M.COM.

SEMESTER - II (CBCS)

RESEARCH METHODOLOGY FOR BUSINESS

SUBJECT CODE : 66701

Prof. Suhas Pednekar
Vice-Chancellor,
University of Mumbai

Prof. Ravindra D. Kulkarni
Pro Vice-Chancellor,
University of Mumbai

Prof. Prakash Mahanwar
Director,
IDOL, University of Mumbai

Programme Co-ordinator : **Prof. Rajashri Pandit**
Asst. Prof. in Economic,
Incharge Head Faculty of Commerce,
IDOL, University of Mumbai, Mumbai

Course Co-ordinator : **Mr. Sambhaji Shivaji Shinde**
Assistant Professor, IDOL,
University of Mumbai, Mumbai

Course Writer : **Dr. Vinod Kamble**
Assistant Professor,
Bal Bharti's MJP College of Commerce,
Kandivali (W), Mumbai

: **Dr. Vinit Joshi**
Assistant Professor,
SIES College of Arts, Commerce & Science,
Nerul, Navi Mumbai - 400706

: **Dr. Bhalchandra Karbhari**
Assistant Professor,
Akbar Peerbhoy College of Commerce & Economics,
Mumbai - 400008

: **Dr. Megha K. Juvekar**
Assistant Professor,
Nirmala Memorial Foundation College of Commerce and Science,
Mumbai - 400101

May 2021, Print - 1

Published by : Director,
Institute of Distance and Open Learning ,
University of Mumbai,
Vidyanagari, Mumbai - 400 098.

DTP Composed : Ashwini Arts
Vile Parle (E), Mumbai - 400 099.

Printed by :

CONTENTS

SYLLABUS
M.COM
SEMESTER - II
RESEARCH METHODOLOGY FOR BUSINESS

Sr. No.	Modules / Units
1.	**Introduction Research**
	• Features and Importance of research in business, Objectives and Types of research- Basic, Applied, Descriptive, Analytical and Empirical Research. • Formulation of research problem, Research Design, significance of Review of Literature • Hypothesis: Formulation, Sources, Importance and Types • Sampling: Significance, Methods, Factors determining sample size
2.	**Research Process**
	• Stages in Research process • Data Collection: Primary data: Observation, Experimentation, Interview, Schedules, Survey, • Limitations of Primary data • Secondary data: Sources and Limitations, • Factors affecting the choice of method of data collection. Questionnaire: Types, Steps in Questionnaire Designing, Essentials of a good questionnaire
3.	**Data Processing and Statistical Analysis**

	• Data Processing: Significance in Research, Stages in Data Processing: Editing, Coding, Classification, Tabulation, Graphic Presentation
	• Statistical Analysis: Tools and Techniques, Measures of Central Tendency, Measures of Dispersion, Correlation Analysis and Regression Analysis.
	• Testing of Hypotheses –
	  • Parametric Test-t test, f test, z test
	  • Non-Parametric Test -Chi square test, ANOVA, Factor Analysis
	• Interpretation of data: significance and Precautions in data interpretation

4.	**Research Reporting and Modern Practices in Research**

	• Research Report Writing: Importance, Essentials, Structure/ layout, Types
	• References and Citation Methods:
	  • APA (American Psychological Association)
	  • CMS (Chicago Manual Style)
	  • MLA (Modern Language Association)
	• Footnotes and Bibliography
	• Modern Practices: Ethical Norms in Research, Plagiarism, Role of Computers in Research

Reference Books

- Research Methodology – Text and Cases with SPSS Applications, by Dr S.L. Gupta and Hitesh Gupta, International Book House Pvt Ltd
- Business Research Methodology by T N Srivastava and Shailaja Rego, Tata Mcgraw Hill Education Private Limited, New Delhi
- Methodology of Research in Social Sciences, by O.R. Krishnaswami, Himalaya Publishing House
- Research Methodology by Dr Vijay Upagude and Dr Arvind Shende
- Business Statistics by Dr S. K Khandelwal, International Book House Pvt Ltd
- Quantitative Techniques by Dr S. K Khandelwal, International Book House Pvt Ltd
- SPSS 17.0 for Researchers by Dr S.L Gupta and Hitesh Gupta, 2nd edition, Dr S. K Khandelwal, International Book House Pvt Ltd
- Foundations of Social Research and Econometrics Techniques by S.C. Srivastava, Himalaya publishing House
- Statistical Analysis with Business and Economics Applications, Hold Rinehart & Wrintston, 2nd Edition, New York
- Business Research Methods, Clover, Vernon T and Balsely, Howard L, Colombus O. Grid, Inc
- Business Research Methods, Emary C.Willima, Richard D. Irwin In. Homewood

- Research Methods in Economics and Business by R. Gerber and P.J. Verdoom, The Macmillan Company, New York

- Research and Methodology in Accounting and Financial Management, J.K Courtis

- Statistics for Management and Economics, by Menden Hall and Veracity, Reinmuth J.E

- Panneerselvam, R., Research Methodology, Prentice Hall of India, New Delhi, 2004.

- Kothari CR, Research Methodology- Methods and Techniques, New Wiley Ltd., 2009

Scheme of Examination:

The performance of the learners will be evaluated in two components. One component will be the Internal Assessment component carrying 40% marks and the second component will be the Semester End Examination component carrying 60% marks.

Internal Assessment:

The Internal Assessment will consist of one class test of 40 marks for each course excluding projects. The question paper pattern will be shown as below:

Question Paper Pattern
(Internal Assessment)

Maximum Marks: 40 marks
Questions to be set: 03
Duration: 1 hours

Question No.	Particular	Marks
Q - 1	Objective Questions Students to answer 10 sub questions out of 15 sub questions. *(*Multiple choice/ True or False/ Match the columns/ Fill in the blanks)* ***OR*** Objective Questions A) Sub Questions to be asked 08 and to be answered any 05 *B)* Sub Questions to be asked 08 and to be answered any 05 *(*Multiple choice/ True or False/ Match the columns/ Fill in the blanks)*	10 Marks
Q - 2	Concept based short questions Students to answer 5 sub questions out of 8 sub questions.	10 Marks
Q - 3	Practical problems or short questions Students to answer 02 sub questions out of 03 sub questions	20 Marks

Question Paper Pattern
(Theoretical Courses)

Maximum Marks: 60
Questions to be set: 04
Duration: 2 hours
All Questions are Compulsory Carrying 15 Marks each.

Question No.	**Particular**	**Marks**
Q - 1	Full length Question OR	15 Marks
Q - 1	Full length Question	15 Marks
Q - 2	Full length Question OR	15 Marks
	Full length Question	15 Marks
Q - 3	Full length Question OR	15 Marks
	Full length Question	15 Marks
Q - 4	Objective Question (Multiple Choice/ True or False/ Fill in the Blanks/ Match the Columns/ Short Questions.) OR	15 Marks
	Short Notes (Any three out off Five)	15 Marks

Note :

Full length question of 15 marks may be divided into two sub questions of 08 and 07 marks.

Sr.	Particular
01	**Standard of Passing** The learner to pass a course shall have to obtain a minimum of 40% marks in aggregate for each course where the course consists of Internal Assessment & Semester End Examination. The learner shall obtain minimum of 40% marks (i.e. 16 out of 40) in the Internal Assessment and 40% marks in Semester End Examination (i.e. 24 out of 60) separately, to pass the course and minimum of Grade E in the project component, wherever applicable to pass a particular semester. A learner will be said to have passed the course if the learner passes the Internal Assessment & Semester End Examination together.
02	**Allowed to Keep Terms (ATKT)** 1) A learner shall be allowed to keep term for Semester II irrespective of number of courses of failure in the semester I. 2) A learner shall be allowed to keep term for Semester III if he/she passes each of the semester I and Semester II OR a learner fails in not more than two courses of Semester I and not more than two courses of Semester II.

MODULE – I
INTRODUCTION TO RESEARCH

1

INTRODUCTION TO RESEARCH

Unit Structure

1.0 OBJECTIVE

After studying this unit the student will be able to -

- Understand the concept of Research
- To understand the characteristics of Research
- Know the Importance of Research in Business
- Explain various Types of Research

1.1 INTRODUCTION

1.1.1 MEANING

The word 'Research' is derived from the Middle French word **'recherche'** meaning **'to go about seeking**'. Research is a careful and detailed study into a specific problem, concern, or issue using the scientific method. Also research is a systematic investigation to search for new facts in any branch of knowledge. It helps to find solutions to certain problems and arrive at new conclusions.

1.1.2 DEFINITION

According to The Organization for Economic Cooperation and Development (OECD), "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."

According to John W. Creswell, who states that "research is a process of steps used to collect and analyze information to increase our

understanding of a topic or issue". It consists of three steps: pose a question, collect data to answer the question, and present an answer to the question

William C. Emory defines “research is any organized inquiry designed and carried out to provide information for solving a problem.”
Robert Ross defines “Research is essentially an investigation, a recording and analysis of evidence for the purpose of gaining knowledge.”

1.1.3 FEATURES OF RESEARCH

1) **Systematic Process:** Research is a systematic process. No research can be conducted haphazardly. Each step must follow other. There are set of procedures that have been tested over a period of time and are thus suitable to use in research: The steps are as follows:
   - Formulating the research problem
   - Review of Literature
   - Define Research objectives
   - Preparing Research Design
   - Collection and analysis of data
   - Interpretation of data
   - Preparation of report
   - Follow-up of report

2) **Objective and Logical / Empirical:** Researcher needs to make every possible effort to avoid biasness in every stage of research process. In order to make research objective and logical / empirical, there is a need to collect relevant and accurate data to investigate into the research problem. After collection of data, the researcher need to undertake processing of data, analyse and interpret it and arrive at logical conclusion. So the research has to be conducted following rigorous scientific methods and procedures. Each step in the research has been tested for accuracy.

3) **Development of Principles and Theories:** A systematic research helps to develop new principles and theories. Such principles and theories can be useful to several organizations to manage and deal with people and things in a better way. Eg. Prof. Alfred Marshall used the inductive method of research in economics. On the basis of the market analysis he framed ‘Law of Demand’. According to this law, there exists a negative relationship between the price and quantity demanded. When price increases, demand falls and vice versa. Another example could be, ‘14 Principles of Management by Henry Fayol’. They are developed gradually with thorough research work. Systematic observation and experiments are conducted in various organizations before developing them.

4) **Multipurpose Activity:** Research is multipurpose activity. It helps to achieve multiple purposes such as:

- Discover new facts or verify old facts.
- Predict future events and control such events
- Establishes relationship between variables
- Develop new scientific tools, concepts and theories

5) **Basic and Applied Research: Basic or fundamental or pure research** is a research approach that is entirely theoretical and aimed at improving or expanding the knowledge-base of a particular field of study. The main motivation in basic research is to expand man's knowledge, not to create or invent something. There is no obvious commercial value to the discoveries that result from basic research. It does not have direct commercial objective.

   Applied research is designed to solve practical problems of the modern world, rather than acquire knowledge for knowledge sake. In other words, the purpose of applied research is to know more about a certain real-world problem and take steps to solve it. It has direct commercial objective. Researchers in this field try to find immediate solutions to existing problems facing a society or an industrial or business organization.

6) **Quantitative and Qualitative Research: Quantitative research** refers to as a systematic investigation of phenomena by gathering quantifiable data and performing statistical techniques. Eg. Research is undertaken to find out the number of unemployed graduates. This type of research is usually done by using surveys, experiments, and so on.

   Qualitative research is used to gain an understanding of human behaviour, intentions, attitudes, experience, etc. It is based on the observation and the interpretation of the people. Eg. Research is undertaken to find out reasons as to why employees remain absent from work.

7) **Generalization:** When the researcher conducts a research, he/she selects target population and from this population, small sample is selected for collecting data. So the sample selection must be done systematically so that it represents the whole population or the universe. The findings with this sample is generalized on entire population/universe of research. Eg. A research is undertaken on 'Consumer behavior towards electronic goods of Samsung Company in Mumbai region' among 500 sample size. The findings of these 500 samples may be generalized for people residing in entire Mumbai region.

8) **Reliability:** It is a subjective term which can not be measured precisely, but today there are instruments which can estimate the reliability of any research. Reliability is the extent to which the outcomes are consistent when then experiment is repeated more than once. If research is undertaken with similar population and with

similar procedures, If it yields similar results each time it is called to be a reliable research. . Eg. A research is conducted on 'the effects of single parenting on the class performance of the children' and the results conclude that single parenting causes low grades in class. These results should have to be reliable for another sample taken from a similar population. More the results are similar; more reliability is present in the research.

9) **Validity:** Validity of the research instrument can be defined as the suitability / accuracy of the research instrument to the research problem. Validity is the extent to which the instruments that are used in the experiment measure exactly what you want them to measure. Some researchers say that validity and reliability are co-related, but the validity is much more important than reliability. Without validity, research goes in the wrong direction.

1.1.4 IMPORTANCE OF RESEARCH IN BUSINESS

1) **Helps to predict changes in business environment:** The business management is witnessing constant changes due to changes in external business environment such as:
   - Consumer preferences,
   - Competitor's strategy,
   - Society expectations
   - Economic environment,
   - Technological environment,
   - Legal environment (macro factors) and so on.

This change in business environment can adversely affect a business organization. So the manager can timey predict such changes and save business from heavy losses.

2) **Launching new product:** Business research helps in successful launching of a new product in the market. This is because, a research enables to know the likes, dislikes, preferences and choices of their consumers related to product. Accordingly a business firm can design and launch new product. Such product has lower rejection rate and higher acceptance from consumers. When customers are offered product as per their preferences, it results into customer satisfaction.

3) **Helps to design effective marketing strategy:** Business research helps to design effective marketing strategy. Research enables to a business organization to:
   - Design quality product
   - Decide right price
   - Effective promotion
   - Proper distribution

4) **Achieve organizational goals:** Systematic business research helps to achieve organization goals such as:
   - Customer Satisfaction
   - Increase in sales and profits
   - Expansion of Business
   - Enhance Corporate image
   - Face competition and so on.

5) **Studying the competition:** Companies often use business research to study key competitors in their markets. The company may want to know the percentage of customers in the market who purchase its products versus competitor's products. Also it enables to know the marketing strategy of competitors. Accordingly a business firm can design its own marketing strategy to survive and growth in the highly competitive market.

6) **Facilitates decision making:** With the help of research data available, businessman can take right decision at right time. Research provides a business with a chance to update itself the latest market trends. Such knowledge will prove helpful in the formulating of useful tactics for success in the market. It is through research that a business is able to make educated and informed decisions.

7) **Helps to measure business progress:** Business research enables to gauge (measure) how well business is performing. Early research may highlight problems in services and short falls in the products. Regular market research will show if improvements are being made and if positive, will help to motivate a team.

8) **Availability of competent manpower:** Research also helps in the recruitment and selection of competent manpower. Proper recruitment and selection of employees with the right skills and attitudes help the firm to increase its productivity levels. Further effective training and compensation package can improve morale of employees and motivate them to work with dedication and commitment.

9) **Helps to get right suppliers:** Research helps the firm to get a right supplier who offers raw material at right price and right time. A proper supplier selection enables the firm to get or acquire high quality raw materials which result into production of high quality products that are consumed by end users.

10) **Improves productivity:** Productivity refers to the ratio of output to the input i.e. with one unit of input, how much output is produced. Productivity can be increased with the help of:
    - Training to employees
    - Research and Development
    - Use of Modern Technology

Business research make realize to the business firm to undertake these activities which result into improvement in productivity of the business.

1.1.5 OBJECTIVES OF RESEARCH

1) **To find out solution to problems:** Research can be undertaken to find solutions to solve a specific problem. Data is collected on the problem faced by an organization. Such data is analysed and interpretation is made to find out solution to solve the problem. Eg. An organization may initiate research to find solution to problem of declining sales of their product in the market. So the data is collected to find out reasons for declining sales and analysis of such data may provide solution to the problem.

2) **To obtain Information:** Research is undertaken to obtain information, which may not be easily available. Variety of information can be collected such as consumer preference, competitor's strategy, demand, economic conditions and so on. Such information is vital for a marketer to take crucial marketing decisions.

3) **To make future predictions:** Research enables a businessman to collect past and present data. Based on such data, the researcher can make future predictions about business situation and business stand in near future. Eg. A marketer wishes to launch a new product in the market. With the help of research he can predict the future of that product and then decide whether to come up with that product or not.

4) **To develop new tools and concepts:** Research helps to develop new tools and concepts for better study of an unknown phenomenon. For this purpose, exploratory research is undertaken to achieve new insights into such phenomenon. Eg. Research enables a business firm to know what factors affect consumer satisfaction in modern times. Accordingly marketer can develop new tools and concepts to provide maximum consumer satisfaction.

5) **To verify and test existing laws or theories:** Research may be undertaken to verify and test existing laws or theories. Such verification and testing of existing laws and theories is required to know the relevance of it in present time..

1.2 TYPES OF RESEARCH

1) **Basic Research**: Basic or fundamental or pure research is a research approach that is entirely theoretical and aimed at improving or expanding the knowledge-base of a particular field of study. The main motivation in basic research is to expand man's knowledge, not to create or invent something. There is no obvious commercial value to the discoveries that result from basic research. It does not have direct

commercial objective. Eg. “A study on socio-economic status of dwellers of Dharavi area, Mumbai” OR “A study on impact of mobile phone on studies of secondary school children.”

2) **Applied Research:** Applied research is designed to solve practical problems of the modern world, rather than acquire knowledge for knowledge sake. In other words, the purpose of applied research is to know more about a certain real-world problem and take steps to solve it. It has direct commercial objective. Researchers in this field try to find immediate solutions to existing problems facing a society or an industrial or business organization. Eg. “Investigating factors that improve worker's productivity.” OR “A study on declining sales of ABC soap.”

3) **Descriptive Research:** Descriptive research is a type of research that provides an in-depth description of the situation, phenomenon or population under study. This research is also an appropriate option when researcher does not have any information about the research problem, and primary information gathering is required to establish a hypothesis. The descriptive research provides the answer to the “what” part of a research and does not answer the questions “why”. Researcher has no control over the variables and researcher has to report the actual conditions. For example, an apparel brand that wants to understand the fashion purchasing trends among Mumbai buyers. They will conduct a demographic survey of this region, gather population data and then conduct descriptive research on this demographic segment. The study will then uncover details on “what is the purchasing pattern of Mumbai buyers,” but not cover any investigative information about “why” the patterns exits. Because for the apparel brand trying to break into this market, understanding the nature of their market is the study’s objective.

4) **Analytical Research:** Analytical research is a critical evaluation based on information that is available. The researcher makes use of facts or information already available and analyse these to make a critical evaluation of the material. It is primarily concerned with testing hypothesis. It specifies and interprets relationships by analyze the facts or existing information. The analytical research provides the answer to the “why” part of a research. It is usually concerns itself with cause-effect relationships. Eg. Explaining why and how US trade balance move in a particular way over time.

5) **Conceptual Research:** This research is a type of research that is generally related to abstract idea (existing in thought or as an idea but not having a physical or concrete existence) or concept. It does not involve any practical experiment. This research is generally used by philosophers and thinkers to develop new concepts or reinterpret existing ones.

6) **Empirical Research:** Empirical research relies on the observation and experience with hardly any regard for theory and system. Such research is data based, which often comes up with conclusions that can be verified through experiments or observation. For example: A research is being conducted to find out if listening to happy music while working may promote creativity? An experiment is conducted by using a music website survey on a set of audience who are exposed to happy music and another set who are not listening to music at all, and the subjects are then observed. The results derived from such a research will give empirical evidence if it does promote creativity or not.

1.3 SUMMARY

This unit is about features of research and how research holds significance in business. It also comprises objectives of undertaking research and different types of research. So this unit gives basic ideas about research which helps to plan the research activity and achieve its objectives. So every researcher must know the features of research, its importance and objectives as well as different types of research before he/she actually starts with a research activity.

1.4 EXERCISE

FILL IN THE BLANKS

1) _____ is any organized inquiry designed and carried out to provide information for solving a problem
   A) Sampling **B) Research** C) Hypothesis D) Research Design

2) The research should be _____
   A) Empirical B) Biased C) Subjective D) Inaccurate

3) _____ research refers to as a systematic investigation of phenomena by gathering quantifiable data and performing statistical techniques
   A) Qualitative B) Census **C) Quantitative** D) Historical

4) _____ research is an appropriate option when researcher does not have any information about the research problem
   A) Analytical **B) Descriptive** C) Conceptual D) Basic

5) Research in business facilitates _____
   A) Design ineffective marketing strategy
   B) Design various schemes for upliftment of backward regions
   C) Availability of incompetent manpower
   D) Forecast changes in business environment

TRUE OR FALSE

1. Research can be conducted haphazardly. **FALSE**
2. Research helps to measure business progress. **TRUE**
3. Applied research has direct commercial objective. **FALSE**
4. Analytical research is a critical evaluation based on information that is available. **TRUE**
5. In research, reliability is the extent to which the outcomes are consistent when then experiment is repeated more than once. **TRUE**

MATCH THE PAIRS

Group A	**Group B**
1) Basic Research	a) Achieve organizational goals
2) Empirical Research	b) Gathering quantifiable data
3) Quantitative Research	c) Risarch
4) Objective of Research in Business	d) Recherche
5) The word 'Research' is derived from French word	e) Relies on the observation and experience
	f) Also known as pure or fundamental research

(1 – f, 2 – e, 3 – b, 4 – a, 5 – d)

ANSWER IN BRIEF

1) Define Research. Explain its features
2) "Research is important in Business." Explain.
3) Discuss the objectives of research.
4) **Write a note on:**
   - Basic Research
   - Applied Research
   - Descriptive Research
   - Analytical Research
   - Empirical Research

1.5 REFERENCES

- https://en.wikipedia.org/wiki/Research#:~:text=The%20word%20research%20is%20derived,the%20term%20was%20in%201577.
- https://readingcraze.com/index.php/characteristics-research-2/
- https://www.slideshare.net/darious91/importances-of-research-in-business
- https://research-methodology.net/descriptive-research/
- https://www.marketing91.com/descriptive-research/
- https://www.ukessays.com/essays/economics/descriptive-research-vs-analytical-research-economics-essay.php
- https://www.youtube.com/watch?v=UqtckUep840

❖❖❖❖

2

PLANNING OF RESEARCH

Unit Structure

2.0 OBJECTIVES

After studying this unit the student will be able to -

- Describe formulation of Research Problem
- Understand the concept of Research design
- Discuss significance of Review of Literature
- Explain the concept of Hypothesis
- Elucidate the concept of Sampling

2.1 INTRODUCTION

Understanding the research process is an important step towards executing a thorough research or study. Let us examine the different phases in research planning as well as the stages involved in a research process.

2.2 FORMULATION OF RESEARCH PROBLEM

2.2.1 Meaning

Identification and formulation of a research problem is the first step of the research process. It is the most challenging and difficult phase of the research process.

A research problem is a question that a researcher wants to answer or a problem that a researcher wants to solve. In other words, research

problem is an issues or a concern that an investigator / researcher presents and justifies in a research study.

A research problem is the most important aspect of the research. Researcher must spend time to refine and assess the research problem before getting started with the research activities. A research question must be straightforward, to the point and focused.

Example
Incorrect Research Problem: What are the effects of social media on people?

Correct Research Problem: What effect does using Facebook everyday have on teenagers?

In the above example, the first question is not specific enough to capture accurate feedback. Nobody knows what social media researcher is talking about and what 'people' (target population) researcher is referring to.

2.2.2 Steps to Formulate of Research Problem

1) **Identify the broad research area:** The researcher begins research by identifying a broad research area based on his/her interest, specialty, profession, expertise, and knowledge. For example, a researcher studying about Business Management can select areas like Marketing Management, Human Resource Management, Organizational Management, and Financial Management. These are the broader areas that can be further subdivided into various research topics to figure out marketing strategies.

2) **Divide the broad area into sub-areas:** After the researcher chooses a broad area to study, he/she need to narrow down to a specific topic that is manageable and researchable. To do this, break down the broad area into sub-areas and choose a specific topic. For example, if your broad area is Marketing Management it can be further divided into the following subcategories:
    - Consumer Satisfaction
    - Marketing Mix
    - Consumer Relationship
    - Digital Marketing

3) **Choose a sub-area:** It is not possible to study all the sub-areas due to time and money constraints. Thus, the researcher needs to choose one sub-area of interest and one that is manageable and feasible for him/her. The area selected must have some research significance and must be significant to the researcher's research knowledge. Eg. A researcher selects a sub-area is 'Consumer Satisfaction'.

