian Nursing Council(INC)	1947	http://www.indiannursingcouncil.org/
7	Dental council of India(DCI)	1948	http://www.dciindia.org.in/
8	Indian Council of Medical Research(ICMR)	1949	http://www.icmr.nic.in/
9	University Grants Commission(UGC)	1956	http://www.ugc.ac.in/
10	Medical Council of India(MCI)	1956	http://www.mciindia.org/
11	National Council of Educational Research and Training(NCERT)	1961	http://www.ncert.nic.in/
12	Bar Council of India(BCI)	1961	http://www.barcouncilofindia.org/
13	Indian Council of Social Science Research (ICSSR)	1969	www.icssr.org
14	Central Council of Indian Medicine (CCIM)	1970	https://ccimindia.org/

Serial No	Name	Year	Web Address/email
15	Indian Council of Historical Research (ICHR)	1972	http://ichr.ac.in
16	Council of Architecture(CoA)	1972	https://coa.gov.in/
17	Science and Engineering Research Council (SERC)	1974	http://www.serb.gov.in/home.php
18	Indian Council of Philosophical Research (ICPR)	1977	www.icpr.in
19	Central Council for Research in Homoeopathy (CCRH)	1978	http://ccrhindia.org/
20	Veterinary Council of India(VCI)	1984	http://vci.nic.in
21	Distance Education Council (DEC)/ Distance Education Bureau	1985	http://www.ugc.ac.in/deb/index.html
22	Rehabilitation Council of India(RCI)	1992	http://www.rehabcouncil.nic.in/
23	National Council for Teacher Education	1993	http://ncte-india.org/
24	National Council of Rural Institutes (NCRI)	1995	www.ncri.in

14.5 Conclusion

Based on the recommendations of the National Knowledge Commission (2005) and the Committee on Renovation and Rejuvenation of Higher Education (2009), steps were initiated during the Eleventh Plan to create a new legislative framework and provide a new governance structure for higher education in the country. For this purpose several new laws are currently under consideration. These include (*i*) The Prohibition of Unfair

Practices in Technical Educational Institutions, Medical Educational Institutions and Universities Bill aimed at checking unfair practices relating to capitation fees and misleading advertising through mandatory disclosures by academic institutions;(*ii*) The National Accreditation Regulatory Authority for Higher Educational Institutions Bill that seeks to make accreditation by independent accreditation agencies mandatory for all higher educational institutions; (*iii*) The Education Tribunals Bill to create a Central tribunal and State-level tribunals for expeditious resolution of disputes relating to institutions,

faculty, students and regulatory authorities; (*iv*) Foreign Educational Institutions (Regulation of Entry and Operations) Bill to enable quality foreign education institutions to enter and operate in India and regulate operations of foreign education providers;

(*v*) National Commission for Higher Education and Research (NCHER) Bill to create an umbrella regulatory authority subsuming the UGC, and current regulators, AICTE, NCTE and DEC; and (*vi*) The National Academic Depository Bill, 2011, to create a repository of all academic credentials in the country.

These new laws together reflect the Government's focus on quality, accountability, access, and inclusion and on preparing the country's higher education system for a more competitive globalising world. These reforms would enable and facilitate innovative and high-quality institutions to grow, while making it difficult for poor-quality institutions to operate. In the next few years, a new governance structure at the national-level consisting primarily of the NCHER, National- and State-level Tribunals and the National Authority for Accreditation would be in place.

In the meantime, the UGC and other regulatory agencies have an opportunity to revitalise themselves to ensure a smooth transition to the NCHER. In this context, a review of internal processes and staff capabilities is essential and agencies should draw up year-wise transformative action plans. In addition, the UGC could immediately implement a number of innovative financing schemes that could impact the state of higher-education significantly. For example, (*i*) the UGC could shift from its current scheme-based approach to more effective programmatic interventions including norm-based financing of institutions; (*ii*) it could consider a move from historically determined detailed operational budgets to formula-based funding for general operations; (*iii*) it could start strategic funding of innovative programmes to promote certain activities/ changes/investments based on institutional proposals evaluated selectively and competitively; (*iv*) finally, the UGC or some other Central agency could further play a leading role in longitudinal profiling of students as they transition through the higher educational cycle into the workplace and could also play a role in institutional benchmarking on a longitudinal basis.

The structure of governance of higher education and their legislative framework varies widely across the States. All States will be encouraged to undertake a review of their current legislative and governance arrangements with a view to preparing themselves for the unique challenges they face in higher education.