4) **Formulate research question/problem:** After a specific sub-area is chosen, researcher needs to formulate research question/problem that deems important for the research study. Many question/problem may arise but narrow down and choose the most important and impact ful question/problem. Eg. "A study on Consumer Satisfaction from After-Sales-Services provided by Samsung Company in Mumbai region"

5) **Set research objectives:** After research question/problem is framed, researcher must draw a plan about the objectives of the research that he/she need to explore. The objectives of the research study help to identify the research question/problem. There is a difference between the research question/problem and the research objective. The difference is the way they are written. Research question/problem generally consists of an interrogative tone. On the other hand, the research objectives are aim-oriented. They include terms like to examine, to investigate, to explore, and to find out. Example of Research Objectives:
   - To investigate factors affecting consumer satisfaction from after-sales-services provided by Samsung company
   - To find out various problems faced by consumers while availing after-sales-service of Samsung company

2.3 RESEARCH DESIGN

2.3.1 Meaning

After deciding the basic aspects of research project such as formulating research problem, objectives of research, data requirement, sample design etc. and before the commencement of work of research project, the researcher has to prepare research design. Decisions relating to what, where, when, how much, by what means concerning a research study constitute a research design.

Research design is a logical and systematic outline of research project prepared for directing, guiding and controlling a research work. It means to prepare detailed plan and procedure for the conduct of the research project. It acts as a broad outline of the research work and acts as a master plan / blue print for the conduct of formal investigation. It is the basic plan that guides researcher in the execution of research project undertaken.

2.3.2 Elements of Research Design

1) Nature of the research and Objectives of study
2) Time period of research study
3) Universe and sample size of respondents
4) The location where the study would be conducted
5) The resources required to conduct the research
6) Type and source of research data required
7) Techniques of data collection and analysis

2.3.3 Need and Importance of Research Design

1) **Provides Guidelines:** Research design provides guidelines to the researcher in respect of:
    - Time period of research work
    - Type and source of data to be collected
    - Technique of data collection and analysis
    - Sampling design

    Therefore, the researcher will be able to collect right data from the right source at the right time. It also helps him to complete the research within stipulated time.

2) **Organizing Resources:** Research design enables organizing of resources required to carry on research activity. The resources required are:
    - Funds
    - Equipment / instruments and materials
    - Manpower

    Availability of right amount of resources facilitates smooth conduct of research activity. Research activity would be difficult to carry on without availability of proper amount of resources.

3) **Selection of Techniques of data collection and analysis:** Research design helps to select appropriate technique for **data collection** such as
    - Survey or Interview
    - Observation
    - Experimentation
    - Internet
    - Library etc.

Research design also helps to select appropriate technique for **data analysis** such as :

- Measures of Central Tendency (Mean, Median, Mode)
- Time Series (Simple Moving Average, Weighted Moving Average)
- Correlation Techniques etc.

4) **Collection of Relevant Data:** The research design indicates :
    - Area of research
    - Universe/Population of research
    - Sample Size etc.

    Accordingly researcher can select right area of research and target audience. He can decide his universe/population from which samples are selected to collect the relevant data.

5) **Objectives of Research:** Research design specifies objectives of research. Research design provides right direction to the researcher to carry on with research activity. This in turn will help to attend the research objectives.

6) **Monitoring of Expenditure:** Research design includes allocation of research budget for various research activities. There is proper control over expenditure. Wastage of funds does not happen. The research is successfully conducted with allocated funds.

7) **Execution of Research Work:** The research design indicates the start time and completion time of research activity. Therefore, there is timely execution of research work. If research design does not indicate the time frame, there is every possibility of research work getting delayed and objectives may not be achieved.

8) **Motivation to Research staff:** A systematic research design motivates the staff to collect the right data from the right source. Also due to timely completion of research activity, the research staff may be rewarded with monetary and non-monetary incentives.

9) **Improvement in Decision-Making:** Research design enables researcher to collect right data from right source. Right data enables to take right decision. Wrong data may lead to misleading decision making.

2.4 REVIEW OF LITERATURE

2.4.1 Meaning

The researcher must consult the available publications such as books, journals, magazines, research reports and similar other publications before starting his/her own research activity. ROL refers to extensive review of literature relating to research problem which researcher intends to undertake. Such ROL provides good insides into research problems and get familiar with previous research studies undertaken.

2.4.2 Significance of Review of Literature (ROL)

1) **Get background knowledge of research problem:** A research problem is a question that a researcher wants to answer or a problem that a researcher wants to solve. ROL helps researcher to get background knowledge of Research Problem.

2) **Helps to identify gaps in research:** Research gap refers to the areas which are not explored in the past researches. ROL enables the researcher to identify the gap in research, conflicts in previous studies, open questions left from other research. The researcher can make an attempt to fill this gap by undertaken research activity.

3) **Help to formulate research hypotheses:** Hypothesis is an assumption made to explain certain fact or provide basis for further investigation. It is tentative in nature and it may prove to be correct or incorrect. Past studies or ROL helps researchers to frame hypothesis for his/her current studies. The researcher collects data which may prove or disprove the hypothesis. Based on the result of hypothesis testing a conclusion can be drawn.

4) **Get familiar with methodology adopted by other researchers:** Research methodology is the specific procedures or techniques used to identify, select, process, and analyze information/data about a research problem. ROL enables researcher to get familiar with methodology that is used by other researchers in their researches. Accordingly he/she can decide his/her own methodology in terms or target population, sample size, method and technique of data collection and analysis and so on.

5) **Prepare research design:** Research design is a logical and systematic outline of research project prepared for directing, guiding and controlling a research work. With the help of ROL, a researcher can prepare his/her research design. Research design keeps research work on right track and help to complete research in time.

6) **Prepare sample design:** A sample design is the framework, or road map, that serves as the basis for the selection of a survey sample. In a research, it is not possible to collect data from entire population/universe due to constraints of time, money and energy on the part of researcher. So researcher needs to select samples from population/universe of research. Sample selected must be true representative of the population/universe of research. ROL helps researcher to prepare proper sample design.

7) **Get familiar with data collection sources and data analysis techniques:** For the purpose of researcher primary or secondary or both data can be used by a researcher. The primary data can be collected by observation, survey or experiment method. Secondary data can be collected from library, internet, reports etc. Collected data needs to be analysed to draw conclusion out of it. Various statistical tools can be used such as correlation, measures of central tendency and so on for the purpose for data analysis. ROL facilitates to get familiar with data collection sources and data analysis techniques used by other researchers.

8) **Understand findings of other researchers and their conclusions:** ROL helps researcher to understand findings of other researchers and their conclusion. It can be basis for researcher's own further research activity.

9) **To compile bibliography:** Bibliography is a list of sources used in our research. The main purpose of a bibliography entry is to give credit to authors whose work researcher has consulted in his/her research. ROL helps researcher to refer bibliography of others researchers to find out more about the topic by exploring into their research.

10) **Understand the structure of research report:** Research report is a written document containing key aspects of research project. After the research work is completed, the findings along with recommendations are presented in the form of research report to the authority for the purpose of further decision making. So ROL enables researcher to understand the structure of research report.

2.5 HYPOTHESIS

2.5.1 Meaning

Hypothesis is an assumption made by the researcher to explain certain fact or provide basis for further investigation. It states what the researcher thinks the outcome of the study will be. The researcher makes hypothesis and collects data that either support the hypothesis or do not support it. So the hypothesis may be proved to be correct or incorrect. Hypotheses are essential to all research studies with the possible exception of some descriptive studies whose purpose is to answer certain specific questions.

Example - A manager may hypothesize that salespersons who show the highest job satisfaction will be the most productive salespersons. Another example, organizational researcher may believe that if workers' attitudes toward an organizational climate are changed in a positive direction, there will be an increase in organizational effectiveness among these workers.

2.5.2 Definition

Webster's Dictionary defines hypothesis as "an unproved theory, proposition, supposition etc. tentatively accepted to explain certain facts or to provide a basis for further investigation, argument, etc."

2.5.3 Formulation of Hypothesis

1) **Identification of Research Problem and its causes:** The researcher must identify the research problem which needs to be investigated. Also he/she needs to identify cause of such problem. Eg. The research problem could be "**Decline in Sales of Lux soap in Mumbai Region".** The possible causes of such decline in sales could be:
    - Poor quality of the products
    - Higher price of the product
    - Ineffective promotion-mix
    - Faulty distribution network

2) **Formulate the Hypotheses:** The researcher may undertake extensive Review of Literature (ROL) or discuss with expert or by his/her own experience formulates the hypothesis. Eg. Hypotheses relating to the above problem can be formulated as:
    - Sales are declining on account of poor quality of the products
    - Sales are declining on account of higher price of the product
    - Sales are declining on account of ineffective promotion-mix
    - Sales are declining on account of Faulty distribution network

3) **Pilot Test the Hypotheses:** The researcher may conduct pilot study to test the hypothesis. Small sample respondents are selected and data is collected from them to conduct the pilot study. All the hypotheses are put to test. The pilot study may indicate the most probable cause of the problem. This may help to select the best hypothesis for the purpose of detailed investigation. Suppose the pilot study states that most probable cause of problem is poor quality of the product.

4) **Select the Best Hypothesis:** After selecting the best hypothesis on the basis of pilot study, the researcher proceeds for investigation of the problem and find out the validity of the hypothesis. The researcher may specify the null hypothesis and alternative hypothesis.

    - **Null Hypothesis:** It states that there is no relationship between two or more variables. A researcher hopes to reject or disprove the null hypothesis. Eg. There is no relationship between poor quality of product and decline in sales.

    - **Alternative Hypothesis:** It states there is relationship between two or more variables. Eg. There is relationship between poor quality of product and decline in sales.

5) **Conduct the Research:** After formulating the final hypothesis, the researcher proceeds to conduct the research. He/she may prepare research design to conduct research in right direction. The researcher collects data and analyse the same to draw conclusion. He may use T-test, Z-test, Chi-Square, ANOVA, Correlation etc. tests for the purpose for testing hypothesis.

6) **Acceptance or Rejection of Hypothesis:** After testing the hypothesis, the researcher may reject the null hypothesis or the researcher may fail to reject the null hypothesis. Generally, when the researcher rejects the null hypothesis, the researcher may accept the alternative hypothesis. At times, the alternative hypothesis may also be rejected.

2.5.4 Sources of Hypothesis

1) **Intuition or Hunch:** A person may get ideas to develop hypothesis due to one's own intuition or hunch. Ideas can strike like a flash. Eg. The story of Laws of Gravitation propounded by Newton at the sight of falling apple is the case of intuition.

2) **Past Researches:** Findings of the past researches done by others can be used for framing the hypotheses. Eg. A researcher found in the past researches that rise in rate of commission of salesman resulted in increase in sales of the company. A researcher may use this finding to formulate his research hypothesis as "Increase in rate of commission of salesman leads to increase in sales."

3) **Consultations:** The researcher can hold discussion with experts to develop hypothesis. In academic research, the research students can take help of a research guide who is expert in his/her own subject. In applied (commercial) research, the researcher may take help of marketing manager. In social research, the researcher may take help of an NGO.

4) **Observation:** Hypothesis can be developed through own observation. Eg. One can observe general pattern of buying behaviour in the market, and develop a hypothesis such as "Educated customers prefer braded items as compared to illiterate or less educated customers."

5) **Continuity of Research:** Some researches are carried on for several numbers of years. The research may be divided into different phases. At each phase the researcher may get different findings based on which he/she develops hypothesis for next phase.

6) **Culture:** While formulating a hypothesis for a problem, culture should be studied. If a researcher wants to study trends towards female education in a particular area, for this purpose he/she needs to study traditions, family system, Norms, Values, region and education system of that area.

7) **Theory:** Logical deduction from the theory lead to new hypothesis. The hypothesis must be valid, if the theory holds true. Eg. The theory on human relations in management states that effective human relations help to improve productivity. On the basis of this theory, a hypothesis can be developed that "Effective management-labour relations facilitates higher productivity."

8) **Personal Experience:** On the basis of personal experience, researcher uses his mind and suggests hypothesis. Eg. A researcher experienced poor services in the Government hospitals. He/she may develops a hypothesis "Poor quality of services results into less footfall in Government Hospitals."

2.5.5 Importance

1) **Helps to explore unknown facts:** The hypothesis provides the researcher with the most efficient instrument for exploring and explaining the unknown facts. It stimulates the researcher for further research studies.

2) **Enables to prepare research design:** The hypothesis helps in preparing research design. It may suggest research objectives, sample design, data requirement, techniques of data collection, tests and tools to analyse data etc.

3) **Identifies need for data:** A Hypothesis specifies the need of data i.e. whether research will require primary data or secondary data. Hypothesis would enable to collect required data. Without hypothesis much useless data may be collected and important data would be omitted.

4) **Identifies sources of data:** A Hypothesis also specifies the source of data i.e. survey, experiment, observation, library, reports, internet etc. Therefore, the researcher would consider only the relevant source of data, which in turn would speed up the research activity.

5) **Development of theory and principles:** Hypothesis also facilitates development of theory and principles. Eg. The theory of consumer behaviour which presupposes that no two consumers think and behave alike. Similarly, 14 Principles of Management by Henry Fayol, states that practicing these principles in an organization, results into increasing organizational efficiency.

6) **Provides specific direction:** When hypothesis is finalized a definite and specific direction is provided to the research work. It makes way to the progress of investigation. In the absence of hypothesis it becomes extremely difficult to focus on research problem.

7) **Prevents blind research:** Hypothesis lights up the path of research. It distinguishes between scientific and unscientific inputs. It acts as a guide. Accuracy and precision is possible through hypothesis. Therefore, hypothesis prevents blind research.

8) **Economical:** Developing hypothesis in business research is economical. It saves time, money and energy of a researcher because it guides the researcher in the right direction. Hypothesis provides the basis for proper data collection. Relevant and correct information collected by a researcher through properly formulated hypothesis proves resource saving.

2.5.6 Types of Hypothesis

1) **Simple Hypothesis:** It reflects the relationship between one dependent variables (DV) and one independent variable (IDV).

   Examples:

   - Higher the unemployment (IDV), higher would be the rate of crime (DV) in society

- Lower the use of fertilizers (IDV), lower would be agricultural productivity (DV).
- Higher the poverty (IDV) in the society, higher would be the rate of crimes (DV).

2) **Complex Hypothesis:** It reflects the relationship between two or more dependent variables and two or more independent variables.

   Examples:

   - Higher the poverty (IDV) leads to higher rate of illiteracy (DV) in the society, higher would be the rate of crime (DV).
   - Lower use of fertilizer (IDV) and modern equipments (IDV), lower would be the agricultural productivity (DV)

3) **Directional Hypothesis:** A directional hypothesis is a prediction made by a researcher regarding a positive or negative change, relationship, or difference *between two variables* of a population. This prediction is typically based on past research, accepted theory, extensive experience, or literature on the topic. For example- "There will be a positive relationship between extra coaching and academic achievement"

4) **Non-Directional Hypothesis:** This form of hypothesis is used in studies where there is no sufficient past research available on which predication can be made about relation between variables. It does not stipulate the direction of the relationship. It is a statement that a relationship exists between two variables, without predicting the exact nature (direction) of the relationship. Eg. "Teacher – student relationship influence student's learning."

5) **Null Hypothesis:** This is a hypothesis that proposes no relationship or difference between two variables. It involves a statement that says there is no relationship between two groups that the researcher compares on a certain variable. It is denoted by "H_0". Example –
   - There is no relation between poverty and crime in a society.
   - 'There is no difference in the academic performance of high school students who participate in extracurricular activities and those who do not participate in such activities'

6) **Alternative Hypothesis:** This hypothesis proposes a relationship between two or more variables. Alternative hypothesis is denoted by "H_1". Example –
   - There is relation between poverty and crime in a society.
   - 'There is difference in the academic performance of high school students who participate in extracurricular activities and those who do not participate in such activities'

7) **Causal Hypothesis:** Causal hypotheses propose a cause and effect interaction between two or more variables. This hypothesis predicts the effect of independent variable on the dependent variable. Eg. 'High school students who participate in extracurricular activities spend less time studying which leads to a low grades.'

8) **Associative Hypothesis:** These hypotheses aim to determine if relationships exist between a set of variables. Do not indicate cause and effect.

9) **Testable Hypothesis:** These hypotheses predict relationship between the independent variable and the dependent variable. These variables are testable or measurable.

2.6 SAMPLING

2.6.1 Meaning

Sampling is a technique of selecting a subset (part) of the population to make statistical inferences (conclusion) from them and estimate characteristics of the whole population. Different sampling methods are widely used by researchers in market research so that they do not need to research the entire population to collect data. It is also a time-convenient and a cost-effective method and hence forms the basis of any research design.

In other words, Sampling means the process of selecting a part of the population. A population is a group of people that is studied in a research. It is difficult for a researcher to study the whole population due to limited resources such as time, cost and energy. Hence, the researcher selects a part of the population for his study, rather than studying whole population. This process is known as sampling. It makes the research activity manageable and convenient for the research.

2.6.2 Definition

According to Bogrdus, "Sampling is the selection of certain percentage of a group of items according to a predetermined plan."

2.6.3 Significance of Sampling

1) **Time Saving:** Since using a sample reduces the number of people that have to be reached out to, it reduces time. Sampling helps to save time in respect of data collection and its analysis. The data can be collected at faster rate, so also data analysis. Therefore, the researcher can get quick research results and accordingly can take timely action.

2) **Economical:** Since using a sample reduces the number of people that have to be reached out to, it also reduces cost. For any research, availability of funds is a constraint. A smaller sample requires less funds not only for data collection but also for processing and interpretation of data.

3) **Reduced resource deployment:** It is obvious that if the number of people involved in a research study is much lower due to the sample, the resources required are also much less. The workforce needed to research the sample is much less than the workforce needed to study the whole population.

4) **Convenient:** Sampling offers convenient to the researcher to collect the data. The work of data collection becomes easy, quick and economical. A researcher can complete his research project in time.

5) **Quality of Research Work:** The quality of research work may be improved due to sampling. The field staff will get sufficient time to collect the data from sample respondents. They need not to rush through the collection of data. Also, data analysis staff gets sufficient time for data analysis purpose. Therefore, overall quality of research work improves.

6) **Reduce Complexities:** Sampling helps to reduce complexities in research work. If a limited sample is used, then fewer respondents are required to collect data. As a result, the researcher may require less time for editing, coding and interpretation of data. Therefore, analysis can be quick and without complexities.

7) **Motivation to Research Staff:** Limited sample size brings relief to the research staff. They get motivated to collect the right information. This is because they get sufficient time for collection and analysis of data. They may also get higher rewards due to good quality research work.

8) **Detailed Information:** Due to sampling, the researcher can collect detailed information from the sample respondents. They can ask more questions than questions in questionnaire. Since there are lesser respondents, the data collected from a sample is intense and thorough. More time and effort is given to each respondent rather than having to collect data from a lot of people.

9) **Infinite Population:** If the population is too larger then the sampling method is the best way to find out solution.

10) **Accuracy of data:** Since the sample is indicative of the population, the data collected is accurate. Also, since the respondent is willing to participate, the survey dropout rate is much lower, which increases the validity and accuracy of the data.

2.6.4 Methods of Sampling

Probability Methods	Non-Probability Methods
- Simple Random Sampling	- Convenience Sampling
- Systematic Sampling	- Judgement Sampling
- Stratified Random Sampling	- Quota Sampling
- Cluster Sampling	- Snow-Ball Sampling

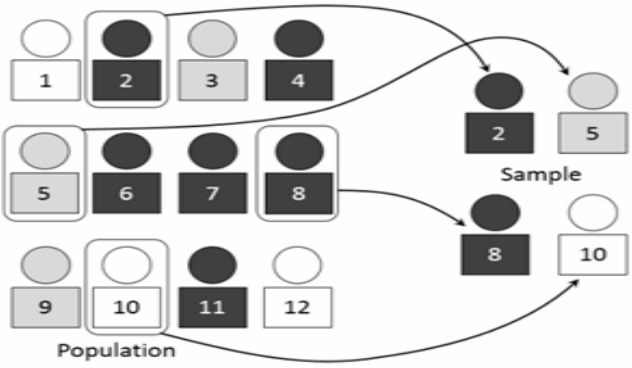
2.6.5 Probability Sampling Method

Probability sampling is a method of deriving a sample where the objects are selected from a population-based on the theory of probability. This method includes everyone in the population, and everyone has an equal chance of being selected. Hence, there is no bias whatsoever in this type of sample. Each person in the population can subsequently be a part of the research. The selection criteria are decided at the outset of the market research study and form an important component of research.

The various probability sampling methods are discussed as below:

1) **Simple Random Sampling:** This is the most popular method which is normally followed to collect research data. This technique provides every member an equal chance of being selected in the sample. The members are selected randomly and purely by chance. There are two sub-methods:

   - **Lottery Method:** Where each member is given a number and then the numbers are mixed and by drawing of lots, the sample is selected.



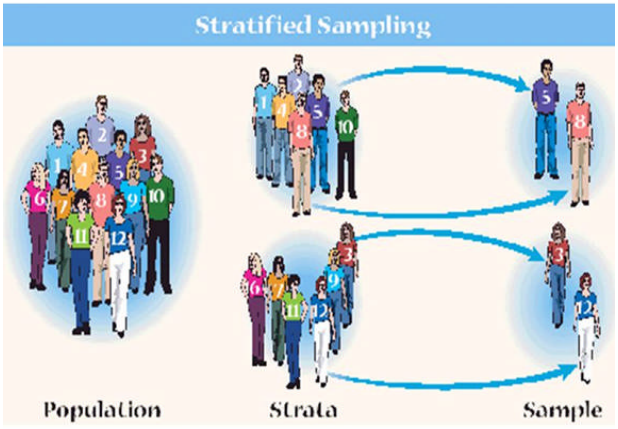
   - **Random Tables:** The members are given numbers and the numbers are placed in rows. The sample is selected from rows at random.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

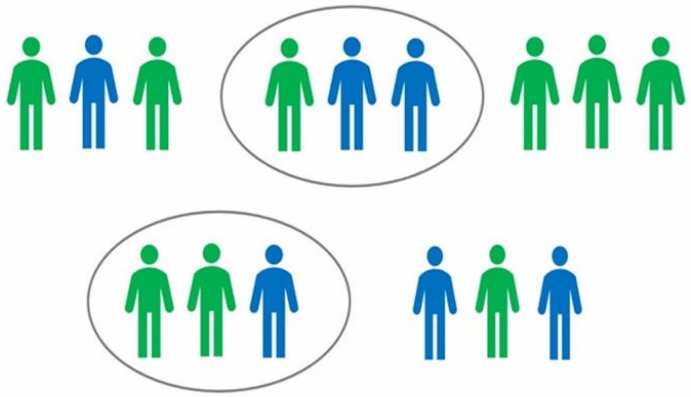
2) **Systematic Sampling:** Systematic sampling is a sampling method where the researcher chooses respondents at equal intervals from a population. Every member of the population is listed with a number, but instead of randomly generating numbers, individuals are chosen at regular intervals. Example : If the total population is 100 and the sample size is 10. Each respondent is given a number from 1 to100. A certain number is selected say no. 3. So number consist of 3, 13, 23, 33, 43, 63, 73, 83, 93 will be selected as sample.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

3) **Stratified Random Sampling:** This sampling method is appropriate when the population has mixed characteristics, and researcher wants to ensure that every characteristic is proportionally represented in the sample. Researcher divides the population into subgroups (called strata) based on the relevant characteristic (e.g. gender, age range, income bracket, job role). The strata are formed by researcher. Then he/she uses random or systematic sampling to select a sample from each subgroup. Example - The company has 800 female employees and 200 male employees. Researcher wants to ensure that the sample reflects the gender balance of the company, so he/she sorts the population into two strata based on gender. Then researcher uses random sampling on each group, selecting 80 women and 20 men, which give researcher a representative sample of 100 people.



4) **Cluster Sampling:** Cluster sampling also involves dividing the population into subgroups, but each subgroup should have similar characteristics to the whole sample. The clusters are naturally formed. Instead of sampling individuals from each subgroup, researcher randomly selects entire subgroups. If the clusters themselves are large, researcher can select sample from each cluster using simple random or systematic sampling method. This method is good for dealing with large and dispersed populations, but there is more risk of error in the sample, as there could be substantial differences between clusters. It's difficult to guarantee that the sampled clusters are really representative of the whole population. Example - The Company has offices in 10 cities across the country (all with roughly the same number of employees in similar roles). Researcher doesn't have the capacity to travel to every office to collect data, so he/she uses random sampling to select 3 offices – these are the clusters.

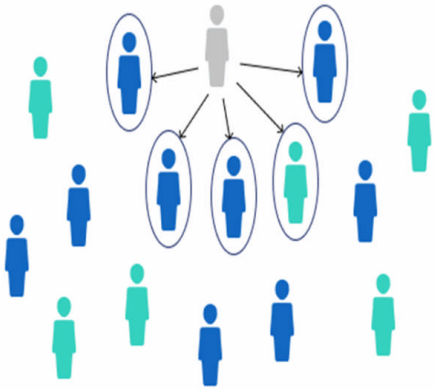


2.6.6 Non-Probability Sampling Method

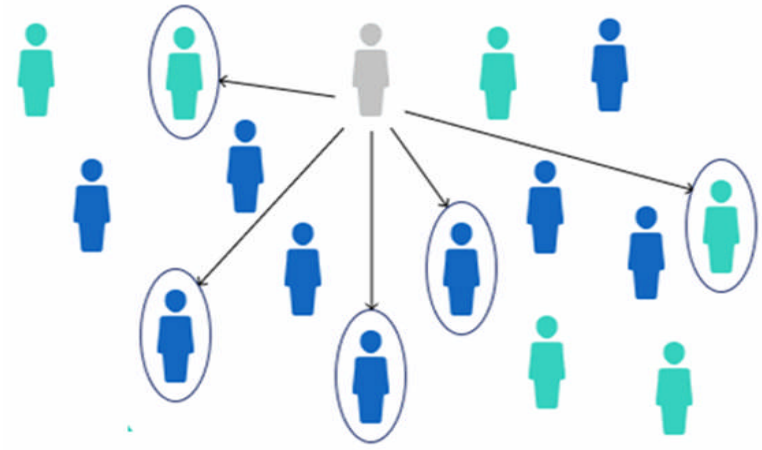
Non-probability sampling is defined as a sampling technique in which the researcher selects samples based on the subjective judgment of the researcher rather than random selection. It is a less stringent method. This sampling method depends heavily on the expertise of the researchers. It is carried out by observation, and researchers use it widely for qualitative research.

1) **Convenient Sampling:** It is a type of where samples are selected from the population only because they are conveniently available to the researcher. Only those members are selected which are easily accessible to the researcher. Eg. A researcher may visit a college or a university and get questionnaires filled in by volunteer students. Similarly, a researcher may stand in a market and interview the volunteer persons.

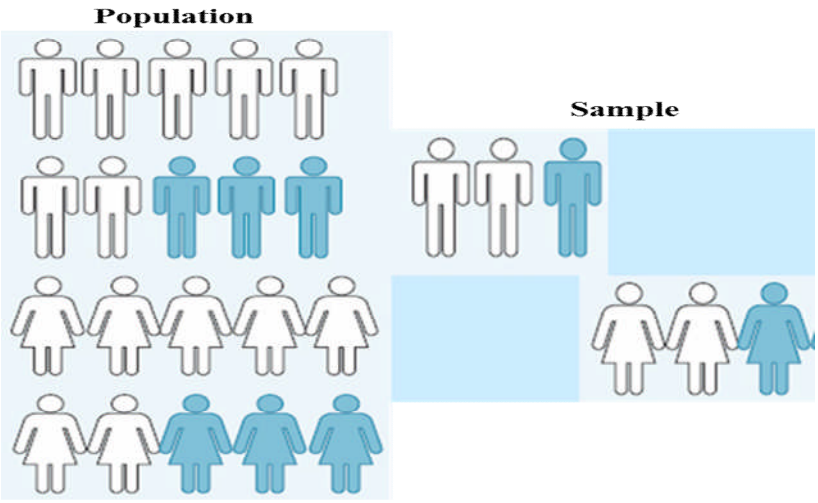
   Ideally, in research, it is good to test a sample that represents the population. But, in some research, the population is too large to examine. It is one of the reasons why researchers rely on convenience sampling, which is the most common non-probability sampling method, because of its speed, cost-effectiveness, and ease of availability of the sample.



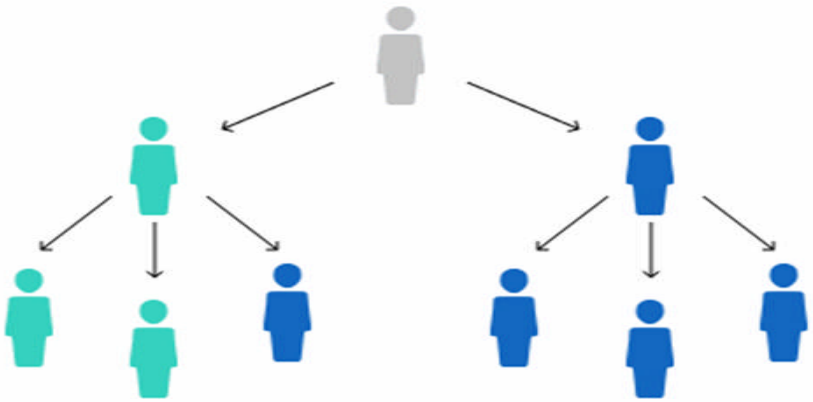
2) **Judgment or Purposive Sampling:** In this method of sampling researchers select the samples based purely on the researcher's knowledge and credibility. In other words, researchers choose only those people who they deem fit to participate in the research study. Judgmental or purposive sampling is not a scientific method of sampling, and the downside to this sampling technique is that the preconceived notions of a researcher can influence the results. Thus, this research technique involves a high amount of ambiguity.



3) **Quota Sampling:** Under this method, the researcher allocates certain quota to certain groups under study. The quotas may differ from each area depending upon certain factors like age, occupation, income etc. Eg. A researcher studying the newspaper reading habits of college students may select 10 colleges for data collection. He may fix quota for each college based on certain criteria. He may select 100 students from one college, may be because the number of students is more in that college; and he may select only 20 students from another college because of less number of students in that college.



4) **Snow-Ball Sampling:** Snowball sampling helps researchers find a sample when they are difficult to locate. Researchers use this technique when the sample size is small and not easily available. This sampling system works like the referral program. It is a sampling design in which respondents selected earlier are asked to identify other sample members. Eg. To find out Mercedez Benz car owners in the city like Mumbai. In this case, researcher may locate one Mercedez Benz owner and collect the names of 2-3 other Mercedez Benz owners.



2.6.7 Factors determining sample size

1) **Area of Research:** The number of sample respondents depends on the area of research. If the research is conducted at national level, it may require more number of respondent. If the research is conducted at local level, it may require less number of respondents.

2) **Availability of Funds:** Generally, the researcher may be constrained by the limitation of funds to conduct the research. Therefore, when the researcher has limited amount of funds allocated to the research activity, the sample size would be lesser as compared to when the researcher has larger amount of funds.

3) **Availability of Manpower:** The researcher may require manpower to conduct surveys, interviews or for conducting experiments, observation etc. Eg. If the researcher has a good number of filed staff to conduct interviews, he may select the larger sample size of respondents and vice-versa.

4) **Time Frame:** The sample size may depend on the time frame of research. If the researcher has lot of time available to conduct the research, he may select a large sample size of respondents and vice-versa.

5) **Nature of Research:** The nature of research may influence the sample size of respondents. Eg. In case of academic research, the researcher may be constrained with the limitations of funds, and therefore, he may select a smaller sample size. However, in the case of census survey of population, the sample size will be the entire population of the country.

6) **Method of Sampling:** The method of sampling may influence the sample size of respondents. Eg. If convenience sampling method is used, the researcher may consider a smaller sample size to obtain responses. However, in case of stratified sampling or cluster sampling, the researcher needs to select a larger sample size of respondents.

7) **Method / Sources of Data Collection:** The method of data collection may influence the sample size of respondents. Eg. If researcher collects data through interviews, he may select a larger sample size of respondents. However, if the researcher adopts observation method, he may consider a smaller sample size for the purpose of observation.

8) **Judgment of the Researcher:** At times, the researcher may use his judgement in deciding in the sample size. He may consider a smaller sample size, if he is confident in getting the adequate data from a smaller sample size. However, if the researcher feels that he needs to select a larger sample to collect responses, he may select a larger sample size.

2.7 SUMMARY

This unit talks about formulation of research problems, which is a first step in research activity. Proper formulation of research problems enables researchers to carry on research activities accurately and researchers understand what kind of research data is required to collect and achieve research objectives.

Another part of this unit is about Research Design. Research design is a plan of research. It enables us to plan the various activities of research such as sampling method, data collection and analysis method, resources required etc. Research design enables us to start and end the research on time. Delayed research may not hold any importance.

Next part of this unit Review of Literature. It refers to the previously done research. This provides insight to the researcher and provides direction as to how he can carry on his/her research activity.

Hypothesis is another part of this unit. It refers to the assumption made by the researcher, which he/she tries to cross check after collecting the data. It can be proved to be correct or incorrect. That the researcher can confirm only after collecting data from the respondents.

Sampling is the last part of this unit. Since the researcher cannot collect data from the entire population/universe under study, he/she can select samples by adopting different methods of sampling. Research data is collected from these samples and the result of it is generalized on the entire population / universe under study.

2.8 EXERCISE

FILL IN THE BLANKS

1. _____ is a question that a researcher wants to answer or a problem that a researcher wants to solve
   A) Hypothesis B) Data Analysis
   C) Research Problem D) Research Report

2. _____ is a logical and systematic outline of research project prepared for directing, guiding and controlling a research work
   A) Universe / Population **B) Research Design**
   C) Hypothesis Testing D) Review of Literature

3. Review of Literature enables the researcher _____
   A) To pirate the research software
   B) To undertake plagiarism
   C) To identify gaps in research
   D) To collect data from entire population

4. _____ hypothesis proposes no relationship or no difference between two variables.
   A) Simple B) Associative C) Alternative **D) Null**

5. _____ is one of the probability methods of sampling
   A) Cluster Sampling B) Quota Sampling
   C) Judgemental Sampling D) Snow-Ball Sampling

TRUE OR FALSE

1. "A study on E-Commerce" is a correctly formulated research problem. **FALSE**
2. Research design contains details regarding nature of the research and objectives of study, time period of research study, universe and sample size of respondents, type and source of research data required and techniques of data collection and analysis. **TRUE**
3. Sampling refers to consult the available publications such as books, journals, magazines, research reports and similar other publications before starting his/her own research activity. **FALSE**
4. Intuition is one of the sources of generating hypothesis. **TRUE**
5. Directional hypothesis is used in studies where there is no sufficient past research available on which predication can be made about relation between variables. **FALSE**
6. Snow-ball sampling is a sampling method where the researcher chooses respondents at equal intervals from a population. **FALSE**

MATCH THE PAIRS

Group A	**Group B**
6) Review of Literature	g) Outline of the research work
7) Research Design	h) Helps to get familiar with previous research studies undertaken
8) Alternative Hypothesis	i) Non-Probability Sampling Method
9) Judgement Sampling	j) Denoted as "H_1"

(1 – b, 2 – a, 3 – d, 4 – c)

ANSWER IN BRIEF

5) What is Review of Literature? Elucidate its significance.
6) How to formulate research hypotheses?
7) Describe the different sources of generating hypotheses.
8) Highlight the importance of hypothesis in research.
9) Briefly explain types of hypothesis.
10) Define the term 'Sampling'. Explain its significance.

11) Discuss various methods of sampling.

12) What are the factors determine sample size?

13) Write a note on:

- Formulation of Research Problem
- Research Design

2.9 REFERENCES

- https://www.youtube.com/watch?v=UqtckUep840
- https://www.slideshare.net/maheswarijaikumar/a-research-problem
- https://www.questionpro.com/blog/research-problem/
- https://www.researchgate.net/publication/325846748_FORMULATING_AND_TESTING_HYPOTHESIS
- https://www.questionpro.com/blog/non-probability-sampling/
- https://www.scribbr.com/methodology/sampling-methods/

❖❖❖❖

Module - II
RESEARCH PROCESS

3

STAGES IN RESEARCH PROCESS

Unit Structure

3.0 OBJECTIVES

After going through this unit the learner will understand
1. Various Stages in Research Process
2. Primary data
3. Different sources of collecting Primary Data
4. Limitations of Primary Data

3.1 INTRODUCTION

Research refers to the systematic investigation into and study of materials and sources in order to establish facts and reach new conclusions. To attain correct results Research has to be done with a predefined process. In this unit we will learn What are different stages in Research Process, How Primary Data is collected

3.2 STAGES IN RESEARCH PROCESS

3.2.1 MEANING

Each and every researcher needs to follow appropriate research process for successful completion of his or her research project. A researcher who is conducting scientific research needs to follow a systematic process to study the research problem and to arrive at a conclusion. The scientific research process includes a sequence of various steps that needs to be followed while undertaking the research project. Every research problem is unique and it requires research work to be

carried out accordingly. Following different steps to be followed in the research process.

I] Identifying and selection of research problem:

The most important step in the research process is to identify and select research problem, it is often said a well defined problem is half solved. High degree of intellectual work is needed to identify and define the research problem. A properly defined research problem will help researcher to collect reliable data to solve the research problem and in that case he will be in a better position to arrive at a suitable conclusion based on Analysis of data.

Following are some of the Essentials of a good research problem

1. Researchable : the identified research problem can be studied through collection and Analysis of data
2. Understandable : The research problem should be understandable it should be well formulated and logically structured
3. Ethical : The selected research problem should not harm the researcher respondents and Society. You should not create any embarrassment every of the stakeholder in the world in research process
4. Purposeful : The research problem should be defined in such a way that after concluding this study it should be able to solve the situation.
5. Manageable : the research problem should be manageable the researcher should take care that the problem should be within limits of his skills resources and time.

II] Review of literature:

Review of literature refers to study of available knowledge in respect of research problem. A researcher can study Various publications such as journals books research reports other published matter particularly the researcher should go to the similar research study that were conducted previously. Literature review is important because it is the duty of research to find out what study has been already done literature review will help the researcher to find out gaps in earlier study. On the basis of those gaps researcher can formulate his Research Design Literature review on both matic as well as methodological direction to the researcher.

III] Formulation of the hypothesis & Research Design

Once the researcher has clearly defined the research problem and has made in-depth literature review he needs to formulate Research Design for his study. The researcher also need to formulate hypothesis for his research.

This is a tentative assumption made to test its logical and empirical consequences. The hypothesis should be formulated on the basis of index knowledge of the research problem. A well defined hypothesis will clearly

identify the kind of data required by the researcher to conduct the study which will help him to create a suitable Research Design.

Research Design is a systematic and logical plan prepared to conduct a research study we can call it as a blueprint for research study. The Research Design will include guidelines in respect of collection, measurement and analysis of data required for the research study.

IV] Sampling Design :

A researcher needs to collect Information for his study however it is not possible to collect information from each and every member of the universe hence he needs to select sample for data collection. The research needs to select an appropriate sampling method which is suitable for the study. The selected sample must be representative of the universe. Its size must be flexible and sufficient enough to provide required information which can be analysed and used to test the hypothesis.

V] Designing Questionnaire :

The researcher needs to collect primary data for his study. He cannot rely only on secondary data. Primary data can be collected through various sources however the most commonly used is Questionnaire. Questionnaire is a list of questions that will be asked by research to the respondents for Collection of data. While designing questionnaire researcher need to consider various aspects such as what type of information is needed, what type of technique will be used for conducting the research and he also need to take care regarding proper wording and sequence of the questions.

VI] Collection of data :

The researchers need to collect all relevant information in respect of his research problem as per Research Design through various sources. He can use primary data and secondary data for purpose of research. While collecting data researcher need to consider that information collected is up-to-date and free from any bias. the information must be complete in all aspects and the information must be relevant to the research problem.

VII] Processing and analysing data :

The researcher collects data from primary as well as secondary sources however the data collected is in raw form and it needs to be processed for further analysis.

Processing of data in hall editing, coding, classification, and tabulation of data. Editing of data refers to removing unwanted and irrelevant information it is necessary to check errors and omissions in data collection. coding refers to assigning different codes to the collected data which helps in further processing of the information classification of data refers to grouping of data under different categories and tabulation of data

involves transferring all classified data into tabular form tabulation of data helps in analysis and interpretation of data.

Once data is organised the researcher needs to analyse the data. Analysis of data is very crucial as it tries to establish relationship between information and research problem. Once that relationship is established interpretation of data and finding out logical conclusions of study is possible.

VII] Hypothesis Testing :

Once analysis and interpretation of data is over the researcher must test hypothesis. hypothesis testing is necessary because researcher need to confirm whether the finding of the research supports the hypothesis or provides a contrary picture. Researcher may apply various tests such as Chi-square test, F-test, t-test for testing the hypothesis. Once hypothesis is tested its result will either lead its acceptance or its rejection.

IX] Preparation of Research Report :

The researcher must prepare a research report which includes all the findings and conclusions of his study. The report can be divided into three parts.

1. Preliminary Content. : This includes preface, table of contents and all other related authorisations and declarations in respect of the Research report.
2. Main body - This includes introduction, literature review, research methodology, data analysis, conclusion and recommendations.
3. Appendix - this will include all the annexure bibliography and other supportive documents related to the research.

3.3 PRIMARY DATA

In statistical analysis, collection of data plays a significant part. The method of collecting information is divided into two different sections, namely primary data and secondary data. In this process, the primary data is assembling data or information for the first time, whereas the secondary data is the data that has already been gathered or collected by others.

3.3.1 Definition of Primary Data:

Primary data is the data that is collected for the first time through personal experiences or evidence, particularly for research. It is also described as raw data or first-hand information

3.3.2 Characteristics of Primary Data

Following is list of characteristics of Primary Data.

1. Primary data is original data

2. Collection of primary data is an expensive exercise
3. Collection of primary data is a time consuming
4. Data Lake to be collected considering research problem
5. Primary data is collected from relevant respondents
6. There are several methods of collecting primary data such as Survey, Observation experimentation etc.
7. Primary data need to be processed and analysed before its use.
8. Primary data is considered to be basic input in research study.

3.4 METHODS OF COLLECTING PRIMARY DATA

Primary data is collected by researchers by interacting with or observing the respondents; it can be collected through various methods such as survey, interview, observation and experimentation.

1. Survey
2. Interview
3. Observation
4. Experimentation

3.4.1. Survey Method :

The Survey method is the technique of gathering data by asking questions to people who are thought to have desired information. This data is comprehensive information gathered from a target audience about a specific topic to conduct research. Survey can be a Census or Sample Survey

Census : In case of Census Survey the entire universe is contacted for collecting information.

Sample Survey : In case of sample survey information is collected from selected respondents of the universe.

Methods of Survey:

The researcher can apply different methods of Survey based on the research problem, No of respondents from whom information to be collected, time available to the researcher and his financial limitations. Today technology available with the researcher and respondents also plays a crucial role in selecting the survey method. Following are different methods of survey.

1. **Interview :** In this method researcher meets respondents personally and asks questions regarding his research problem and collects the relevant information. Which is the most traditional, expensive and most effective method of collecting data.
2. **Telephonic Survey :** In this method the researcher does not meet respondents personally. However the survey is conducted through telephonic discussions. The researcher asks various questions to the respondents on telephone and collects the required information.

3. **Mail Survey :** This is another method of Survey it is generally used when the number of respondents is very high and there is a geographical challenge for a researcher. In this method a questionnaire is prepared and sent to the respondents via mail or it can be advertised in newspapers and magazines and respondents are asked to fill in the information and send it back to the researcher.

4. **Online Surveys :** Now a days this method of collecting information is gathering momentum today we have high penetration of internet services in India particularly in urban and semi-urban areas, in this method the researcher can collect information from respondents by sending the Google Form or Microsoft from through email or other communication applications such as WhatsApp or Telegram.

Each and every method has its own Advantages and disadvantages. Following are Advantages and disadvantages of Survey

3.4.2. Advantages of Survey Method

1. **Reliable and Detailed Information :** Information collected through survey is more reliable. The researcher collects information from respondents directly sold is assumed to be more reliable and correct however accuracy of the information depends upon the type of questions the survey method and the respondents bias towards the research problem

2. **Versatile :** This method is most versatile. A researcher can restructure or modify the questions according to the responses he has received. He can delete certain questions according to the situation, he may ask questions to the respondents in the different languages to obtain responses from the respondents.

3. **Personal touch :** This method allows the researcher to create personal relationship with respondents which help the researcher to obtain candid responses from the respondents on the questions which otherwise could have been avoided by the respondent.

4. **Cost effective :** This method is cost effective in terms of quality of information obtained through this method.

3.4.3. Disadvantages of Survey Method.

1. **Time consuming :** This method is time consuming, a lot of time is needed to prepare a questionnaire, to take appointment of the respondent and obtain responses from him.

2. **Personal Bias :** The quality of information obtained from this particular method can be affected due to personal bias of interviewer as well as respondent. The respondent may not provide correct information while answering the questions and at the same time interviewer / researcher may bypass or twist the questions which prohibit respondents from providing the correct information.

3. **Expensive :** Survey method is expensive and it is difficult for a researcher with limited financial strength.

3.4.4. Interview Method :

In the case of an interview the researcher or interviewer interacts with respondents personally by meeting the respondent. In this method there is face-to-face interaction between researcher and the respondent. It is also called a Personal Interview. In this method the interviewer asks questions to the respondent and collects information through its answer given by the respondent. There are various types of personal interviews.

1. **Structured Interview** - In this type of interview the interviewer uses a set of predetermined questions and highly standardized techniques of recording.

2. **Unstructured Interview** - In this method interviewer does not follow a system of predetermined questions and standardized techniques of recording information and it is fully based on flexibility and requirements.

3. **Clinical Interview -** This type of interview is concerned with broad underlying feelings or motivations or with the course of an individual's life experience.

4. **Non-directive Interview -** In this type the interviewer simply to encourage the respondent to talk about the topic with a bare minimum of direct questioning.

5. **Focused Interview -** In this type of interview the task of the interviewer is to confine the respondent to discuss about given experience to them and its effects.

Nowadays such interviews can be taken through video conferencing platforms such as Zoom, Google Meet or Microsoft Teams.

3.4.5. Advantages of Interview

- Higher response rate.
- When the interviewees and respondents are face-to-face, there is a way to adapt the questions if this is not understood.
- More complete answers can be obtained if there is doubt on both sides or a particular information is detected that is remarkable.
- The researcher has an opportunity to detect and analyze the interviewee's body language at the time of asking the questions and taking notes about it.

3.4.6. Disadvantages of Interview

- They are time-consuming and extremely expensive.
- They can generate distrust on the part of the interviewee, since they may be self-conscious and not answer truthfully.
- Contacting the interviewees can be a real headache, either scheduling an appointment in workplaces or going from house to house and not finding anyone.

3.4.7 Observation Method :

This is another method of collecting data, in this method the researcher does not Ask any question to the respondent but observes the behaviour of respondents. The observation Method can be defined as "Observation is a technique that involves systematically selecting, watching, listening, reading, touching, and recording behavior and characteristics of living beings, objects, or phenomena."

3.4.8. Advantages of Observation Method

1. **No Respondent Bias :** There is no any respondent bias involved in collecting data. in this method behaviour of respondents is observed and respondents may not be aware that they are being observed hence they behave naturally without any bias.
2. **Accurate and Reliable :** The information collected through this method is accurate and reliable as it is collected without any bias and it is collected at the point of action or reaction of the respondent.

3.4.9. Disadvantages of Observation Method

1. **Time Consuming :**It is time consuming method for collection of data, researchers need to observe behaviour of all respondents patiently.
2. **Expensive :** It is an expensive method of Data Collection as a researcher may have to appoint trained staff for observing behaviour of respondents.
3. **Non Verbal Method :** It is a nonverbal method of Data Collection, in depth collection of information like interview method is not possible through this method.