It would be desirable for each State (except small States) to set up a State Council for Higher Education to lead the planned and coordinated development of higher education in the State and to foster sharing of resources between universities, benefit from synergy across institutions, lead academic and governance reforms at the institution level, maintain databanks on higher education and conduct research and evaluation studies. In small States, the main affiliating university can perform this role. Private universities and colleges form a bulk of higher education in several States. States could also establish independent agencies to regulate private HEIs. Institutional Level Governance Academic institutions primarily rely on individual initiative and creativity to develop their unique institutional culture and tradition over a long period of time. Principles of academic freedom, shared governance, meritocratic selection, promotion of diversity and institutional accountability are defining features of a well-governed academic institution. Moreover, the oversight, governance and management of HEIs should be closely tied to their mission. For this the current practice of treating all institutions alike will need to be abandoned. There is a need to move away from enforcing standardisation of education and processes to allow for diversity in institutional types, missions, resources and privileges. This would require a categorisation of institutions of higher educa

Ministry of Human Resource Development in its draft national education policy, 2016 has given following policy framework for the development of education and under chapter 4, S no.4.20 under the heading research, innovation and new knowledge:

Although India's overall share of research publications in the world has risen in the past decade, the quality of research has not made a significant mark. Barring a few pockets of excellence, the system is marked by mediocrity. Research minded students and faculty prefer to go abroad as they do not find the research climate in our institutions conducive. Favourable conditions need to be created in the country to promote high quality research. The country needs to develop an enabling condition for research and innovations by creating an administrative and academic environment complementing higher education. In the context of India's emergence as a soft power, there is a need to promote generation of new domains of learning required for a knowledge society. The following policy initiatives will be taken:

1. Over the next decade, at least 100 new centres/ departments of excellence, in the field of higher education, both in the public and the private sector, will be established to promote excellence in research and encourage innovations. Private trusts, philanthropists and foundations will be given freedom to establish such Centres of Excellence.

2. A clear reorientation of research agenda of National University of Educational Planning and Administration (NUEPA) will be undertaken to reflect actual issues on the ground.

3. Steps will be taken to promote generation of new knowledge and their applications and introduction of these new domains into the curricula of higher education to consolidate and strengthen India's position as a soft power.

4. In order to promote innovation, creativity and entrepreneurship, 100 more incubation centres will be established in HEIs over a period of next 5 years.

International collaborations and networks will be promoted for developing human resources required to sustain new knowledge with special focus on inter-disciplinary (Ministry of Human Resource Development, Government of India, 2016) research and studies. (National Knowledge Commission, 2009)

One of the ways to measure the rigorousness of research and innovation in a country is by looking at its patent filings. The Indian government has also been indicating that it is keen to see an increase in domestic patent filings across all industry sectors and technology areas.

The World Intellectual Property Organisation's (WIPO) IP Indicators report for 2013 revealed that of the 43,663 patents applications made in India in 2012-

2013, only 22% were filed by domestic entities. "Domestic filings need to be increased by encouraging research and development in India.

14.6 Summary

In this unit activities as well as their publications such as guidelines, statues, brochures, curriculum structure, books, journals, newsletter etc. and library resources of Research Promoting Agencies their role in various policy making and providing advice to the Government of India has been discussed for developing awareness of the research scholars. Learners can visit the websites mentioned and enhance their ideas about these organisations. They can go through their readily available publications also.

14.7 Questions /Self Assessment Questions

1. What are roles of Research Promotional Agencies in India?
2. Name the councils providing advice to research institutes in India and describe its activities.
3. Descibe the role of ICMR,ICSSR NCTE.
4. Describe the present status of research in India.
5. Describe the history of research in India

14.8 References

Ministry of Human Resource Development, Government of India. (2016). *Some inputs for draft National Education Policy ,2016.* MHRD.

National Knowledge Commission. (2009). *National Knowledge Commission : report of the nation 2006-2009.* New Delhi: National Knowledge Commission.

Planning Commission,Government of India. (2013). *Twelfth five year plan,2012-2017:social sector.* New Delhi: Sage.

মানুষের জ্ঞান ও ভাবকে বইয়ের মধ্যে সঞ্চিত করিবার যে একটা প্রচুর সুবিধা আছে, সে কথা কেহই অস্বীকার করিতে পারে না। কিন্তু সেই সুবিধার দ্বারা মনের স্বাভাবিক শক্তিকে একেবারে আচ্ছন্ন করিয়া ফেলিলে বুদ্ধিকে বাবু করিয়া তোলা হয়।

—রবীন্দ্রনাথ ঠাকুর

ভারতের একটা mission আছে, একটা গৌরবময় ভবিষ্যৎ আছে; সেই ভবিষ্যৎ ভারতের উত্তরাধিকারী আমরাই। নতুন ভারতের মুক্তির ইতিহাস আমরাই রচনা করছি এবং করব। এই বিশ্বাস আছে বলেই আমরা সব দুঃখ কষ্ট সহ্য করতে পারি, অন্ধকারময় বর্তমানকে অগ্রাহ্য করতে পারি, বাস্তবের নিষ্ঠুর সত্যগুলি আদর্শের কঠিন আঘাতে ধূলিসাৎ করতে পারি।

—সুভাষচন্দ্র বসু

Any system of education which ignores Indian conditions, requirements, history and sociology is too unscientific to commend itself to any rational support.

—Subhas Chandra Bose

Price : Rs. 150.00

(Not for sale to the students of NSOU)

Published by : Netaji Subhas Open University, DD-26, Sector-I, Salt Lake, Kolkata-700 064
and Printed at SEVA MUDRAN, 43, Kailash Bose Street, Kolkata-700 006