3.4.10. Experimentation Method :

Experimentation method of data collection is normally followed in the field of science. In this method research study cause and effect relationship between two or more variables through experiments.

An experiment is a structured study where the researchers attempt to understand the causes, effects, and processes involved in a particular process. This data collection method is usually controlled by the researcher, who determines which subject is used, how they are grouped and the treatment they receive.

3.4.11. Advantages of Experimentation Method

1. Researcher have full control on the experiment
2. Researchers can obtain reliable and accurate information through experiments.
3. It eliminates any kind of bias in Collection of data.

3.4.12. Disadvantages of Experimentation Method

1. Expensive method of Data Collection
2. It could be time consuming
3. Small error could result into failure of experiment

3.4.13. Schedules

The researcher may use a schedule while collecting data from interview or observation. Schedule is an instrument used to collect data from the respondents Schedule contains questions, statements (on which respondents are expected to give their opinion) It provides blank spaces/tables to the respondents to fill up their responses. Schedule is important as it help researcher to attain objectivity by reminding him different aspects to be observed.

The features of schedules are :

1. The schedule is presented by the interviewer.
2. The questions are asked and the answers are noted down by him.
3. The list of questions is a more formal document, it need not be attractive.
4. The schedule can be used in a very narrow sphere of social research.

There are several kinds of schedule.

1. **Rating Schedules** : It is a schedule which contains positive and negative statements of opinion on the phenomenon. The respondents are to express their opinions, preferences etc, respondents over statements on the phenomenon studied and schedule is used to record their responses.
2. **Documents Schedules :** are used to collect data/information from recorded evidences and/or case histories. Here the blanks, functional issues related blanks and the like to be filled up from records and documents are present.
3. **Survey Schedules :** Survey schedules are like questionnaires, it is a list of questions that need to be answered by respondents.
4. **Observation Schedules :** Observation schedules are schedules used when the observational method of data collection is used.

3.4.14. LIMITATIONS OF PRIMARY DATA

1. **Expensive :** Primary data collection methods are expensive comparing it to secondary data. It requires appointment and training of staff for collecting information through Interviews, surveys or observation. It also requires sophisticated equipments for conducting experiments. However such expenses are not required when secondary data is used by researcher.
2. **Time consuming :** Collecting data through primary sources could be time-consuming as it requires time to collect information from each and every respondent, observe the sample as well as experiments.
3. **Bias :** Reliability and accuracy of primary data could be affected due to bias of researcher or respondent. Respondents may provide fake information to the researcher regarding sensitive topics. Interviewer / researcher may also not take sufficient efforts to collect information from the respondents.
4. **Processing of data :** the information collected from respondent need to be edited, coded, classified and tabulated before its analysis only after properly analysing the data useful inferences can be drawn.
5. **Sampling Error :** While collecting information it may not be possible to collect information from all the respondents of the Universe. Researcher may have to collect information from sample i.e. selected respondents from the entire population, sample selection could be wrong which may lead to collection of wrong information and wrong conclusions from the research.

3.5. SUMMARY

This module has discussed the Research Process thoroughly, the unit also discusses Primary Data its Features Advantages and Disadvantages. It also discusses various methods of collecting primary data thoroughly. The unit has also provided Limitations of Primary Data. However a researcher may not rely only on Primary Data, he may have to use secondary data as well. It is discussed in the next unit

3.6 EXERCISE

1. Describe Stages in Research Process
2. Explain Importance of Primary Data
3. Explain methods of collecting primary data
4. Explain limitations of primary data
5. Write Short Notes on
   a. Interview
   b. Schedule
   c. Observation Method
   d. Experimentation Method
   e. Survey Method

State following statements are True or False

1. The research problem should be clearly defined
2. Literature review to provide good insight into research problem
3. Tabulation of data helps in analysis and interpretation of data
4. Collecting Primary data is inexpensive
5. There are no disadvantages of experimentation

(True - 1,2,3, False - 4,5)

3.7 REFERENCE

- Research Methodology : Michale Vaz, Aurora Vaz : Manan Prakashan
- Research Methodology Methods and Techniques : C. R. Kothari, New Age International Publishers
- Research Methodology : R Panneerselvam, PHL L:earning

4

SECONDARY DATA

Unit Structure

4.0 OBJECTIVES

After going through this unit the learner will understand

1. What is Secondary Data ?
2. What are different methods of collecting Secondary Data ?
3. What are limitations of collecting secondary data ?
4. Factors influencing selecting methods of collecting data
5. The process of Questionnaire

4.1 INTRODUCTION

In earlier unit we learned about primary data. However the researcher also needs to look into Secondary Data. The researcher should go through all published material regarding his area of study, that published material such as Journal, Magazines, Research Papers, Websites, books etc is called Secondary data. Based on available secondary data the researcher will identify the area in which he can conduct research and collect necessary primary data.

4.2 SECONDARY DATA

This data is readily available for researchers. It is available from published sources such as newspaper, magazines, research journals, research papers etc. Generally any researcher makes an attempt to obtain information through secondary data and if secondary data is insufficient or redundant the researcher will try to obtain primary data.

4.2.1 Features of Secondary Data

1. **Readily available data** : Secondary data is collected from published as well as unpublished sources. Secondary data is available internally (from own records, accounts etc) or externally. It is already collected, processed and published by other researcher.
2. **Quickly collectible :** Secondary data is readily available hence it is easy to collect. The researcher can easily collect secondary data as compared to primary data.
3. **Less time consuming and Less Expensive :** Secondary data is easy to collect as it is readily available hence it can be collected in less time. It is also less expensive to coolest secondary data.
4. **Comprehensive** : Secondary data provide quantitative data. Information available through secondary data is huge. Secondary data is available for almost all subjects.
5. **Relevance :** Secondary data is available for almost all subjects. However the secondary data available may not be related with the research topic hence the researcher need to evaluate the available secondary data.

4.2.2 Advantages of Secondary Data

Followings are advantages of Secondary Data.

1. **Less Processing of Data :** Secondary data is already processed earlier hence it requires less processing of data. It becomes easy for researcher to use secondary data for his research.
2. **Cost Effective :** Collecting secondary data is cost effective. The secondary data is available to researcher through internal or secondary sources. The researcher does not require to appoint and trained staff to collect data or make any conduct survey or experiments.
3. **Complimentary to Primary Data** : The Secondary data is information collected and processed by other researcher for some other issue hence it may not be hundred percent relevant for the research problem of the researcher. In such scenario Secondary Data can be used to support the primary data collected by the researcher.
4. **Less Time Consuming :** The secondary data is readily available to the researcher trough various internal and external sources hence collection of secondary data is quick and less time consuming for the researcher.
5. **No Sampling errors** : The secondary data is readily available to the researcher trough various internal and external sources. There is no any requirement of sampling for collecting of secondary data hence there os no any issue of sampling error ion Secondary Data.

4.2.3 Limitations of Secondary Data

1. **Unsuitable :** Secondary data may not be suitable for all types of research. Secondary data is not useful in case of commercial research. In case data is required to solve a business problem for example customer satisfaction. In such case Secondary Data is not suitable researcher need to collect primary data to solve this problem.
2. **Accuracy :-** Secondary data may not be accurate. Secondary data available to researcher may not be genuine which results in inferior quality of secondary data. Researcher need to be careful while selecting secondary data to be used for his research.
3. **Relevance :** The secondary data available to the researcher may not be collected for the research which is undertaken by the researcher. Secondary data is data collected by some other researcher earlier for some other research problem. Considering above Secondary Data may not be relevant for the research.
4. **Biased Information :** The researcher does not have any control over quality of secondary data. Secondary data may be affected due to bias of researcher and respondents when it was originally collected.
5. **Adequacy :** The available secondary data may not be adequate for the researcher. Secondary data is available information which was not collected to research the problem under consideration. In such a case researcher can not rely only on secondary data but he has to collect primary data for his study.

4.3. METHODS / SOURCES OF COLLECTING SECONDARY DATA

1. **Internal Sources :** Internal sources refers to data available within the organisation through its own records. An organisation generates huge data which could be useful for decision making. Following are varius internal sources of Secondary Data.
   a. Purchase and Sales Records
   b. Record of receivables and payables
   c. Record related to production
   d. \Financial Statements such as Profit & Loss, Balance Sheet and Cashflow.
   e. Marketing Information
   f. Records related to employees etc.

2. **External Sources :** Secondary data can be collected through various external sources also. It includes information collected from various published sources such as books, magazines, newspapers reports, research paper online sources etc.

1. **Government Publications :** Central, State and local government bodies collect huge data which can be used by researchers. Following data is collected by various government organisations.

   Census which provide demographic details of population by Registrar General of India. Statistical Information on National Income and its various components is published by The Central Statistical Organisation. The Director General of commercial Intelligence provide information on Imports and Exports. Information on Price of various commodities is provided by Ministry of Commerce & Industry.

   Apart from above various other bodies organisations such as Planning Commission, Reserve Bank of India, Ministry of Finance, National Sample Survey various boards collect and publish information which can be used as secondary data by the researcher.

2. **Private Organisations:** Apart from government there are various private organisations which collect data and make it available to anyone by charging some fees. In India there are organisations such as The Operation Research Group, The India Market Research Bureau, Pathfinder etc who provide required information to the researcher.

3. **General Publications :** It includes various publications such as newspapers, magazines, trade and professional journals, research papers, publications of various Commerce and Trade Associations. It can also use publications made by International Organisations such as WHO, WTO, World Bank, IMF, United Nations etc. Apart from this researcher can also use information from various specialised libraries, reference libraries for collection of secondary data.

4.4. FACTORS AFFECTING THE CHOICE OF METHODS OF DATA COLLECTION

The researcher needs to collect data for his research. It is a crucial decision for the researcher to choose a method of data collection for his study. The most desirable approach with regard to the selection of the method depends on the nature of the particular problem and on the time and resources (money and personnel) available along with the desired degree of accuracy needed for the study. There are several factors that affect the decision to choose a particular method of collecting data which are as follows.

1. **Nature, scope and object of enquiry:** This constitutes the most important factor affecting the choice of a particular method. The method selected should be such that it suits the type of enquiry that is to be conducted by the researcher. This factor is also important in deciding whether the data already available (secondary data) are to

be used or the data not yet available (primary data) are to be collected.

2. **Availability of funds:** Availability of funds for the research project determines to a large extent the method to be used for the collection of data. When funds at the disposal of the researcher are very limited, he will have to select a comparatively cheaper method which may not be as efficient and effective as some other costly method. Finance, in fact, is a big constraint in practice and the researcher has to act within this limitation.
3. **Time factor:** Availability of time has also to be taken into account in deciding a particular method of data collection. Some methods take relatively more time, whereas with others the data can be collected in a comparatively shorter duration. The time at the disposal of the researcher, thus, affects the selection of the method by which the data are to be collected.
4. **Precision required:** Precision required is yet another important factor to be considered at the time of selecting the method of collection of data.
5. **Availability of Research Staff** : This is also an important factor while selecting a method to collect primary data researcher may have to appoint trained staff who can collect and process data. Without trained staff it is highly difficult for the researcher to collect and process data hence it is an important factor in selecting data collection method.
6. **Availability of Respondent :** Availability of respondents is also a crucial factor while deciding data collection method. No of respondents from which data can be collected in a given time frame decide the method of collecting data. If no of respondents are less, Researcher can prefer personal interview, if the number is large then he may prefer mail survey.

4.5 QUESTIONNAIRE

4.5.1. Meaning

Questionnaire is an important tool for collecting primary data. It was first developed by the Statistical Society of London in 1838 and has been in continuous use ever since. A Questionnaire is a list of questions that researchers ask to the respondents for collection of information related to the research problem. There are different types of questionnaire which are listed below.

1. **Formal standardized questionnaire :** Standardized questionnaire is also known as structured questionnaire. This type of questionnaire is scientifically designed and segmented to collect accurate information from the respondents. This questionnaire is used to collect quantitative data which is information recorded as a count or numerical value.

Standardized questionnaires are best used when researcher have already formed an initial hypothesis. These questionnaires can be further classified into

a. **Close Form Questionnaire :** This tye of questionnaire do not give too many choices to the respondents, rather respondent have to select answer from given options only. For example options provide could be Yes / No True / False Agree / Disagree

b. Open Ended Questionnaire : In this type of questionnaire respondent is not provided with options to select as answer rather respondent can answer the questions in descriptive manner.

2. Exploratory questionnaire : Exploratory Questionnaire is also known as unstructured questionnaires. They're used to collect qualitative data which is information that can be observed and recorded but isn't numerical in nature. In this type of questionnaire there is no any specific segmentation of questions for collection of information.Exploratory questionnaires are ideal for conducting interviews. The researcher or Interviewer with the help of an unstructured questionnaire can obtain information far better than a structured questionnaire, reason for this is an unstructured questionnaire gives flexibility to the interviewer while conducting the interview. It is best suitable when a researcher is in the early stages and want to learn more about a topic before designing a solution or hypothesis.

3. Scaled questionnaires: In this type of questionnaire the researcher designed the questionnaire in a such way that the while answering the questions respondents are asked to scale the answers based on a given rating prescribed by the question.

4. Pictorial Questionnaire : This type of questionnaire is not regularly used. In this type of questionnaire researcher provide various pictures / images and ask respondents to choose from the given set of images as their response to the given question.

5. Codified and Uncodified Questionnaire : In case of codified questionnaire researcher assign codes to the expected answers and respondents are expected to fill in the code while answering the questions. An uncodified questionnaire is a simple plain questionnaire without using any codes

4.5.2. STEPS IN DESIGNING QUESTIONNAIRE

Designing a proper questionnaire is a key for success in research. Researcher need to be careful while designing the questionnaire which will be used for collecting data for his research. Followings are steps involved in the development of a questionnaire:

1. Decide the information required.
2. Define the target respondents.
3. Choose the method(s) of reaching your target respondents.

4. Decide on question content.
5. Develop the question wording.
6. Put questions into a meaningful order and format.
7. Check the length of the questionnaire.
8. Pre-test the questionnaire.
9. Develop the final survey form.

1. **Deciding on the information required :** While designing the questionnaire researcher needs to understand his research problem. Through understanding of the research problem in deciding what information is required to be collected for the research. This understanding helps the researcher in drafting relevant questions.
2. **Define the target respondents :** The researcher needs to know from whom the information needs to be collected for his research. The questionnaire needs to be framed taking into consideration respondents familiarity of the language, status, technical words, age, education etc of the respondents.
3. **Choose the method(s) of reaching target respondents :** Considering the information needed and respondents from whom the information is collected. The researcher needs to decide on a method of collecting information. There are different methods of collecting data such as Personal Interview, group interview, mailed questionnaire etc. The method of collecting data plays an important role in designing questionnaire.
4. **Decide on question content :** The researcher should carefully decide on contents of questionnaire. He need to be vigilant on contribution of question towards achievement of research objectives and its use in hypothesis testing. The researcher should avoid all irrelevant questions the he may be tempted to include in the questionnaire. The researcher also need to decide on type of questions to be used for collecting information for eg. Open Ended Questions, Closed Ended Questions, Pictorial Questions etc.
5. **Develop the question wording :** After deciding questions to be used in questionnaire the researcher need to focus on framing the questions using appropriate language. The researcher need to be careful while framing the questions, he must ensure simple and easy to understand words are used in the questionnaire. He should draft question carefully to avoid any ambiguity, confusion and care also need to be taken that question does not offend respondents.
6. **Put questions into a meaningful order and format :** The next step in designing a questionnaire is to decide the logical sequence of the questions to be included in the questionnaire. Deciding a sequence of questions is important as it helps in collecting accurate information and avoiding confusion of the respondents while answering the questions.

7. **Check the length of the questionnaire :** Once the logical sequence is decided the researcher needs to check the length of questionnaire. It means he needs to check no of questions asked in the questionnaire. A good questionnaire tries to obtain maximum information in minimum possible questions. If required, researchers will edit the questions if required and finalise the draft of the questionnaire.

8. **Pre-test the questionnaire :** The researcher must undertake a pilot study before starting collecting data for the study. He should evaluate problems faced by the interviewer and respondents while answering the questions asked in the questionnaire. He also needs to look into information collected from the questionnaire. This step is really important as it checks the usefulness of the questionnaire in the research.

9. **Develop the final survey form :** After conducting a pilot survey the researcher may edit the questionnaire. This is necessary to obtain reliable and accurate data which decides success of the research work. Once final editing is done the researcher prepares the final draft of the questionnaire which will be used for collecting data for the research work.

4.5.3. ESSENTIALS OF GOOD QUESTIONNAIRE

To draft a questionnaire or schedule is an art. The success of statistical investigation depends on proper drafting of the questionnaire. It is a highly specialized job and following points should be borne in mind:

1. Brief and Limited Questionnaire: The number of questions in a questionnaire should be brief and limited as possible. Only relevant questions to the problem under investigation should be included in the questionnaire. This is important because respondent may run out of patience while answering lengthy questionnaire and may not answer all questions.

2. Simple and Clear: The questions asked in the questionnaire should be simple, clear and precise. Its language used for the questions should be easy so that informants may easily understand and respond to the question. If respondent does not understand the question he may skip the question or may provide irrelevant and wrong information

3. Unambiguous Questions : The questionnaire should not contain any ambiguous question. If any question creates ambiguity for the respondents it will irritate him and again he may provide wrong or irrelevant information while answering such question

4. No Personal Questions: The researcher while designing questionnaire should refrain from asking personal questions to the respondents. The respondents may avoid answering such questions.

5. Avoidance of Calculations: Questions should not be based on calculations. Only those questions should be asked which the respondents may reply immediately. Moreover, questions should avoid memories.

6. Sequence of the Questions: The questions in a questionnaire should have a logical sequence A questionnaire with logically arranged questions encourage respondents to provide accurate information to the researcher.

7. Pre-testing: The researcher must conduct a pilot study before sending the questionnaire to the respondents. Such pretesting is needed to ensure questionnaire is properly designed and useful in collecting required information.

8. Instructions: The questionnaire must provide instruction to the respondents in respect of How to fill the questionnaire. Such instructions should be given as a footnote in the questionnaire.

9. Cross Examination: The questionnaire should be set in such a way that there may be cross examination of the information supplied by the informants. In fact, it is a check on false or inaccurate answers.

10. Attractive Questionnaire: Apart from the quality of questions in the questionnaire the researcher needs to take care of the physical appearance of the questionnaire. A questionnaire with good physical appearance encourage respondent to answer the questions truthfully.

4.6. SUMMARY

This unit had given in depth information on Secondary Data. It also discusses various factors which affect the decision of choosing a method of data collection. It provide detailed information in respect of questionnaire, its types, questions asked in questionnaire etc. Both units of Research process collective provide basic information to learner regarding Research Process, Data collection, various methods of data collection, its advantages and disadvantages etc

4.7. EXERCISE

1. Explain sources of Secondary Data
2. What are factors influencing choosing method of data collection
3. Explain steps in designing questionnaire
4. Explain different types of questionnaire
5. Explain essentials of good questionnaire

State following statements are True or False

1. Questionnaire is used to collect Secondary Data
2. Generally lengthy questionnaires are advisable for collecting data.
3. Tabulation of data helps in analysis and interpretation of data

4. Internal sources of secondary data collection include Government Statistics.

5. Collection of Secondary Data is time consuming than primary data

(True - 3 False - 1,2,4,5)

4.8 REFERENCE

- Research Methodology : Michale Vaz, Aurora Vaz : Manan Prakashan
- Research Methodology Methods and Techniques : C. R. Kothari, New Age International Publishers
- Research Methodology : R Panneerselvam, PHL Learning

❖❖❖❖

Module III
DATA PROCESSING AND STATISTICAL ANALYSIS

5

DATA PROCESSING

Unit Structure

5.0. OBJECTIVES :

1. To know the concept of data processing
2. To explain the different statistical methods in research analysis
3. To understand the concept of hypothesis testing and various statistical test
4. To understand the significance and precautions in data analysis

5.1. INTRODUCTION

Quantitative information may be found almost everywhere. But not all numerical data is statistical so it is necessary to examine a few definitions of statistics and to understand the features of statistical data. One of the most important objectives of statistical analysis is to get one single value that describes the characteristics of the entire mass of unwieldy data. Such a value is called central value.

5.2. METHODS OF DATA PROCESSING IN RESEARCH

5.2.1. Data Processing :

Data processing is a set of methods that are used to input, retrieve, verify, store, organize, analyse or interpret a set of data. Data processing enables information to be automatically extracted from data. Data processing starts with data in its raw form and converts it into a more readable format (graphs, documents, etc.), giving it the form and context

necessary to be interpreted by computers and utilized by employees throughout an organization.

The essence of data processing in research is data reduction. Data reduction involves sorting out the irrelevant from the relevant data and establishing order from chaos and giving shape to a mass of data.
Data processing in **research** consists of five important steps. They are following:

5.2.2. Editing of Data :

Editing is the first step in **data processing**. Editing is the process of examining the data collected in questionnaires/schedules to detect errors and omissions and to see that they are corrected and the schedules are ready for tabulation. When the whole data collection is over a final and a systematic check up is made.

Mildred B. Parten in his book points out that the editor is responsible for seeing that the data are;

i. Accurate as possible
ii. Consistent with other facts secured
iii. Uniformly entered
iv. As complete as possible
v. Acceptable for tabulation and arranged to facilitate coding tabulation.

5.2.3. Different types of editing.

i. **Editing for quality** asks the following questions:
    - Are the data forms complete?
    - Are the data free of bias?
    - Are the recordings free of errors?
    - Are the inconsistencies in responses within limits?
    - Are there evidences to show dishonesty of enumerators or interviewers?
    - Are there any excessive manipulation of data? Etc.

i. **Editing for tabulation** does certain accepted modification to data or even rejecting certain pieces of data in order to facilitate tabulation. Orfor instance, extremely high or low value data item may be ignored or bracketed with suitable class interval.

ii. **Field Editing** is done by the respondent. The schedule filled up by the respondent might have some abbreviated writings, illegible writings etc. These are rectified by the enumerator. This should be done soon after the enumeration or interview.

iii. **Central Editing** is done by the researcher after getting all schedules or questionnaires or forms from the respondents. Obvious errors can be corrected. For missed data or information, the editor may substitute data or information by reviewing information provided by likely placed other respondents. A definite inappropriate answer is removed and "no answer" is entered when reasonable attempts to get the appropriate answer fail to produce results.

5.2.4. Coding of Data :

Meaning

Coding is necessary for efficient analysis and through it the several replies may be reduced to a small number of classes which contain the critical information required for analysis.

Coding decisions should usually be taken at the designing stage of the questionnaire. This makes it possible to pre-code the questionnaire choices and which in turn is helpful for computer tabulation as one can straight forward key punch from the original questionnaires. But in case of hand coding some standard method may be used. One such standard method is to code in the margin with a colored pencil. The other method can be to transcribe the data from the questionnaire to a coding sheet. Whatever method is adopted, one should see that coding errors are altogether eliminated or reduced to the minimum level.

Coding is the process/operation by which data/responses are organized into classes/categories and numerals or other symbols are given to each item according to the class in which it falls. In other words, coding involves two important operations;

i. deciding the categories to be used and
ii. allocating individual answers to them.

These categories should be appropriate to the research problem, exhaustive of the data, mutually exclusive and unidirectional. Since the coding eliminates much of information in the raw data, it is important that researchers design category sets carefully in order to utilize the available data more fully.

In the case of pressing – coded questions, coding begins at the preparation of interview schedules. Secondly, coding frame is developed by listing the possible answers to each question and assigning code numbers or symbols to each of them which are the indicators used for coding. The coding frame is an outline of what is coded and how it is to be coded. That is, coding frame is a set of explicit rules and conventions that are used to base classification of observations variable into values which are which are transformed into numbers.

After preparing the sample frame the gradual process of fitting the answers to the questions must be begun. Lastly, transcription is undertaken

i.e., transferring of the information from the schedules to a separate sheet called transcription sheet. Transcription sheet is a large summary sheet which contain the answer/codes of all the respondents. Transcription may not be necessary when only simple tables are required and the number of respondents are few.

5.2.5. Classification of Data :

Classification or categorization of data is the process of grouping the statistical data under various understandable homogeneous groups for the purpose of convenient interpretation. A uniformity of attributes is the basic criterion for classification and the grouping of data is made according to similarity. Classification becomes necessary when there is diversity in the data collected for meaningless presentation and analysis. However, it is meaningless in respect of homogeneous data. A good classification should have the characteristics of clarity, homogeneity, equality of scale, purposefulness and accuracy.

5.2.6. Objectives of Classification of data are below:

i. The complex and scattered data is organized into logical and intelligible form.

ii. It is possible to make the characteristics of similarities and dis – similarities clear.

iii. Comparative studies is possible.

iv. Understanding of the significance is made easier and thereby good deal of human energy is saved.

v. Underlying unity amongst different items is made clear and expressed.

vi. Data is so arranged that analysis and generalization becomes possible.

5.2.7. Types of Classification of data :

Classification of data is of two types

ii. **Quantitative classification :** It is on the basis of variables or quantity

iii. **Qualitative classification :** It is classification according to attributes.

The quantitative classification is the way of grouping the variables, say, quantifying the variables in cohesive groups, while the qualitative groups the data on the basis of attributes or qualities.

Again, it may be multiple classification or dichotomous classification. The former is the way of making many (more than two) groups on the basis of some quality or attributes while the latter is the classification into two groups on the basis of presence or absence of a certain quality. Grouping the workers of a factory under various income (class intervals) groups come under the multiple classification and making

two groups into skilled workers and unskilled workers is the dichotomous classification. The tabular form of such classification is known as statistical series, which may be inclusive or exclusive.

5.2.8. Tabulation of Data :

Tabulation is the process of summarizing raw data and displaying it in compact form for further analysis. Therefore, preparing tables is a very important step. Tabulation may be by hand, mechanical, or electronic. The choice is made largely on the basis of the size and type of study, alternative costs, time pressures, and the availability of computers, and computer programmes. If the number of questionnaire is small, and their length short, hand tabulation is quite satisfactory.

Table may be divided into: (i) Frequency tables (ii) Response tables (iii) Contingency tables (iv) Uni-variate tables (v) Bi-variate tables (vi) Statistical table and (vii) Time series tables.

Necessary steps in the preparation of table:

i. **Title of table:** The table should be first given a brief, simple and clear title which may express the basis of classification.

ii. **Columns and rows:** Each table should be prepared in just adequate number of columns and rows.

iii. **Captions and stubs:** The columns and rows should be given simple and clear captions and stubs.

iv. **Ruling:** Columns and rows should be divided by means of thin or thick rulings.

v. **Arrangement of items :** Comparable figures should be arranged side by side.

vi. **Deviations:** These should be arranged in the column near the original data so that their presence may easily be noted.

vii. **Size of columns:** This should be according to the requirement.

viii. **Special emphasis:** This can be done by writing important data in bold or special letters.

ix. **Foot – notes:** These may be given below the table.

x. **Total:** Totals of each column and grand total should be in one line.

xi. **Source** : Source of data must be given. For primary data, write primary data.

It is always necessary to present facts in tabular form if they can be presented more simply in the body of the text. Tabular presentation enables the reader to follow quickly than textual presentation. A table should not merely repeat information covered in the text. The same information should not, of course be presented in tabular form and graphical form. Smaller and simpler tables may be presented in the text

while the large and complex table may be placed at the end of the chapter or report.

5.2.9. Data Diagrams :

Diagrams are charts and graphs used to present data. These facilitate getting the attention of the reader more. These help presenting data more effectively. Creative presentation of data is possible.

The data diagrams classified into as follows :

i. **Charts:** A chart is a diagrammatic form of data presentation. Bar charts, rectangles, squares and circles can be used to present data. Bar charts are uni-dimensional, while rectangular, squares and circles are two-dimensional.

ii. **Graphs:** The method of presenting numerical data in visual form is called graph, A graph gives relationship between two variables by means of either a curve or a straight line.
Graphs may be divided into two categories.
   a. Graphs of Time Series and
   b. Graphs of Frequency Distribution.

5.2.10 Significance of Data Processing in Research:

- Data processing helps to make reports easy because the data has been processed, it can be used directly. These processed data can be organized in such a way that they can help to conduct analysis quickly. Predefined data helps experts in making conclusions faster.

- Data processing maintains accuracy and speed in data analysis. Complex data can be processed in a minute and can store necessary data. In the processing of research data, the system will automatically check and process invalid data. Therefore, such processes help researchers to ensure high accuracy in data management.

- Data processing is very helpful in breaking a macro problem into a micro problem. It helps to detect errors and omissions.

- Data processing acts like a filter when it comes to acquiring meaningful insights out of a huge data set.

- Data processing helps in keeping human bias away from the research conclusion with the help of proper statistical treatment.

- Data processing further helps in cost reduction, ease in storage, distributing and report making followed by better analysis and presentation are other advantages.

5.3. STATISTICAL ANALYSIS: TOOLS AND TECHNIQUE

Masseurs of Central Tendency:

5.3.1. Definition:

A measure of central tendency is a representative number that characterizes the "middleness" of an entire set of data. The three measures of central tendency are the mean, the median, and the mode.

The term central tendency or average have been defined by various researchers in following ways

Simpson and Kafka " A measure of central tendency is a typical value around which other figures congregate"

Clark "Average is an attempt to find one single figure do describe whole of figures"

It is clear from the above definitions that an average is a single value which represents a whole series and is supposed to contain its major characteristics.

5.3.2. Types of Averages :

Measure of central tendency or averages are usually of the following types :

1) Mathematical Averages :

I. Arithmetic Average or mean
II. Geometric mean
III. Harmonic mean

2) Averages of position :

II. Median
III. Mode

Of the above mentioned five important averages, Arithmetic mean, median and mode are the most popular ones. Besides these there are other averages like Quadratic mean, moving average and progressive average etc. Moving and progressive averages are used in the analysis of commercial statistics. Quadratic mean not so used in analysis.

5.3.3. Arithmetic Mean :

The most widely used measure of representing the entire data by one value is generally called an average and what the statistician call is arithmetic mean. Its value is obtained by adding together all the items and by dividing this total by the number of items. Arithmetic means may either be :

i. Simple arithmetic mean
ii. Weighted arithmetic mean

i. Calculation of Simple Arithmetic mean :

The process of computing mean is case of individual series. At first add together the various values of the variable and divide the total by the number of items. Symbolically,

$$\mu = \frac{X1 + X2 + \cdots Xn}{N}$$

μ (pronounced "mu") = the symbol for the population mean i.e. Arithmetic mean ;

X1+X2+…….Xn = Value of variable
N = Number of observations

Short cut Method :

Arithmetic mean cab be calculated by using an arbitrary origin when deviations are taken from an arbitrary origin. The formula for calculating mean is

$$\mu = A + \frac{\Sigma d}{N}$$

A = assumed mean
d = deviations values form assumed mean

Calculation of mean Discrete series :

In discrete series mean (μ) mat be computed by applying direct method of shortcut method.

Direct Method :

The formula for computing mean is $\mu = \frac{\Sigma fx}{N}$

f = frequency
x = variables in questions
N = total number of items

Short Cut method :

The formula for computing mean is $\mu = A + \frac{\Sigma fd}{N}$

Calculation of Arithmetic mean: Continuous Series

In continuous series arithmetic mean may be computed by applying any of the following methods.

Direct Method : $\mu = \frac{\Sigma fm}{N}$

m = mid value of class
f = the frequency of each class

Shortcut method : $\mu = A + \frac{\Sigma fd}{N} \times C$

A= Assumed mean

d = deviations of mid points from assumed mean

N = total number of observation

5.3.4. Merits and limitations of Arithmetic mean :

Merits :

i. It is the simplest average to understand and to compute
ii. It is affected by the value of every item in the series
iii. It is suitable for further mathematical treatment
iv. It is least affected by the fluctuations of sampling

5.3.5. Limitations :

i. Sometime the abnormal items may considerably affect the average value when the number of items is not large.

ii. In a distribution with open and classes the value of mean cannot be computed without making assumption regarding the size of the class interval of the open end class.

5.3.6. Median

Meaning

Another measure of central tendency, the median, is used in situations in which the mean might not be representative of a distribution. The median by definition refers to the middle value in a distribution.

Yule and Kendall "the median may be defined as the middle most or central value of the variable when the values are arranged in ascending or descending order. In case of a frequency curve the median may be defined as that value of the variable which divides the area of the curve into two equal parts"

The median is the middle score in a distribution after the scores have been arranged from highest to lowest or lowest to highest. The point to remember is that the median is not affected by extreme scores in a distribution because it is only a positional value. The mean is affected by extreme scores because its value is determined by a calculation that has to include the extreme values.

Median is called as a positional average. The terms 'position' refers to the place of a value in a series. The median is calculated by averaging the two middle scores. In other words, we determine the middle point between the two middle scores. Median is thus the central value of the distribution or the value that divides the distribution into two parts.

Calculation of Median :

The calculation of median involves two basic steps.

i. The location of the middle item an d
ii. Finding out its value

The middle item in series of individual observation and also in a discrete series is

$$M = \frac{N+1}{2}$$

Where N is the total number of observations.

In case of continues series $(N/2)^{th}$ item is the middle item of the series.

Once the middle item is located its value has to be found out. In a series of individual observation if the total number of items is an odd figure, the value of the middle item is the median value. If N is even, median is half the sum of the two middle values.

Problem :

Find out the median of the following items
7,9,15,17,22,25,29,35,40

Solution :

These items would first be arranged in ascending order of magnitude the series then would be as follows

Sr. No	Size of items
1	7
2	9
3	15
4	17
5	22
6	25
7	29
8	35
9	40

$$M = \left[\frac{N+1}{2}\right]^{th}\text{item}$$

$$M = \left[\frac{9+1}{2}\right]^{th}\text{item}$$

M= 5^{th} item
Thus M = 22

Computation of Median in a Discrete series:

The various steps in the computation of median in a discrete series are as follows:

i. Arrange the values in ascending or descending order of magnitude
ii. Find out the cumulative frequencies

iii. Find out the diddle item by the formula $M = \left[\frac{N+1}{2}\right]_{th}$ item

iv. Now find out the value of $M = \left[\frac{N+1}{2}\right]_{th}$item. It can be found by first locating the cumulative frequency which is equal to $\frac{N+1}{2}$ or next higher to it and then determining the value corresponding to it. This will be the value of the median.

Solution :

Find out then value of median from the following data

Marks: 4, 6,8,10,12

Students : 5,1,4,2,3

Marks in ascending order (X)	Students (Frequency)	Cumulative frequency (C.F)
4	1	1
6	4	5
8	5	10
10	3	13
12	2	15
	N = 15	

Median is the value of $M = \left[\frac{N+1}{2}\right]_{th}$item

$$M = \left[\frac{15+1}{2}\right]_{th}\text{item}$$

$$M = 7.5\ ^{th}\text{item}$$

7.5^{th} item comes under in the group of 10 Cumulative frequency.

Thus the median value would be 8

Computation of Median in a Continuous series :

While computing the value of median in a continuous series the middle item is the $(N/2)^{th}$ item and not $M = \left[\frac{N+1}{2}\right]_{th}$ item

In a continues frequency distribution the value of median would be in class interval. To set a precise value of median we assume that the frequency of the median class is uniformly spread over the whole class interval. On this assumption the value of the median can be located by the following formula.

$$M = l1 + \frac{\frac{N}{2} - c.f.}{f}.i$$

Where

M = the value of median

l_1 = the lower limit of the class in which median lies
N/2 = the middle number
c.f. = the cumulative frequency of the median class
i = the magnitude of the median class interval

Problem and solution :

X	Frequency	c.f.
0-10	15	15
10-20	17	32
20-30	19	51
30-40	27	78
40-50	19	97
50-60	13	110
	N = 10	

Median value of N/2 is 55^{th} item which is lies in the 30-40 class intervals.\now applying the formula of interpolation

$$M = l1 + \frac{\frac{N}{2} - c.f.}{f} .i$$

$$M = 30 + \frac{\frac{110}{2} - 51}{27} \times 10$$

$$M = 30 + \frac{55 - 51}{27} \times 10$$

$$M = 30 + \frac{4}{27} \times 10$$

$$M = 30 + \frac{40}{27}$$

$$M = 30 + 1.48$$

$$M = 31.48$$

5.3.7. Merits of Arithmetic median :

i. It is not affected by the values of the extreme items and as such is sometimes more representative than arithmetic average.

ii. Even if the value of extremes is not known median can be calculated if the number of items is known

iii. It can be located merely by inspection in many cases.

5.3.8. Limitations of Arithmetic median:

i. For calculating median it is necessary to arrange the data other average do not need any arrangement

ii. Since it is a positional average, its value is not determined by each and every observations

iii. The value of median is affected more by sampling fluctuations than the value if arithmetic mean

5.3.9. Mode.

The third measure of central tendency is the mode—the score in a distribution that occurs with the greatest frequency. The mode is the only indicator of central tendency that can be used with nominal data. Although it can also be used with ordinal, interval, or ratio data, the mean and median are more reliable indicators of the central tendency of a distribution, and the mode is seldom used.

The mode refers to that value in a distribution which occurs most frequently. It is an actual value which has the highest concentration of items in and around it. The value of the variable at which the curve reaches a maximum is called the mode.

Calculation of Mode :

i. Individual Observations :

For ungrouped data or series of individual observations mode is often found by mere inspection.

Example : 30,31,33,34,33,34,37,33,35,38,
Since 33 number appeared three times
Thus Mode = 33

ii. Discrete series :

In discrete series very often mode can be located merely by inspection as the value having the highest frequency would generally be the modal value.

Example :

Value	7	9	10	12	15
Frequency	3	15	12	9	10

Clearly the value of mode is 9 as it carries the maximum frequency of 15. However on all cases the maximum frequency may not necessarily signifies maximum frequency density.

Calculation of mode: Continuous Series

i. By preparing grouping table and analysis table or by inspection ascertain the modal class

ii. Determine the value of mode by applying formula

$$Mo = l1 + \frac{f1 - fo}{2f1 - fo - f2} .i$$

Where

l_1 = lower limit of modal class

f_1 = frequency of the modal class

fo = frequency of the modal class preceding the modal class

f_2 = frequency of the modal class succeeding the modal class

i = Size of the class interval of modal class

Example:

X	20-30	30-40	40-50	50-60	60-70	70-80	80-90
F	8	20	25	22	10	8	7

The modal class is 40-50 since it have highest frequency i.e. 25

$$Mo = l1 + \frac{f1 - fo}{2f1 - fo - f2} .i$$

l_1 = 40 , f_1 = 25 , fo = 20 , f_2 = 22 , i = 10

Mode (Mo) = 46.25

5.4 MEASURES OF DISPERSION

5.4.1. Meaning and Definition

As we know that the various measures of central value gives us one single figure that represents the entire data. But the averages alone cannot adequately describes a set of observations unless all the observations are the same. It is necessary to describe the dispersion of the observations. Thus measures of dispersion help us in studying important feature of a distribution.

Brooks and Dick "Dispersion or spread is the degree of the scatter variation of the variables about a central value."

A measure of central tendency provides information about the "middleness" of a distribution of scores but not about the width or spread of the distribution. To assess the width of a distribution, we need a measure of variation or dispersion. A measure of variation indicates the degree to which scores are either clustered or spread out in a distribution.
As an illustration, consider the two very small distributions of exam scores shown in following table.

Class 1	Class 2
0	45
50	50
100	55
Σ = 150	Σ =150
$\mu = 50$	$\mu = 50$

Notice that the mean is the same for both distributions. If these data represented two very small classes of students, reporting that the two classes had the same mean on the exam might lead you to conclude that the classes performed essentially the same. Notice, however, how different the distributions are. Providing a measure of variation along with a measure of central tendency conveys the information that even though the distributions have the same mean, their spreads are very different.

We will discuss three measures of variation: the range, the mean deviation, and the standard deviation. The range can be used with ordinal, interval, or ratio data however, the standard deviation and average deviation are appropriate for only interval and ratio data.

A. Range :

The simplest measure of variation is the range—the difference between the lowest and the highest scores in a distribution. The range is usually reported with the mean of the distribution. To find the range, we simply subtract the lowest score from the highest score. In our hypothetical distributions of exam scores in above table, the range for Class 1 is 100 points, whereas the range for Class 2 is 10 points. Thus, the range provides some information concerning the difference in the spreads of the distributions. In this simple measure of variation, however, only the highest and lowest scores enter the calculation, and all other scores are ignored. Thus, the range is easily distorted by one unusually high or low score in a distribution.

B. Mean Deviation:

It is the more sophisticated measures of variation use all of the scores in the distribution in their calculation. The average which is frequently used in computing the mean deviation is mean or median.

Mean deviation denoted by Greek letter, small " δ ". The sign of average taken in deviation s used as subscript.

δ_{xor} = mean deviation form mean

δ_M= mean deviation from median

δ_{M0} = Mean deviation from mode

Coefficient of mean Deviation :

Mean deviation when is divided by the average used for calculating it we get coefficient of mean deviation.

$$Coefficent\ of\ M.D. = \frac{\mathbf{M.D}}{\mu \mathbf{\ or\ M\ or\ M0}}$$

Computation of Mean Deviation : Individual series

There are two methods of calculating the mean deviation from a series of individual observations.

i. Direct Method :

The word deviation means to diverge, move away from, or digress. In this method the mean deviation would be calculated by totaling the deviations and dividing the total by the number of items.

Mean Deviation $= (\Sigma |d|)/N$

In the short cut method mean or median is calculated and the total of the values of the items below the mean or median and above it are found out. The former is subtracted from the latter and is divided by the number of items. The resulting figure is the mean deviation.

Computation of Mean Deviation : Discrete series

In discrete series the calculation of mean deviation involve the following steps:

i. Calculating the median or mean of the series (M)
ii. Find the deviations from the median or mean ignoring plus minus signs $dm\ or\ dx$
iii. Multiply the deviations with the respective frequency and get the total $\sum f\,dm$
iv. Divide the total by the number of observations. This will be the value of mean deviation.

Symbolically :

$$\delta m = (\Sigma |f\,dm|)/N \text{ and}$$
$$\delta x = (\Sigma |f\,dm|)/N$$

C. Standard Deviation:

The most commonly used measure of variation is the standard deviation. In other words that might be substituted for the word standard include average, normal, and usual. The standard deviation means the average movement away from something. It is the average movement away from the center of the distribution—the mean. The standard deviation, then, is the average distance of all the scores in the distribution from the mean or central point of the distribution —or, the square root of the average squared deviation from the mean.

Standard deviation first suggested by Karl Person in 1893. It may be defined as “Root-Mean Square Deviation” from the mean. It is usually denoted by the Greek letter σ (sigma).

Calculation of standard deviation : individual observations

In case of individual observations standard deviation may be computed by applying any of the following two methods.

i. Deviation taken from the actual mean :

By Taking deviation of the items from the actual mean. In the method following formula is applied.

$$\sigma = \sqrt{\frac{\Sigma x^2}{N}}$$

Steps :

i. Calculating the actual mean of the series
ii. Tae the deviations of the items from the mean
iii. Square these deviations and obtain that total Σx^2
iv. Divide Σx^2 by the total number of observations i.e. N and extract square root. This gives us the value of S.D.

ii. Deviations taken from Assumed Mean :

When deviation are taken from assumed mean the following formula is applied.

$$\sigma = \sqrt{\frac{\Sigma x^2}{N} - \left(\frac{\Sigma d^2}{N}\right)}$$

Steps :

i. Take the deviations of the items from an assumed mean. Denote these deviations by d. take the total of these deviations i.e. obtain Σd^2
ii. Square these deviations and obtain the total Σd^2
iii. Substitute the values of $\Sigma d^2, \Sigma d^2$ and N in the above formula.

Calculation of S.D : Discrete Series

i. Assumed mean method

When this method is used the following formula is applied

$$= \sqrt{\frac{\Sigma f d^2}{N} - \left(\frac{\Sigma f d}{N}\right)^2}$$

Where d = (X-A)

Steps :

i. Take the deviations of the items from an assumed mean and denote these deviations by d
ii. Multiply these deviations by the respective frequencies and obtain the total $\Sigma f d^2$

iii. Multiply the squared deviations by the respective frequencies and obtain the total Σfd^2. Substitute the value sin the above formula.

Example :

X	F	d= (X-A)	fd	d^2	fd^2
10	8	-30	-240	900	7200
20	12	-20	-240	400	4800
30	20	-10	-200	100	2000
40	10	0	0	0	0
50	7	10	70	100	700
60	3	20	60	400	1200
	N=60		$\Sigma fd^2 = 550$		$\Sigma fd^2 = 15900$

$$= \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$

$$\sigma = \sqrt{\frac{15900}{60} - \left(\frac{550}{60}\right)^2}$$

$$\sigma = \sqrt{265 - 84.02}$$

$$\sigma = \sqrt{180.98}$$

$$\boldsymbol{\sigma = \sqrt{13.45}}$$

ii. Step Deviation method :

When this method is used we take a common factor from the given data. The formula for computing standard deviation is:

$$= \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \cdot C$$

Where $d = \frac{(X - A)}{C}$ and C= common factor

Calculation of standard deviation: Continuous series

In continues series any of the method discussed above for discrete frequency distribution can be used. However in practice it is the step deviation method that is mostly used.

$$= \sqrt{\frac{\Sigma f d^2}{N} - \left(\frac{\Sigma f d}{N}\right)^2} \cdot C$$

X	F	M.V	$d = \frac{(X - A)}{C}$	F × d	fd^2
0-10	18	5	-40	-720	28800
10-20	16	15	-30	-480	14400
720-30	15	25	-20	-300	6000
30-40	12	35	-10	-120	1200
40-50	10	45	0	0	0
50-60	05	55	10	50	500
60-70	02	65	20	40	800
70-80	1	75	30	30	900
	N = 79			$fd = 1500$	$\Sigma fd^2 = 52600$

$$= \sqrt{\frac{\Sigma f d^2}{N} - \left(\frac{\Sigma f d}{N}\right)^2} \cdot C$$

$$\sigma = \sqrt{\frac{52600}{79} - \left(\frac{1500}{79}\right)^2} \times 10$$

$$\sigma = \sqrt{665.82 - (18.98)2} \times 10$$

$$\sigma = \sqrt{665.82 - 360.5} \times 10$$

$$\sigma = \sqrt{305.32} \times 10$$

$$\sigma = 17.47 \times 10$$

$$\sigma = 174.73$$

5.5. SUMMARY :

The central tendency of a distribution is typically contrasted with its dispersion or variability; dispersion and central tendency are the often characterized properties of distributions. Analysis may judge whether data has a strong or a weak central tendency based on its dispersion. In statistics, a central tendency is a central or typical value for a probability distribution. It may also be called a center or location of the distribution.

Colloquially, measures of central tendency are often called averages. A measure of central tendency is a summary statistic that represents the center point or typical value of a dataset. These measures indicate where most values in a distribution fall and are also referred to as the central location of a distribution. You can think of it as the tendency of data to cluster around a middle value.

5.6. EXERCISE:

A. Test your knowledge by choosing the correct option:

Q1. ……. is a set of methods that are used to input, retrieve, verify, store, organize, analyse or interpret a set of data. (Data processing , Central Tendency, Parametric test, Non Parametric Statistical Tests)

Q2. …….. involves sorting out the irrelevant from the relevant data . (Data reduction, Data Processing, Data Collection, Data mining)

Q3……. is done by the respondent. **(Field Editing, Central Editing, Editing for tabulation, Editing for quality)**

Q4. **……..** of data is the process of grouping the statistical data under various understandable homogeneous groups for the purpose of convenient interpretation. (Coding, Editing, Tabulation, Classification)

Q5. The method of presenting numerical data in visual form is called……. (Editing,Graph, Chart, Coding)

Q6. A measure of ……tendency is a representative number that characterizes the "middleness" of an entire set of data. (Central, Positive, Negative, Rational)

Q7. …….averages is not used in the analysis of commercial statistics. (Moving, progressive, Quadratic mean, Arithmetic mean)

Q8. Which is the formula of Arithmetic mean

i. $\mu = \frac{X1 + X2 + \cdots Xn}{N}$ ii. $M = \frac{N+1}{2}$

iii. $Mo = l1 + \frac{f1 - fo}{2f1 - fo - f2} . i$ iv. M D $= (\Sigma \| d \|)/N$

Q9. The ……is the middle score in a distribution after the scores have been arranged from highest to lowest or lowest to highest. (Median, mean, mode, correlation)

Q10. The …….. refers to that value in a distribution which occurs most frequently. (Median, mean, mode, correlation)

Q11. The …… is the average distance of all the scores in the distribution from the mean or central point of the distribution (mean, mode, correlation, standard deviation)

Q12. ……is usually denoted by the Greek letter σ (sigma) (mean, mode, correlation, standard deviation)

(**Answers** – 1. Data processing, 2. Data processing, **3. Field Editing, 4.**Classification, 5. Graph, 6. Central, 7. Quadratic mean, 8. $\mu = \frac{X1 + X2 + \cdots Xn}{N}$ 9. Median, 10. Mode, 11. standard deviation, 12. standard deviation)

B. Short Answer Questions :

1. Define data processing
2. Explain the concept editing of data
3. What is coding of data?
4. Explain Tabulation of data.
5. Define Arithmetic mean
6. Merits and limitations of Arithmetic mean.
7. Merits and limitations of Arithmetic median.
8. What is mode?
9. Define the central tendency.
10. Explain the concept Range.

5.7 REFERENCE

I. Sherri L. Jackson, (2009) , Research Methods and Statistics, A Critical Thinking Approach, ISBN-13: 978-0-495-51001-7

II. BASIC ECONOMETRICS,(2003) Damodar N. Gujarati, *McGraw-Hill Higher Education* ISBN: 978-0-07-233542-2

III. Sachdeva J.K.(2011). Business Research Methodology. New Delhi: Himalaya Publishing House.

IV. Michael V.P.(1997). Research Methodology in Management. Delhi: Himalaya Publishing House

❖❖❖❖

6

STATISTICAL ANALYSIS

Unit Structure

6.0 OBJECTIVE:

After studying this unit learner will be in position;

- To examine the concept of correlation
- To analyse the regression work
- To understand the concept of statistical inferences
- To examine the Parametric and Non Parametric Statistical Tests
- To elaborate the Factor Analysis.

6.1. INTRODUCTION

In the Information Age, data is no longer scarce – it's uncontrollable. The key is to examine through the vast volume of data available and businesses and correctly interpret its implications. But to sort through all this information, we need the right statistical data analysis tools. With the current obsession over big data, analysts have produced a lot of tools and techniques for data interpretation. Statistics is a crucial process behind how we make decisions based on data and make predictions. Statistics tools allow researchers to understand a subject much more deeply.

6.2. CORRELATION

6.2.1. Meaning:

Correlation is a statistic that measures the degree to which two variables move in relation to each other. Correlation shows the strength of a relationship between two variables and is expressed numerically by the correlation coefficient. The correlation coefficient's values range between -1.0 and 1.0. A perfect positive correlation means that the correlation coefficient is exactly 1. A perfect negative correlation means that two variables move in opposite directions, while a zero correlation implies no linear relationship at all.

A correlation coefficient quite close to 0, but either positive or negative, implies little or no relationship between the two variables. A correlation coefficient close to plus 1 means a positive relationship between the two variables, with increases in one of the variables being associated with increases in the other variable.

A correlation coefficient close to -1 indicates a negative relationship between two variables, with an increase in one of the variables being associated with a decrease in the other variable. A correlation coefficient can be produced for ordinal, interval or ratio level variables, but has little meaning for variables which are measured on a scale which is no more than nominal.

The formula for correlation:

$$r = \frac{\sum(X - X)(Y - Y)}{\sqrt{\sum(X - \overline{X})^2}\sqrt{(Y - \overline{Y})^2}}$$

where:

r = Correlation coefficient

$\overline{X}$ = Average of observations of variable X

$\overline{Y}$ = Average of observations of variable Y

6.2.2. Types of correlation:

i. Positive correlation :

Positive correlation refers to the movement of the variable in the same direction or a direct relationship exists between the two variables. This means that an increase in one variable is related to an increase in the other, and a decrease in one is related to a decrease in the other. This type of correlation exist between supply and price of commodity.

ii. Negative correlation :

Negative or inverse correlation refers to when one variable increases or decreases the other moves in the reverse direction. Such a correlation is found between price and demand, when price of a commodity increases its demand goes down or vice-versa.

iii. Simple and multiple correlation :

Under simple correlation the relationship is confined to two variables like between the yield of wheat and the use of chemical fertilizers of between money supply and the general price level.

In case of multiple correlation the relationship between more than two variables is judged. For example the relationship of yield of wheat by judged with reference to say chemical fertilizers irrigation and pesticides.

iv. Partial and total correlation :

These are the two types of a multiple correlation analysis. Under partial correlation the relationship of two or more variable is examined excluding some other variables which are included for calculation of total correlation. \for example coefficient of correlation between yield of wheat and chemical fertilizers excluding the effect of pesticides and manures is called the partial correlation. And the total correlation is based on all the relevant variables.

v. Linear and non-linear correlation :

The distinction between linear and non-linear correlation is based upon the ratio of change between the variables under study. When variations in the values of two variables have constant ratio there will be linear correlation between them. The graph of variables having such a relationship will form a straight line. In non-linear or curve liner correlation the amount of change in the other related variable. For example when we double the use of fertilizers the production of jute would not necessarily be doubled.

6.2.3. Graphic methods of determining correlation:

The different methods of finding out correlations are following.

i. Scatter Diagram :

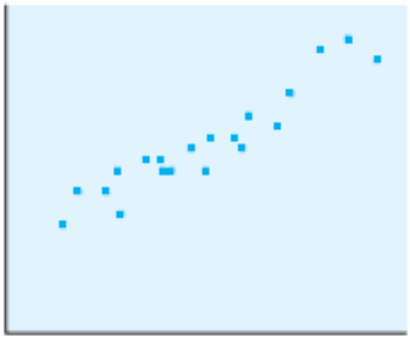
This method is a simple and attractive method of diagrammatic representation of a bivariate distribution for ascertaining the nature of correlation between the variables. Pairs of values (X_1, Y_1) (X_2, Y_2) ……(X_n, Y_n) or two variables X and Y can be plotted as dots (.) on the X-axis and Y axis in the XY- plane. It is customary to take the independent variable along the horizontal or X-axis and the dependent vertical along the vertical or Y axis, if at all there is called a scatter diagram.

If the patterns of points or dots on a scatter diagrams revels an upward or a downward trend the variables are said to be correlated and if the plotted points do not show any trend the two variables have no correlation.

The scatter diagrams may take the following shapes

1) Positive Correlation relationship :

It means that an increase in one variable is related to an increase in the other, and a decrease in one is related to a decrease in the other. The majority of the data points fall along an upward angle (from the lower-left corner to the upper-right corner).

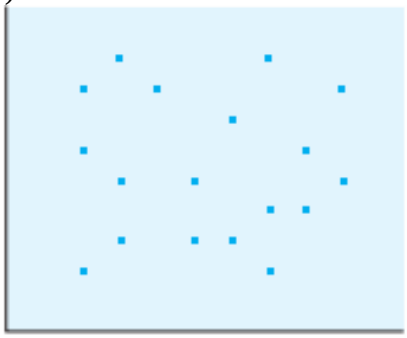


2) **Negative Correlation Relationship :**

In this scatter plot, the data points extend from the upper left to the lower right. This negative correlation indicates that an increase in one variable is accompanied by a decrease in the other variable. This represents an inverse relationship: The more of variable x that we have, the less we have of variable y. Assume that this scatter plots represents the relationship between age and eyesight. As age increases, the ability to see clearly tends to decrease—a negative relationship.

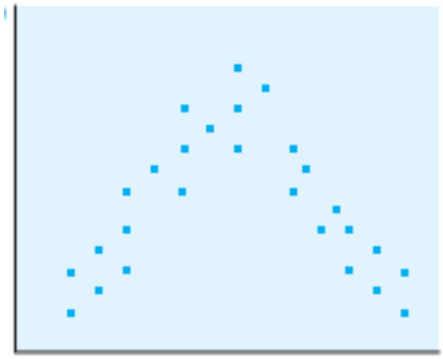
3) No correlation Relationship :

It is also possible to observe no meaningful relationship between two variables. In this scatter plot, the data points are scattered in a random fashion. As you would expect, the correlation coefficient for these data is very close to 0 to (.09).



4) Curvilinear Relationship :

A correlation coefficient of 0 indicates no meaningful relationship between two variables. However, it is also possible for a correlation coefficient of 0 to indicate a curvilinear relationship.



Imagine that above graph represents the relationship between psychological excitement (the x-axis) and performance (the y-axis). Individuals perform better when they are moderately stimulated than when stimulation is either very low or very high. The correlation coefficient for these data is also very close to 0 - (.05). The strong positive relationship depicted in the left half of the graph essentially cancels out the strong negative relationship in the right half of the graph. Although the correlation coefficient is very low, we would not conclude that no relationship exists between the two variables.

As the graph shows, the variables are very strongly related to each other in a curvilinear manner—the points are tightly clustered in an inverted U shape. Correlation coefficients tell us about only linear relationships. Thus, even though there is a strong relationship between the two variables graph, the correlation coefficient does not indicate this because the relationship is curvilinear. For this reason, it is important to examine a scatter plot of the data in addition to calculating a correlation coefficient. Alternative statistics can be used to assess the degree of curvilinear relationship between two variables.

6.3. REGRESSION ANALYSIS

Regression lines are the device used for estimating the value of one variable from the value of the other consists of a line through the points drawn in such a manner as to represent the average relationship between the two variables. Such a line is called the line of regression.

As per the method of least square the two regressions lines are

The general form of each type of regression is:

- **Simple linear regression:** $Y = a + bX + u$
- **Multiple linear regression:** $Y = a + b_1X_1 + b_2X_2 + b_3X_3 + ... + b_tX_t + u$

Where:

- Y = the variable that you are trying to predict (dependent variable).
- X = the variable that you are using to predict Y (independent variable).
- a = the intercept.
- b = the slope.
- u = the regression residual

Multiple regression analysis is used to see if there is a statistically significant relationship between sets of variables. It's used to find trends in those sets of data.

Multiple regression analysis is almost the same as simple linear regression. The only difference between simple linear regression and multiple regression is in the number of predictors ("x" variables) used in the regression.

Simple regression analysis uses a single x variable for each dependent "y" variable. For example: (x1, Y1).

Multiple regression uses multiple "x" variables for each independent variable: (x1)1, (x2)1, (x3)1, Y1).

In one-variable linear regression, you would input one dependent variable (i.e. "sales") against an independent variable (i.e. "profit"). But you might be interested in how different types of sales effect the regression. You could set your X1 as one type of sales, your X2 as another type of sales and so on.

6.3.1. Two lines of regression:

Since the regression relation is irreversible one equation is not sufficient to predict the values of two variables X and Y. moreover the two regression equations are derived under different set of assumptions, therefore one equation is not sufficient to find X and Y.

6.3.2. Properties of Regression coefficients :

i. Both the regression coefficients should be of the same sign.

ii. If both the regression coefficients are positive, correlation coefficient is positive and if both the regression coefficients are negative the correlation coefficient is negative.

iii. Both the regression coefficient are independent of the change of origin but if the change of scale in X and Y is not identical they depend on the change of scale. If the change of scale in X and Y is identical the regression coefficients are independent of the change of scale also.

iv. The correlation coefficient is the geometric mean between the regression coefficients.

6.4. HYPOTHESIS TESTING

6.4.1. Meaning:

Hypothesis testing is an act in statistics where by an analyst tests an assumption regarding a population parameter. The methodology employed by the analyst depends on the nature of the data used and the reason for the analysis. Hypothesis testing is used to assess the plausibility of a hypothesis by using sample data. Such data may come from a larger population, or from a data-generating process.

Hypothesis testing was introduced by Ronald Fisher, Jerzy Neyman, Karl Pearson and Pearson's son, Egon Pearson. Hypothesis testing is a statistical method that is used in making statistical decisions using experimental data. Hypothesis Testing is basically an assumption that we make about the population parameter.

6.4.2. Steps of Hypothesis Testing :

i. All hypotheses are tested using a four-step process:

ii. The first step is for the analyst to state the two hypotheses so that only one can be right.

iii. The next step is to formulate an analysis plan, which outlines how the data will be evaluated.

iv. The third step is to carry out the plan and physically analyze the sample data.

v. The fourth and final step is to analyze the results and either reject the null hypothesis, or state that the null hypothesis is plausible, given the data.

6.4.3. Null hypothesis:

Null hypothesis is a statistical hypothesis that assumes that the observation is due to a chance factor. A statistical hypothesis which is under test, usually a hypothesis of no difference and hence, it's called Null hypothesis. Null hypothesis is denoted by : H_0: $\mu 1 = \mu 2$, which shows that there is no difference between the two population means.

A.R. Fisher "null hypothesis is the hypothesis which is tested for possible rejection under the assumption that it is true."

6.4.4. Alternative hypothesis:

Contrary to the null hypothesis, the alternative hypothesis shows that observations are the result of a real effect. Rejecting null hypothesis implies that it is rejected in favour of some other hypothesis which is accepted. A hypothesis which is accepted when H_0 is rejected is called Alternative hypothesis and is represented by H_1 .

Usually the null hypothesis is expressed as an equality e.g. H_0 :theta $\theta = \theta_0$

6.4.5. Rules of stating Null and Alternative hypothesis :

i. The conclusion expected as a result of the test should be placed in the Alternative hypothesis
ii. The null hypothesis should contain a statement of equality either $\leq, \geq$ or =
iii. The null hypothesis is the hypothesis that is tested.
iv. The null and alternative hypothesis are complementary

6.4.6. Level of significance:

The level of significance is defined as the probability of rejecting Null hypothesis when it is true. It is the maximum size of the type I error which we are prepared to risk. It refers to the degree of significance in which we accept or reject the null-hypothesis. 100% accuracy is not possible for accepting or rejecting a hypothesis, so we therefore select a level of significance that is usually 5%. The higher the level of significance is, the higher the probability of rejecting a null hypothesis when it is true. Level of significance is always fixed in advance before the sample information.

6.4.7. Types of error :

When we test hypothesis there are four possible outcomes.

i. Null hypothesis is rejected when it is true
ii. Null hypothesis is rejected when it is false
iii. Null hypothesis accepted when it is false
iv. Null hypothesis is accepted when it is true.

Outcomes i and iii are undesirable. We may think of these two undesirable outcomes as incorrect actions and distinguish them by referring to them as Type I and Type II errors. The error of rejecting H_0 when H_0 is true is known as type I error and error of accepting H_0 when H_0 is false known as type II error.

If

P (rejecting H_0 when H_0 is true) = P (Type I error) = α

P (accepting H_0 when H_0 is false) = P (Type II error) = β

These are also called the sizes of type I error and type II error respectively. The sizes f type I and type II errors are also known as producer's risk and consumer's risk respectively.

Type I error: When we reject the null hypothesis, although that hypothesis was true. Type I error is denoted by alpha (α). In hypothesis testing, the normal curve that shows the critical region is called the alpha region.

Type II errors: When we accept the null hypothesis but it is false. Type II errors are denoted by beta (β). In Hypothesis testing, the normal curve that shows the acceptance region is called the beta region.

6.4.8. Power of test:

Usually known as the probability of correctly accepting the null hypothesis. The good test should accept the null hypothesis when it is true and reject the null hypothesis when it is false. 1- β is called power of the analysis or probability of type II error. It measures how well the test is working and is called the power of the test. A high value of 1- β implies the test is working quite well. A low value of 1- β impels the test is working very poorly.

A. One-tailed test:

When the given statistical hypothesis is one value like H0: $\mu 1 = \mu 2$, it is called the one-tailed test. For example a test for testing the man of a population

H_0: $\mu = \mu_0$

Against the alternative hypothesis.

H_1: $\mu > \mu_0$(right tailed) or H_1: $\mu < \mu_0$ (left tailed) is a single tailed test. In the right tailed test the critical region lies entirely in the right tail of the sampling distribution, while for the left tailed test the critical region is entirely in the left tail of the distribution of mean.

B. Two-tailed test:

When the given statistics hypothesis assumes a less than or greater than value, it is called the two-tailed test. A test of statistical hypothesis where the alternative hypothesis is two tailed such as

H_0: $\mu = \mu_0$

Against the alternative hypothesis

H_1: $\mu \neq \mu_0$ is known as two tailed test and in such a case the critical region is given by the portion of the area lying in both the tails of the probability curve of the test statistic.

6.5. PARAMETRIC AND NON PARAMETRIC STATISTICAL TESTS :

Meaning:

A parameter in statistics refers to an aspect of a population, as opposed to a statistic, which refers to an aspect about a sample. For example, the population mean is a parameter, while the sample mean is a statistic. A parametric statistical test makes an assumption about the population parameters and the distributions that the data came from.

The parametric test make certain assumptions about a data set namely – that the data are drawn from a population with a specific or normal distribution. It is further assumed in parametric test that the variables in the population are measured based on an interval scale.

Parametric tests are used when the data has a normal distribution and when the measurement scale is interval or ratio

6.5.1. Types of Parametric test :

- Two-sample t-test
- Paired t-test
- Analysis of variance (ANOVA)
- Pearson coefficient of correlation

I. t-Test :

A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. It is mostly used when the data sets, like the data set recorded as the outcome from flipping a coin 100 times, would follow a normal distribution and may have unknown variances. A t-test is used as a hypothesis testing tool, which allows testing of an assumption applicable to a population.

A t-test looks at the t-statistic, the t-distribution values, and the degrees of freedom to determine the statistical significance. To conduct a test with three or more means, one must use an analysis of variance.

A t-test allows us to compare the average values of the two data sets and determine if they came from the same population. Mathematically, the t-test takes a sample from each of the two sets and establishes the problem statement by assuming a null hypothesis that the two means are equal. Based on the applicable formulas, certain values are calculated and compared against the standard values, and the assumed null hypothesis is accepted or rejected accordingly.

If the null hypothesis qualifies to be rejected, it indicates that data readings are strong and are probably not due to chance. The t-test is just one of many tests used for this purpose. Statisticians must additionally use tests other than the t-test to examine more variables and tests with larger sample sizes.

t-Test Assumptions

i. The first assumption made regarding t-tests concerns the scale of measurement. The assumption for a t-test is that the scale of measurement applied to the data collected follows a continuous or ordinal scale, such as the scores for an IQ test.

ii. The second assumption made is that of a simple random sample, that the data is collected from a representative, randomly selected portion of the total population.

iii. The third assumption is the data, when plotted, results in a normal distribution, bell-shaped distribution curve.

iv. The final assumption is the homogeneity of variance. Homogeneous, or equal, variance exists when the standard deviations of samples are approximately equal.

Calculating t-Tests

Calculating a t-test requires three key data values. They include the difference between the mean values from each data set (called the mean difference), the standard deviation of each group, and the number of data values of each group.

The outcome of the t-test produces the t-value. This calculated t-value is then compared against a value obtained from a critical value table (called the T-Distribution Table). This comparison helps to determine the effect of chance alone on the difference, and whether the difference is outside that chance range. The t-test questions whether the difference between the groups represents a true difference in the study or if it is possibly a meaningless random difference.

t-Distribution Tables

The T-Distribution Table is available in one-tail and two-tails formats. The one-tail is used for assessing cases which have a fixed value or range with a clear direction (positive or negative). The two-tails is used for range bound analysis.

T-Values and Degrees of Freedom

The t-test produces two values as its output : t-value and degrees of freedom. The t-value is a ratio of the difference between the mean of the two sample sets and the variation that exists within the sample sets. While the numerator value (the difference between the mean of the two sample sets) is straightforward to calculate, the denominator (the variation that exists within the sample sets) can become a bit complicated depending upon the type of data values involved. The denominator of the ratio is a measurement of the dispersion or variability. Higher values of the t-value, also called t-score, indicate that a large difference exists between the two sample sets. The smaller the t-value, the more similarity exists between the two sample sets.

- A large t-score indicates that the groups are different.
- A small t-score indicates that the groups are similar.

Degrees of freedom refers to the values in a study that has the freedom to vary and are essential for assessing the importance and the validity of the null hypothesis. Computation of these values usually depends upon the number of data records available in the sample set.

Correlated (or Paired) T-Test

The correlated t-test is performed when the samples typically consist of matched pairs of similar units, or when there are cases of repeated measures. This method also applies to cases where the samples are related in some manner or have matching characteristics, like a comparative analysis involving children, parents or siblings. Correlated or paired t-tests are of a dependent type, as these involve cases where the two sets of samples are related.

The formula for computing the t-value and degrees of freedom for a paired t-test is:

$$T = \frac{mean1 - mean2}{\frac{s(\text{diff})}{\sqrt{(n)}}}$$

Where:
Mean 1 and mean 2 = the average values of each of the sample sets
s (diff) = The standard deviation of the differences of the paired data values
n = The sample size (the number of paired differences)
n−1= The degrees of freedom

Equal Variance (or Pooled) T-Test

The equal variance t-test is used when the number of samples in each group is the same, or the variance of the two data sets is similar. The following formula is used for calculating t-value and degrees of freedom for equal variance t-test:

$$\text{T-value} = \frac{mean1 - mean2}{\frac{(n1-1)\times var1^2+(n2-1)\times var2^2}{n1+n2-2} \times \sqrt{\frac{1}{n1} + \frac{1}{n2}}}$$

Where:
Mean 1 and mean 2 = Average values of each of the sample sets
var 1 and var 2 = Variance of each of the sample sets
n1and n2 = Number of records in each sample set
And
Degrees of Freedom= $n1+n2-2$
where:$n1$ and $n2$ = Number of records in each sample set

Unequal Variance T-Test :

The unequal variance t-test is used when the number of samples in each group is different, and the variance of the two data sets is also different. This test is also called the Welch's t-test. The following formula is used for calculating t-value and degrees of freedom for an unequal variance t-test:

$$\text{T-value} = \frac{mean1 - mean2}{\frac{var1^2}{n1} + \frac{var2^2}{n2}}$$

Where:
*mean*1 and *mean*2 = Average values of each of the sample sets
*var*1 and *var*2 = Variance of each of the sample sets
*n*1 and *n*2 = Number of records in each sample set

And

$$\text{Degrees of Freedom} = \frac{\left(\frac{var1^2}{n1} + \frac{var2^2}{n2}\right)^2}{\frac{\left(\frac{var1^2}{n1}\right)^2}{n1-1} + \frac{\left(\frac{var2^2}{n2}\right)^2}{n2-1}}$$

Where:
*var*1 and *var*2 = Variance of each of the sample sets
*n*1 and *n*2=Number of records in each sample set

II. F- test :

An "F Test" is a catch-all term for any test that uses the F-distribution. The f-statistic is used in a variety of tests including regression analysis, the Chow test and the Scheffe Test (a post-hoc ANOVA test).

General Steps for an F Test :

- State the null hypothesis and the alternate hypothesis.
- Calculate the F value.

The F Value is calculated using the formula F = (SSE1 – SSE2 / m) / SSE2 / n-k,

Where, SSE = residual sum of squares, m = number of restrictions and k = number of independent variables.

- Find the F Statistic (the critical value for this test).

The F statistic formula is:

F Statistic = variance of the group means / mean of the within group variances.

You can find the F Statistic in the F-Table.

- Support or Reject the Null Hypothesis.

F Test to Compare Two Variances :

A Statistical F Test uses an F Statistic to compare two variances, s1 and s2, by dividing them. The result is always a positive number (because variances are always positive). The equation for comparing two variances with the f-test is:
$F = s^2_1 / s^2_2$

If the variances are equal, the ratio of the variances will equal 1. For example, if we had two data sets with a sample 1 (variance of 10) and a sample 2 (variance of 10), the ratio would be 10/10 = 1.

We always test that the population variances are equal when running an F Test. In other words, always have to assume that the variances are equal to 1. Therefore, null hypothesis will always be that the variances are equal.

Assumptions :

Several assumptions are made for the test.

- Population must be approximately normally distributed (i.e. fit the shape of a bell curve) in order to use the test.

- The samples must be independent events.
- The larger variance should always go in the numerator (the top number) to force the test into a right-tailed test. Right-tailed tests are easier to calculate.
- For two-tailed tests, divide alpha by 2 before finding the right critical value.
- Standard deviations, must be squared to get the variances.
- If degrees of freedom aren't listed in the F Table, use the larger critical value. This helps to avoid the possibility of Type I errors.

III. Z Test

A Z-test is a type of hypothesis test—a way to figure out if results from a test are valid or repeatable. For example, if someone said they had found a new drug that cures cancer, you would want to be sure it was probably true. A hypothesis test will tell you if it's probably true, or probably not true. A Z test, is used when data is approximately normally distributed (i.e. the data has the shape of a bell curve when you graph it).

When can run a Z Test :

- Sample size is greater than 30. Otherwise, use a t test.
- Data points should be independent from each other. In other words, one data point isn't related or doesn't affect another data point.
- Data should be normally distributed. However, for large sample sizes (over 30) this doesn't always matter.
- Data should be randomly selected from a population, where each item has an equal chance of being selected.
- Sample sizes should be equal if at all possible.

Steps to run Z test :

i. State the null hypothesis and alternate hypothesis.

ii. Choose an alpha level.

iii. Find the critical value of z in a z table.

iv. Calculate the z test statistic.

v. Compare the test statistic to the critical z value and decide to support or reject the null hypothesis.

Following formula is being used to run Z test.

For the normal population with one sample

$$Z = \frac{\bar{x} + \mu}{\frac{\sigma^2}{\sqrt{n}}}$$

Where :
$\bar{x}$ = mean of the sample
μ is the assumed mean
σ is the standard deviation
n is the number of observations

Two Proportion Z-Test :

This tests for a difference in proportions. A two proportion z-test allows to compare two proportions to see if they are the same.

- The null hypothesis (H_0) for the test is that the proportions are the same.
- The alternate hypothesis (H_1) is that the proportions are not the same.

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Where
P_1 and P_2 = means of two sample
$\hat{p}$ = overall sample proportion (standard deviation of the sample)
$n1$ and $n2$ = numbers of observations of two samples

One sample z-test (one-tailed z-test)

- One sample z-test is used to determine whether a particular population parameter, which is mostly mean, significantly different from an assumed value.
- It helps to estimate the relationship between the mean of the sample and the assumed mean.
- In this case, the standard normal distribution is used to calculate the critical value of the test.
- If the z-value of the sample being tested falls into the criteria for the one-sided test, the alternative hypothesis will be accepted instead of the null hypothesis.
- A one-tailed test would be used when the study has to test whether the population parameter being tested is either lower than or higher than some hypothesized value.
- A one-sample z-test assumes that data are a random sample collected from a normally distributed population that all have the same mean and same variance.
- This hypothesis implies that the data is continuous, and the distribution is symmetric.
- Based on the alternative hypothesis set for a study, a one-sided z-test can be either a left-sided z-test or a right-sided z-test.

- For instance, if our H_0: $\mu_0 = \mu$ and Ha: $\mu < \mu_0$, such a test would be a one-sided test or more precisely, a left-tailed test and there is one rejection area only on the left tail of the distribution.
- However, if H_0: $\mu = \mu_0$ and Ha: $\mu > \mu_0$, this is also a one-tailed test (right tail), and the rejection region is present on the right tail of the curve.

Two sample z-test (two-tailed z-test)

- In the case of two sample z-test, two normally distributed independent samples are required.
- A two-tailed z-test is performed to determine the relationship between the population parameters of the two samples.
- In the case of the two-tailed z-test, the alternative hypothesis is accepted as long as the population parameter is not equal to the assumed value.
- The two-tailed test is appropriate when we have H_0: $\mu = \mu_0$ and H_0: $\mu \neq \mu_0$ which may mean $\mu > \mu_0$ or $\mu < \mu_0$
- Thus, in a two-tailed test, there are two rejection regions, one on each tail of the curve.

Z test sample :

If a sample of 400 male workers has a mean height of 67.47 inches, is it reasonable to regard the sample as a sample from a large population with a mean height of 67.39 inches and a standard deviation of 1.30 inches at a 5% level of significance.

Taking the null hypothesis that the mean height of the population is equal to 67.39 inches,
we can write:
H_0 : $\mu = 67.39$“
H_a: $\mu \neq 67.39$“
$x = 67.47$“
$\sigma = 1.30$“
$n = 400$

Assuming the population to be normal, we can work out the test statistic z as under:

$$Z = \frac{\bar{x} + \mu}{\frac{\sigma^2}{\sqrt{n}}} \qquad Z = \frac{67.47 + 67.39}{\frac{1.30^2}{\sqrt{400}}}$$

As H_0 is two-sided in the given question, we shall be applying a two-tailed test for determining the rejection regions at a 5% level of significance which comes to as under, using normal curve area table:

R : | z | > 1.96

The observed value of t is 1.231 which is in the acceptance region since

R: | z | > 1.96, and thus, H_0 is accepted.

6.5.2. Non-parametric test

Non-parametric test are also known is distribution-free test is considered less powerful as it uses less information in its calculation and makes fewer assumption about the data set. A nonparametric test, is which doesn't assume anything about the population parameters. Nonparametric tests include chi-square, Fisher's exact test and the Mann-Whitney test.

When non-parametric tests are used

- When the study is better represented by the median
- When the data has a normal distribution
- When there is ordinal data, ranked data, or outliers can't be removed
- When the sample size is very small
- When the measurement scale is nominal or ordinal

6.5.3. Chi Square test (X^2) :

Chi(X^2)test is used to determine whether observed data comes from a given theoretical distribution whether attributes in a given contingency table are independent and whether the mean values in several population are same. The Chi (X^2) distribution looks like a normal distribution which is skewed to the right. It is a continuous distribution which assumes only positive values. It begins at 0 and extended to infinity in a positive direction.

There are two types of chi-square tests. Both use the chi-square statistic and distribution for different purposes:

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where
C = degrees of freedom
O = observed value and
E = expected value
Σ = calculation for every single data item in your data set.

Due to tedious method of calculation most of time following technology is being used. :

- Chi Square Test in SPSS.
- Chi Square P-Value in Excel.

A chi-square statistic is one way to show a relationship between two categorical variables. In statistics, there are two types of variables: numerical (countable) variables and non-numerical (categorical) variables. The chi-squared statistic is a single number that tells how much

difference exists between observed counts and the counts expected if there were no relationship at all in the population.

There are a few variations on the chi-square statistic. However, all of the variations use the same idea, which is comparing expected values with the values you actually collect. One of the most common forms can be used for contingency tables:

$$c^2 = \sum_{i=1}^{k} \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

Where
O = observed value,
E = expected value and
I = "ith" position in the contingency table.

A low value for chi-square means there is a high correlation between two sets of data. In theory, if observed and expected values were equal ("no difference") then chi-square would be zero. If the chi-square value is more than the critical value, then there is a significant difference.

The Chi-square statistic can only be used on numbers. They can't be used for percentages, proportions, means or similar statistical values. For example, if you have 10 percent of 200 people, you would need to convert that to a number (20) before you can run a test statistic.

Chi Square P-Values.

A chi square test will give a p-value. The p-value exhibits that test results are significant or not. In order to perform a chi square test and get the p-value, fowling two pieces of information needed:

i. Degrees of freedom. That's just the number of categories minus 1.

ii. The alpha level(α). This is chosen by the researcher. The usual alpha level is 0.05 (5%), but could also have other levels like 0.01 or 0.10.

Degrees of freedom are placed as a subscript after the chi-square (X^2) symbol. For example, the following chi square shows 6 df :X^2_6.
And this chi square shows 4 df:X^2_4.

The Chi-Square Distribution

The chi-square distribution is a special case of the gamma distribution. A chi square distribution with n degrees of freedom is equal to a gamma distribution with a = n / 2 and b = 0.5 (or β = 2).

Let's say you have a random sample taken from a normal distribution. The chi square distribution is the distribution of the sum of these random samples squared. The degrees of freedom (k) are equal to the number of samples being summed. For example, if you have taken 10 samples from the normal distribution, then df = 10. The degrees of

freedom in a chi square distribution is also its mean. In this example, the mean of this particular distribution will be 10. Chi square distributions are always right skewed. However, the greater the degrees of freedom, the more the chi square distribution looks like a normal distribution.

6.5.4. The ANOVA Test

An ANOVA test is a way to find out if survey or experiment results are significant. In other words, it helps to figure out if need to reject the null hypothesis or accept the alternate hypothesis.

One-Way or Two-Way :

One-way or two-way refers to the number of independent variables (IVs) in analysis of Variance test.

- One-way has one independent variable (with 2 levels). For example: brand of cereal,
- Two-way has two independent variables (it can have multiple levels). For example: brand of cereal, calories.

Groups or Levels :

Groups or levels are different groups within the same independent variable. In the above example, levels for "brand of cereal" might be Lucky Charms, Raisin Bran, Cornflakes — a total of three levels. Levels for "Calories" might be: sweetened, unsweetened — a total of two levels.
If groups or levels have a hierarchical structure (each level has unique subgroups), then use a nested ANOVA for the analysis.

Replication :

With a two way ANOVA with replication, can have two groups and individuals within that group are doing more than one thing (i.e. two groups of students from two colleges taking two tests). If only have one group taking two tests, can use without replication.

6.5.5. Types of Tests :

There are two main types: one-way and two-way. Two-way tests can be with or without replication.

- One-way ANOVA between groups: It is used when to test two groups to see if there's a difference between them.
- Two way ANOVA without replication: It is used when have one group and double-testing that same group. For example, testing one set of individuals before and after they take a medication to see if it works or not.
- Two way ANOVA with replication: Two groups, and the members of those groups are doing more than one thing. For example, two groups of patients from different hospitals trying two different therapies.

1) One Way ANOVA :

A one way ANOVA is used to compare two means from two independent (unrelated) groups using the F-distribution. The null hypothesis for the test is that the two means are equal. Therefore, a significant result means that the two means are unequal.

Examples of one way ANOVA :

Situation 1: You have a group of individuals randomly split into smaller groups and completing different tasks. For example, you might be studying the effects of tea on weight loss and form three groups: green tea, black tea, and no tea.

Situation 2: Similar to situation 1, but in this case the individuals are split into groups based on an attribute they possess. For example, leg strength of people according to weight. You could split participants into weight categories (obese, overweight and normal) and measure their leg strength on a weight machine. The major limitations of the One Way NOVA test is that it show that at least two groups were different from each other. But it won't shows which groups are different**.**

2) Two Way ANOVA :

A Two Way ANOVA is an extension of the One Way ANOVA. With a Two Way ANOVA, there are two independents. Two way ANOVA is useful when have one measurement variable (i.e. a quantitative variable) and two nominal variables.

For example, to find out if there is an interaction between income and gender for anxiety level at job interviews. The anxiety level is the outcome, or the variable that can be measured. Gender and Income are the two categorical variables. These categorical variables are also the independent variables, which are called factors in a Two Way ANOVA. The factors can be split into levels. In the above example, income level could be split into three levels: low, middle and high income. Gender could be split into three levels: male, female, and transgender. Treatment groups are all possible combinations of the factors. In this example there would be 3 x 3 = 9 treatment groups.

6.5.6. Main Effect and Interaction Effect

The results from a Two Way ANOVA will calculate a main effect and an interaction effect. The main effect is similar to a One Way ANOVA: each factor's effect is considered separately. With the interaction effect, all factors are considered at the same time. Interaction effects between factors are easier to test if there is more than one observation in each cell. For the above example, multiple stress scores could be entered into cells.

6.6. FACTOR ANALYSIS

6.6.1. Meaning:

Factor analysis is a technique that is used to reduce a large number of variables into fewer numbers of factors. This technique extracts maximum common variance from all variables and puts them into a common score. As an index of all variables, we can use this score for further analysis. Factor analysis is part of general linear model (GLM) and this method also assumes several assumptions like :

- there is linear relationship,
- there is no multi co-linearity, it includes relevant variables into analysis, and
- is true correlation between variables and factors.

6.6.2. Types of factoring:

There are different types of methods used to extract the factor from the data set:

i. **Principal component analysis**: This is the most common method used by researchers. PCA starts extracting the maximum variance and puts them into the first factor. After that, it removes that variance explained by the first factors and then starts extracting maximum variance for the second factor. This process goes to the last factor.

ii. Common factor analysis: The second most preferred method by researchers, it extracts the common variance and puts them into factors. This method does not include the unique variance of all variables. This method is used in SEM.

iii. Image factoring: This method is based on correlation matrix. OLS Regression method is used to predict the factor in image factoring.

iv. Maximum likelihood method: This method also works on correlation metric but it uses maximum likelihood method to factor.

v. Other methods of factor analysis: Alfa factoring outweighs least squares. Weight square is another regression based method which is used for factoring.

1) Factor loading:

Factor loading is basically the correlation coefficient for the variable and factor. Factor loading shows the variance explained by the variable on that particular factor. In the SEM approach, as a rule of thumb, 0.7 or higher factor loading represents that the factor extracts sufficient variance from that variable.

2) Eigen values: Eigen values is also called characteristic roots. Eigenvalues shows variance explained by that particular factor out of the

total variance. From the commonality column, we can know how much variance is explained by the first factor out of the total variance.

3) Factor score: The factor score is also called the component score. This score is of all row and columns, which can be used as an index of all variables and can be used for further analysis. .

4) Rotation method: Rotation method makes it more reliable to understand the output. Eigenvalues do not affect the rotation method, but the rotation method affects the Eigenvalues or percentage of variance extracted. There are a number of rotation methods available: No rotation method, Varimax rotation method, Quartimax rotation method, Directoblimin rotation method, and Promax rotation method.

Assumptions:

i. No outlier: Assume that there are no outliers in data.

ii. Adequate sample size: The case must be greater than the factor.

iii. No perfect multi co-linearity: Factor analysis is an interdependency technique. There should not be perfect multi co-linearity between the variables.

iv. Homoscedasticity: Since factor analysis is a linear function of measured variables, it does not require homoscedasticity between the variables.

v. Linearity: Factor analysis is also based on linearity assumption. Non-linear variables can also be used. After transfer, however, it changes into linear variable.

vi. Interval Data: Interval data are assumed.

6.7. INTERPRETATION OF DATA

6.7.1. Meaning

Data interpretation is the process of reviewing data through some predefined processes which will help assign some meaning to the data and arrive at a relevant conclusion. It involves taking the result of data analysis. Data analysis is the process of ordering, categorizing, manipulating, and summarizing data to obtain answers to research questions. It is usually the first step taken towards data interpretation.

It is evident that the interpretation of data is very important, and as such needs to be done properly. Therefore, researchers have identified some data interpretation methods to aid this process.

6.7.2. Significance of Data interpretation

The purpose of collection and interpretation is to acquire useful and usable information and to make the most informed decisions possible.

From businesses, to newlyweds researching their first home, data collection and interpretation provides limitless benefits for a wide range of institutions and individuals.

i. It is through interpretation that the researcher can well understand the abstract principle that works beneath his findings. Through this he can link up his findings with those of other studies, having the same abstract principle, and thereby can predict about the concrete world of events. Fresh inquiries can test these predictions later on. This way the continuity in research can be maintained.

ii. Interpretation leads to the establishment of explanatory concepts that can serve as a guide for future research studies; it opens new avenues of intellectual adventure and stimulates the quest for more knowledge.

iii. Researcher can better appreciate only through interpretation why his findings are what they are and can make others to understand the real significance of his research findings.

iv. The interpretation of the findings of exploratory research study often results into hypotheses for experimental research and as such interpretation is involved in the transition from exploratory to experimental research. Since an exploratory study does not have a hypothesis to start with, the findings of such a study have to be interpreted on a post-factum basis in which case the interpretation is technically described as 'post factum' interpretation.

6.7.3. Data Interpretation Methods

Data interpretation methods are how analysts help people make sense of numerical data that has been collected, analyzed and presented. Data, when collected in raw form, may be difficult for the layman to understand, which is why analysts need to break down the information gathered so that others can make sense of it.

There are two main methods in which this can be done, namely; quantitative methods and qualitative methods.

1) Qualitative Data Interpretation Method

The qualitative data interpretation method is used to analyse qualitative data, which is also known as categorical data. This method uses texts, rather than numbers or patterns to describe data.

Qualitative data is usually gathered using a wide variety of person-toperson techniques, which may be difficult to analyze compared to the quantitative research method. Unlike the quantitative data which can be analyzed directly after it has been collected and sorted, qualitative data needs to first be coded into numbers before it can be analyzed. This is because texts are usually cumbersome, and will take more time and result in a lot of errors if analyzed in its original state. Coding done by the analyst should also be documented so that it can be reused by others and also analyzed.

There are two main types of qualitative data, namely; nominal and ordinal data. These two data types are both interpreted using the same method, but ordinal data interpretation is quite easier than that of nominal data. In most cases, ordinal data is usually labelled with numbers during the process of data collection, and coding may not be required. This is different from nominal data that still needs to be coded for proper interpretation.

2) Quantitative Data Interpretation Method

The quantitative data interpretation method is used to analyze quantitative data, which is also known as numerical data. This data type contains numbers and is therefore analyzed with the use of numbers and not texts.

Quantitative data are of two types namely; discrete and continuous data. Continuous data is further divided into interval data and ratio data, with all the data types being numeric. Due to its natural existence as a number, analysts do not need to employ the coding technique on quantitative data before it is analyzed. The process of analyzing quantitative data involves statistical modelling techniques such as standard deviation, mean and median.

6.7.4. Precautions in Data interpretation

It must be always remember that even if the data are properly collected and analyzed, wrong interpretation would lead to inaccurate conclusions. It is, therefore, absolutely essential that the task of, interpretation be accomplished with patience in an impartial manner and also in correct perspective. Following precautions need to be taken in data interpretation.

i. Identify the Required Data Type

Researchers need to identify the type of data required for particular research. It may be nominal, ordinal, interval, or ratio data. The key to collecting the required data to conduct research is to properly understand the research question. If the researcher can understand the research question, then he can identify the kind of data that is required to carry out the research.

ii. Avoid Biases

There are different kinds of biases a researcher might encounter when collecting data for analysis. Although biases sometimes come from the researcher, most of the biases encountered during the data collection process is caused by the respondent.

There are two main biases, namely, response bias and non-response bias. Researchers may not be able to eliminate these biases, but there are ways in which they can be avoided and reduced to a minimum. Response biases are biases that are caused by respondents intentionally giving wrong answers to responses, while non-response bias occurs when

the respondents don't give answers to questions at all. Biases are capable of affecting the process of data interpretation.

iii. Use Close Ended Surveys

Although open-ended surveys are capable of giving detailed information about the questions and allow respondents to fully express themselves, it is not the best kind of survey for data interpretation. It requires a lot of coding before the data can be analyzed.

Close-ended surveys, on the other hand, restrict the respondents' answer to some predefined options, while simultaneously eliminating irrelevant data. This way, researchers can easily analyze and interpret data.

However, close-ended surveys may not be applicable in some cases, like when collecting respondent's personal information like name, credit card details, phone number, etc.

iv. Researcher must invariably satisfy himself that (a) the data are appropriate, trustworthy and adequate for drawing inferences; (b) the data reflect good homogeneity; and that (c) proper analysis has been done through statistical methods.

v. The researcher must remain cautious about the errors that can possibly arise in the process of interpreting results.

6.8. SUMMARY:

The statistics vary based on whether the study is a between-participants or correlated-groups design. It is imperative that the appropriate statistic be used to analyze the data collected in an experiment. The first point to consider when determining which statistic to use is whether it should be a parametric or nonparametric statistic. This decision is based on the type of data collected, the type of distribution to which the data conform, and whether any parameters of the distribution are known. The second consideration is whether a between participants or correlated-groups design has been used. This information enables us to select and conduct the statistical test most appropriate to the particular study's design and data.

6.9. EXERCISE:

A. Test your knowledge by choosing the correct option:

1. …….is a statistic that measures the degree to which two variables move in relation to each other. (mean deviation , Correlation , standard deviation, Regression)

2. The ……..coefficient's values range between -1.0 and 1.0. (Mean deviation , Correlation , standard deviation, Regression)

3. A perfect positive correlation means that the correlation coefficient is exactly…. (1, -1, 0, -0.5)

4. When price of a commodity increases its demand goes down or vice-versa, its ……. correlation (Positive, Negative, Partial, Multiple)

5. Which method is a diagrammatic representation of a bivariate distribution for ascertaining the nature of correlation between the variables? (Scatter diagram, Regression lines, mean deviation, standard deviation)

6. ……is an act in statistics whereby an analyst tests an assumption regarding a population parameter. (Hypothesis testing, Scatter diagram, Regression lines, mean deviation)

7. A statistical hypothesis which under test is usually a hypothesis of no difference and hence is called …..hypothesis. (Alternative, Null, Second, negative)

8. A hypothesis which is accepted when H_0 is rejected is called ……..hypothesis. (Alternative, Null, Second, negative)

9. ……test are also known is distribution-free test. (Parametric, Non-parametric, Two- sample , Paired T test)

10. A …….is a type of inferential statistic used to determine if there is a significant difference between the means of two groups. (t-test,F test, chi Square, ANOVA)

11. ……. refers to the values in a study that has the freedom to vary and are essential for assessing the importance and the validity of the null hypothesis. (Power of test, Parametric tests, Degrees of freedom, Null hypothesis)

12. Which test is used to determine whether observed data comes from a given theoretical distribution whether attributes in a given contingency table are independent and whether the mean values in several population are same? (F test, -test, Chi (X^2), ANOVA)

13. The …… distribution is a special case of the gamma distribution. (F test, -test, Chi (X^2), ANOVA)

14. …….. helps to figure out if need to reject the null hypothesis or accept the alternate hypothesis. (F test, -test, Chi (X^2), ANOVA)

15. A …… ANOVA is used to compare two means from two independent (unrelated) groups usingthe F-distribution. (one way, two way, Replication, Groups)

16. …… analysis is a technique that is used to reduce a large number of variables into fewer numbers of factors. (One way ANOVA, Two way ANOVA, Factor, Replication)

17. ……is the process of reviewing data through some predefined processes which will help assign some meaning to the data and

arrive at a relevant conclusion. (Data interpretation, Data collection, Hypothesis, Review of literature)

18. The qualitative data interpretation method is used to analyse …….data (Numerical, categorical, Statistical)

(**Answers** : 1. Correlation, 2. Correlation, 3. 1, 4. Negative, 5. Scatter, 6. diagram, 7.Hypothesis testing, 8. Null, 9. Alternative, 10. Non-parametric, 11. t-test, 12. Degrees of freedom, 13. Chi (X^2), 14. Chi (X^2), 15. ANOVA, 16. one way, 17. Factor, 18 Data interpretation, 19. Categorical)

B. Short Answer Questions:

1. Explain the concept correlation
2. Elaborate the types of correlation.
3. Explain the different shapes of Scatter diagram
4. Define line of regression
5. Which are the Properties of Regression coefficients?
6. Explain the level of significance.
7. Differentiate between Type I and Type I errors.
8. Explain the one tailed and two tailed test.
9. Explain the Parametric tests.
10. Which are the assumptions of t-Test?
11. Which are the General Steps for an F Test?
12. Explain the concept of Z test.
13. Explain two way ANOVA.
14. Identify the different types of factoring.
15. Which Precautions are needed in Data interpretation?

6.10 REFERENCE

I. Sherri L. Jackson, (2009) , Research Methods and Statistics, A Critical Thinking Approach, ISBN-13: 978-0-495-51001-7

II. BASIC ECONOMETRICS,(2003) Damodar N. Gujarati, *McGraw-Hill Higher Education* ISBN: 978-0-07-233542-2

III. Sachdeva J.K.(2011). Business Research Methodology. New Delhi: Himalaya Publishing House.

IV. Michael V.P.(1997). Research Methodology in Management. Delhi: Himalaya Publishing House

Module IV
RESEARCH REPORT AND MODERN PRACTICES IN RESEARCH

7

RESEARCH REPORT

Unit structure

7.0 OBJECTIVES

- To understand essentials of research report writing.
- To study different structures and layout used for report writing.
- To differentiate between footnote and bibliography.
- To know different styles of references and citation methods.

7.1 INTRODUCTION:

7.1.1 Meaning

The research work is presented in a written form. The practical utility of research study depends heavily on the way it is presented to those who are expected to act on the basis of research findings. Research report is a written document containing key aspects of research project. Research report is a medium to communicate research work with relevant people. It is also a good source of preservation of research work for the future reference. Many times, research findings are not followed because of improper presentation. Preparation of research report is not an easy task. It is an art. It requires a good deal of knowledge, imagination, experience, and expertise. It demands a considerable time and money.

A research report is a document prepared by an analyst or strategist who is a part of the investment research team in a stock brokerage or

investment bank. A research report may focus on a specific stock or industry sector, a currency, commodity or fixed-income instrument, or on ageographic region or country. Research reports generally, but not always, have actionable recommendations such as investment ideas that investors can act upon.

7.1.2 Essentials of research report:

Report writing differs from person to person depending on personality, imaginative and creative abilities, experience, and training. However, most researchers agree that following general principles must be kept in mind to produce a better research report. These principles are often called as qualities or requirements of a good report.

1. **Accuracy:** The information collected and presented by the researcher must be accurate to the best of his knowledge. Wherever, possible the sources of data can be mentioned to make report more authentic. The consistency must be maintained while writing report it includes the data, language and also the presentation.

2. **Concise:** The research report must be concise, brief and straight to the point. Irrelevant topics can be avoided. There is no minimum or maximum limit of pages that report should consist. Hence, a researcher can stick to the research topic and cover the pointers related to that.

3. **Logical arrangement**: As such there is no standard format to be followed for writing a research project yet, researcher needs to take care that there is a proper flow and logical connectivity among chapters. Eg: Conclusions and suggestions of the study will follow only after the analysis and interpretation of the data.

4. **Date and signature:** These two are the important elements of the research report and therefore its existence in research report makes the report more concrete. It also helps new researchers to know the period of the study. The signatures of officials and researchers makes the report more reliable.

5. **Impersonal style of writing:** The researcher needs to write a research report in “third form” i.e. without using the words like I, me, you, we etc. The language used must be decent and official. Informal words, expressions must be avoided. No need to use sophisticated language even a simple language can serve the purpose of communicating research outcomes to readers.

6. **Timely submission of report:** For every research report a particular time frame is allotted and researcher must stick to that time. It requires proper planning and allocation of work so that the research repot can be completed in a given time frame.

7. **References:** This is the most important section of research report. The researcher refers several research works, books, research papers etc and researcher may take some content from same, hence, the credit must be given to all those authors and publishers. Thus, references must be mentioned at the end of the research project.

8. **Attractive presentation:** If research report contains only text and data, the reader will not be interested to read in detail. Hence, wherever possible, data must be converted in colourful charts and pointers or variables can be put forward in the form of diagrams. It will attract the attention of readers.

7.1.3 Significance of research report:

1. **Decision Making Tool:** Today's complex business organizations require thousands of information. Â reports provide the required information a large number of important decisions in business or any other area are taken on the basis of information presented in the reports. This is one of the great importance of report.

2. **Transmission of knowledge**: The knowledge that has been obtained on the basis of research need transmission for proper utilization of the resources invested. Because of that reason, it is always advisable to prepare to report in a written manner so that it can also provide knowledge to layman in understanding various social problems.

3. **Investigation:** Whenever there is any problem, a committee or commission or study group investigates the problem to find out the reason behind the problem and present the findings with or without the recommendation in the form of a report.

4. **Inspiration for Further Research:** Research report inspires others to undertake further research in the same line or in any other inter-disciplinary fields. If the report appears to be interesting and a novel one, it is more likely to draw the attention of the social scientists.

5. **Evaluation:** Large scale organizations are engaged in multidimensional activities. It is not possible for a single top executive to keep personal watch on what others are doing.

   So, the executive depends on reports to evaluate the performance of various departments or units.

6. **Development of skill:** Report writing skill develops the power of designing, organization coordination, judgment and communication.

7. **Presentation of finding:** Society is more concerned with the finished product in terms of output of research which has the input of immense money, human resources and precious time. Therefore, the social

utility of the research report lies in its exposure to the layman as well as submission to the sponsoring agency.

8. **Professional Advancement:** Report also plays a major role in professional achievement. For promotion to the rank-and-file position, satisfactory job performance is enough to help a person. But for promotion to high level position, intellectual ability is highly required. Such ability can be expressed through the report submitted to higher authority.

9. **A managerial Tool:** Various reports make activities easy for the managers. For planning, organizing, coordinating, motivating and controlling, manager needs help from a report which acts as a source of information.

10. **Encountering Advance and Complex Situation:** In a large business organization, there is always some sort of labor problems which may bring complex situations. To tackle that situation, managers take the help of a report.

7.2 STRUCTURE OF THE RESEARCH REPORT:

7.2.1 Meaning

Research reports are recorded data prepared by researchers or statisticians after analyzing information gathered by conducting organized research. A research report is a reliable source to recount details about a conducted research and is most often considered to be a true testimony of all the work done to gather details of research.

Researchers can prepare report at their ease as there are no such set rules or procedure of writing reports. However, following general guidelines can help for writing research reports:

7.2.2 Preliminary content:

1. A certificate stating the details of university/institution where the research project is to be submitted. The certificate must be duly signed by researches and research guide.
2. **Acknowledgement:** Researcher should take this as an opportunity to thank all those people who have helped to conduct the research work successfully. Eg: Statistician for statistical help, respondents for filling up the forms, etc.
3. **Table of content**: It is also known as an Index. This will help the reader to trace the content on the given page number.
4. **List of tables and graphs**: Tables and graphs are part of every research and thus a separate index can be prepared mentioning the table and chart displayed on a particular page.

5. **Abbreviations:** There are certain words which are used repetitively in the report and thus, an abbreviation can be used instead of writing a complete word. The list of abbreviations used by researcher must be given at the preliminary stage, so that, reader gets an idea.

7.2.3 Body of the report:

Chapter 1 – Introduction

This is the first chapter of the research work, wherein the reader is introduced to the basis of the research topic. The reader gets an idea about the complete work from this chapter.

This chapter may consist of following points:

- Objectives of the study
- Significance of the study
- Scope of the study
- Limitations of the study
- Origin of the study
- Chapter scheme of the research

Chapter 2 - Review of literature:

This chapter consists of a brief summary framed by researcher about the past research or studies done by other researchers. It may include research done at national or international level as well. This chapter will help the reader to know the past contributions made by similar researchers.

Chapter 3 – Research methodology

In this chapter, a complete process and research tools are used in research is been highlighted. It includes:

- Hypotheses of the research work.
- A brief report on pilot study
- Reliability check of questionnaire
- Conceptual framework, if relevant to the topic
- Research design
- Population and Sample size
- Techniques used for data analysis in research process.
- Description of questionnaire.
-

Chapter IV – Core concept/ Population of the research

This chapter includes a complete information about a core concept selected by the researcher to conduct research. Eg: Research topic: A study on causes and effects of stress among youth. In this case, a separate chapter can be framed on Stress (Core concept) and another chapter can be framed on Youth (Population). The chapter may include:

- Meaning of stress
- Causes of stress
- Impact of stress
- Symptoms of stress
- Theory put forward by earlier researchers.

Chapter V – Findings and observations

This chapter is the heart of the research project as it consists of a compilation of data collected by researcher. The data is presented in a form of tables and charts as it becomes easy for a stranger to relate it to the topic. Wherever, required justification and interpretation of the presented data must be given. Researcher can also write their observation that they come across while interacting with respondents. This chapter also gives the explanation for the hypothesis and objectives framed by the researcher.

Chapter VI – Conclusion and suggestions

The complete research process is undertaken to come up with this chapter. In this chapter, researcher gives suggestions based on analysis and data collected by the researcher. A concrete conclusion to sum up the topic must also be framed in this chapter. Researchers can link their suggestion with theories or they can come up with their own model in the form of suggestion.

Chapter VII – Road ahead

This is the last chapter of the project, wherein, researcher chalks out the area which is not covered and there is a scope for future researchers to do in depth study. This chapter is very brief and concise in nature.

7.2.4 Supplementary material

This section includes all those extra documents used for conducting research must be attached and described. It includes following documents:

a. Questionnaire:

Every researcher uses a questionnaire to collect primary data from respondent. A copy of questionnaire must be attached. If the questionnaire, is translated in different language then questionnaire in both the languages must be attached.

b. Letters:

If the researcher has received any letters from companies or respondent, the same must be attached at the end of the report. If researcher has submitted any letters for the process of data collection, same can be attached at the end.

c. Report of plagiarism check:

Now-a- days plagiarism check is must and therefore, every researcher needs to attach a plagiarism certificate, stating the content is authentic and first hand. Many software are available, researcher can have a plagiarism check and get the certificate.

7.2.5 References/ Bibliography:

All those books, research papers, sites, newspapers referred by researcher needs to be mentioned under this heading. All these references must be mentioned in different categories in alphabetical manner. It becomes a huge source of references for new researchers to refer.

7.2.6 Images:

Although, this section is optional but, if research demands images, those can be included under this section with proper heading and brief description. It can make the research more interesting and attractive.

7.3 REFERENCES AND CITATION METHODS:

7.3.1 Meaning:

A citation is a way of giving credit to individuals for their creative and intellectual works that is been utilized by a researcher for the work. It can also be used to locate particular sources and combat plagiarism. Typically, a citation can include the author's name and date. A citation style dictates the information necessary for a citation and how the information is ordered, as well as punctuation and other formatting. A citation tells the readers where the information came from.

The citation can be maintained in three forms, they are as under:

a. The name-and-year system

Citations: When researcher cite the source of information in the report,the names of the authors and the date of publication is mentioned.

Jenkins and Busher (1979) report that beavers eat several kinds of herbaceous plants as well as the leaves, twigs, and bark of most species of woody plants that grow near water.

Beavers have been shown to be discriminate eaters of hardwoods (Crawford, Hooper, and Harlow 1976).

b. The alphabet-number system.

Citations: When researcher cite the source of information in the report, a number in parentheses are given that corresponds to the number of the source in the alphabetical listing in the "References."

Jenkins and Busher report that beavers eat several kinds of herbaceous plants as well as the leaves, twigs, and bark of most species of woody plants that grow near water (4).

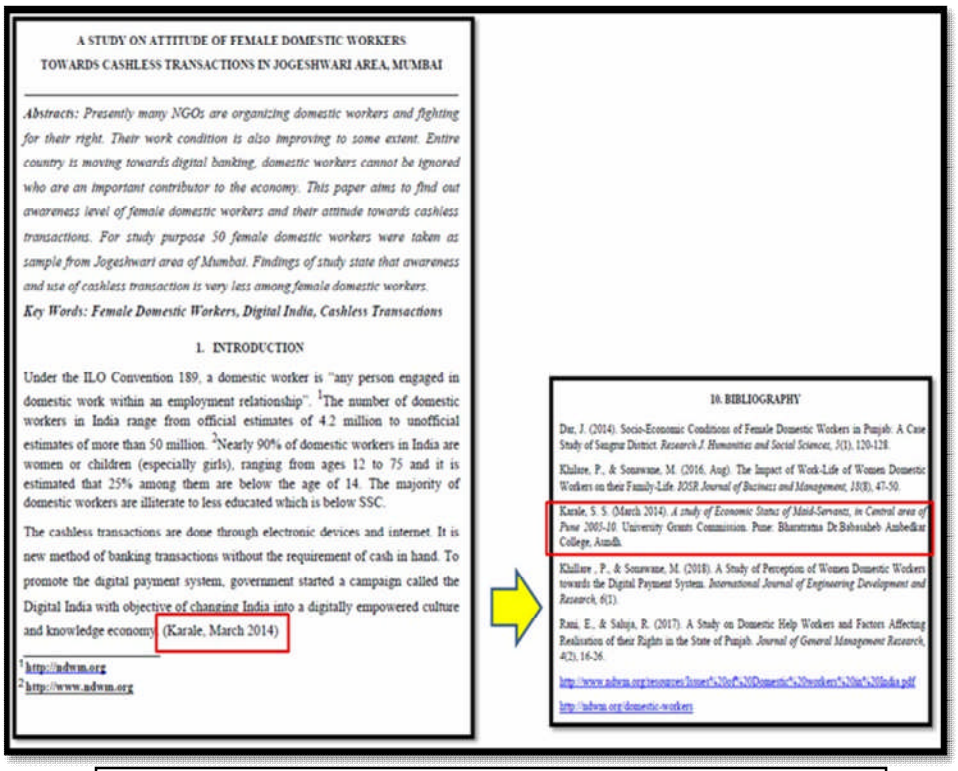
Beavers have been shown to be discriminate eaters of hardwoods (3).

7.3.2 The Citation-Order System (typically used in engineering--IEEE documentation).

Citations: When researcher cite the sources of information in the report, a given number in brackets that corresponds to the number of the source listed in the order in which they appear in the report, the source listed first as [1], the next source [2], etc.

Jenkins and Busher report that beavers eat several kinds of herbaceous plants as well as the leaves, twigs, and bark of most species of woody plants that grow near water [1].

Beavers have been shown to be discriminate eaters of hardwoods [2].



A STUDY ON ATTITUDE OF FEMALE DOMESTIC WORKERS
TOWARDS CASHLESS TRANSACTIONS IN JOGESHWARI AREA, MUMBAI

Abstracts: Presently many NGOs are organizing domestic workers and fighting for their right. Their work condition is also improving to some extent. Entire country is moving towards digital banking, domestic workers cannot be ignored who are an important contributor to the economy. This paper aims to find out awareness level of female domestic workers and their attitude towards cashless transactions. For study purpose 50 female domestic workers were taken as sample from Jogeshwari area of Mumbai. Findings of study state that awareness and use of cashless transaction is very less among female domestic workers.

Key Words: Female Domestic Workers, Digital India, Cashless Transactions

1. INTRODUCTION

Under the ILO Convention 189, a domestic worker is "any person engaged in domestic work within an employment relationship". [1]The number of domestic workers in India range from official estimates of 4.2 million to unofficial estimates of more than 50 million. [2]Nearly 90% of domestic workers in India are women or children (especially girls), ranging from ages 12 to 75 and it is estimated that 25% among them are below the age of 14. The majority of domestic workers are illiterate to less educated which is below SSC.

The cashless transactions are done through electronic devices and internet. It is new method of banking transactions without the requirement of cash in hand. To promote the digital payment system, government started a campaign called the Digital India with objective of changing India into a digitally empowered culture and knowledge economy (Karale, March 2014)

[1] http://adwm.org
[2] http://www.adwm.org

10. BIBLIOGRAPHY

Dar, J. (2014). Socio-Economic Conditions of Female Domestic Workers in Punjab: A Case Study of Sangrur District. *Research J. Humanities and Social Sciences, 5*(1), 120-128.

Khilare, P., & Sonawane, M. (2016, Aug). The Impact of Work-Life of Women Domestic Workers on their Family-Life. *IOSR Journal of Business and Management, 18*(8), 47-50.

Karale, S. S. (March 2014). *A study of Economic Status of Maid-Servant, in Central area of Pune 2005-10.* University Grants Commission. Pune: Bharatratna Dr.Babasaheb Ambedkar College, Aundh.

Khillare , P., & Sonawane, M. (2018). A Study of Perception of Women Domestic Workers towards the Digital Payment System. *International Journal of Engineering Development and Research, 6*(1).

Rani, E., & Saluja, R. (2017). A Study on Domestic Help Workers and Factors Affecting Realisation of their Rights in the State of Punjab. *Journal of General Management Research, 4*(2), 16-26.

http://www.adwm.org/resources/Issue%20of%20Domestic%20workers%20in%20India.pdf

http://adwm.org/domestic-workers

Citation in the research report

7.4 REFERENCES:

7.4.1 Meaning

A reference gives the readers details about the source so that they have a good understanding of what kind of source it is and could find the source themselves if necessary. The references are typically listed at the end of the research report.

7.4.2 APA style:

APA is the style of documentation of sources used by the American Psychological Association. This form of writing research papers is used mainly in the social sciences, like psychology, anthropology, sociology, as well as education and other fields.

Material Type	In-text Citation	Bibliography
A book	(Sapolsky, 2017)	Sapolsky, R. M. (2017). *Behave: The biology of humans at our best and worst*. Penguin Books.
Chapter in an edited book (If the chapter is from an authored book, use the book citation)	(Dillard, 2020)	Dillard, J. P. (2020). Currents in the study of persuasion. In M. B. Oliver, A. A. Raney, & J. Bryant (Eds.), *Media effects: Advances in theory and research* (4th ed., pp. 115–129). Routledge.
An article in a print journal	(Weinstein, 2009)	Weinstein, J. (2009). "The market in Plato's Republic." *Classical Philology*, *104*(4), 439-458.
An article in an electronic journal	(Grady et al., 2019)	Grady, J. S., Her, M., Moreno, G., Perez, C., &Yelinek, J. (2019). Emotions in storybooks: A comparison of storybooks that represent ethnic and racial groups in the United States. *Psychology of Popular Media Culture*, *8*(3), 207–217. https://doi.org/10.1037/ppm0000185
A website	(Bologna, 2019)	Bologna, C. (2019, October 31). *Why some people with anxiety love watching horror movies*. HuffPost. https://www.huffpost.com/entry/anxiety-love-watching-horror-movies_l_5d277587e4b02a5a5d57b59e

Source: https://pitt.libguides.com/citationhelp/apa7

7.4.3 Chicago style:

Chicago is a documentation style that has been published by the Chicago University Press since 1906. This citation style incorporates rules of grammar and punctuation common in American English. Typically, Chicago style presents two basic documentation systems:

(1) Notes and bibliography
(2) Author-date.
Choosing between the two often depends on subject matter and the nature of sources cited, as each system is favored by different groups of scholars.The notes and bibliography style is preferred by many in the humanities, including those in literature, history, and the arts. This style presents bibliographic information in notes and, often, a bibliography.

Material Type	**Notes/Bibliography Style**
A book in print	Note Style: 1. Michael Pollan, *The Omnivore's Dilemma: A Natural History of Four Meals* (New York: Penguin, 2006), 99–100. Duplicate Note: 2. Pollan, *Omnivore's Dilemma,* 3. Bibliography: Pollan, Michael. *The Omnivore's Dilemma: A Natural History of Four Meals*. New York: Penguin, 2006.
An article in a print journal	Note Style: 1. Joshua I. Weinstein, "The Market in Plato's *Republic*," *Classical Philology* 104 (2009): 440. Duplicate Note: 2. Weinstein, "Plato's *Republic*," 452–53. Bibliography: Weinstein, Joshua I. "The Market in Plato's *Republic*." *Classical Philology* 104 (2009): 439–58.
An article in an electronic journal	Note Style: 1. GueorgiKossinets and Duncan J. Watts, "Origins of Homophily in an Evolving Social Network," *American Journal of Sociology* 115 (2009): 411, accessed February 28, 2010, doi:10.1086/599247. Duplicate Note: Kossinets and Watts, "Origins of Homophily," 439. Bibliography: Kossinets, Gueorgi, and Duncan J. Watts. "Origins of Homophily in an Evolving Social Network." *American Journal of Sociology* 115 (2009): 405–50. Accessed February 28, 2010. doi:10.1086/599247.
A website	Note Style: 1."Google Privacy Policy," last modified March 11, 2009, http://www.google.com/intl/en/privacypolicy.html. Duplicate Note: "Google Privacy Policy." Bibliography: Google. "Google Privacy Policy." Last modified March 11, 2009. http://www.google.com/intl/en/privacypolicy.html.

Source: https://pitt.libguides.com/citationhelp/chicago

The author-date style has long been used by those in the physical, natural, and social sciences. In this system, sources are briefly cited in the text, usually in parentheses, by author's last name and date of publication. The short citations are amplified in a list of references, where full bibliographic information is provided.

Author/Date Style	**In-text Citation**	**Bibliography**
A book	(Pollan 2006, 99–100)	Pollan, Michael. 2006. *The Omnivore's Dilemma: A Natural History of Four Meals*. New York: Penguin.
An article in a print journal	(Weinstein 2009, 440)	Weinstein, Joshua I. 2009. "The Market in Plato's *Republic*." *Classical Philology* 104:439–58.
An article in an electronic journal	(Kossinets and Watts 2009, 411)	Kossinets, Gueorgi, and Duncan J. Watts. 2009. "Origins of Homophily in an Evolving Social Network." *American Journal of Sociology* 115:405–50. Accessed February 28, 2010. doi:10.1086/599247.
A website	(Google 2009)	Google. 2009. "Google Privacy Policy." Last modified March 11. http://www.google.com/intl/en/privacypolicy.html.

Source: https://pitt.libguides.com/citationhelp/chicago

7.4.4 MLA Style:

MLA (Modern Language Association) style for documentation is widely used in the humanities, especially in writing on language and literature. MLA style features brief parenthetical citations in the text keyed to an alphabetical list of works cited that appears at the end of the work.

Material Type	**Works Cited**
Book in print	Card, Claudia. *The Atrocity Paradigm: A Theory of Evil*. Oxford UP, 2005.
eBook	Gaither, Milton. *Homeschool: An American History*. Palgrave Macmillan, 2017. *SpringerLink*, doi-org.pitt.idm.oclc.org/10.1057/978-1-349-95056-0.

An article in a print journal	Doggart, Julia. "Minding the Gap: Realizing Our Ideal Community Writing Assistance Program." *The Community Literacy Journal*, vol. 2, no. 1, 2007, pp. 71-80.
An article in an electronic journal	Sherrard-Johnson, Cherene. "'A Plea for Color': Nella Larsen's Iconography of the Mulatta." *American Literature*, vol. 76, no. 4, 2004, pp. 833-869, doi: 10.1215/00029831-76-4-833.
A encyclopaedia entry	"Patanjali." *Benét's Reader's Encyclopedia*, edited by Bruce Murphy, 4th ed., HarperCollins Publishers, 1996, p. 782.
A government publication	United States, Federal Maritime Commission. *Hawaiian Trade Study: An Economic Analysis*. Government Printing Office, 1978.
An interview you conducted	Brandt, Deborah. Personal interview. 28 May 2008. (Note: List the interview under the name of the interviewee)
A film/DVD	Note: This depends on the focus of your work. Please see the MLA Style blog for a detailed explanation.
A Page on a Website with no author	"Stunning Lakeside View on Lake Erie." *VisitPA*, Commonwealth of Pennsylvania, 7 June 2018, www.visitpa.com/article/stunning-lakeside-views-lake-erie.
A Page on a Website with an author	Del Castillo, Inigo. "How Not to Kill Your Houseplants, According to Botanists." *Apartment Therapy*, 29 Jan. 2020, www.apartmenttherapy.com/houseplant-tips-botanists-36710191.
Artwork - from website	Sherald, Amy. *Former First Lady, Michelle Obama*. 2018. *National Portrait Gallery*, npg.si.edu/object/npg_NPG.2018.15.

Source: https://pitt.libguides.com/citationhelp/mla8thedition

7.5 FOOTNOTES:

7.5.1 Meaning

While preparing research report the materials, books, articles, published or unpublished material etc should be given proper

acknowledgement by way of footnotes and bibliography. Oxford Dictionary refers report as an ascertained fact of record.

Footnotes are used to give the credit for borrowed words, ideas, symbols or other forms of expression should be given and their sources should be stated in the text or footnotes. Footnotes are of two kinds they are as follows:

- Content notes
- Reference notes

The objectives of having footnotes in research project are as follows:

1. To acknowledge in the author of original work
2. To refer the source of information and to establish the validity of evidence.
3. To amplify or clarify the ideas or information presented in the text.
4. To give original version of material.
5. To provide cross reference to various parts of thesis.

The name and complete detail form where the information and paragraph has been taken is given in footnotes. If the information is on one single page it could be preceded by "P". If it is on two or more pages it could be "PP". Following format of footnotes must be followed:

- Author's name
- Title of the work
- Place of publication
- (Wherever required)
- Publisher's name
- Year
- Page no

Examples of footnotes are as follows:

A) Citing Book:

1. Mann, Social change and social research, New Delhi, concept publishing company, 1988. P – 25.
2. Ibid, PP 20 – 24

B) Newspaper articles:

1. Kumar, Naresh, "Exploring sea for economic progress", The Economic Times (Bangalore) August 8, 1989. P – 6.
2. Kamal Shal, "Effects of recession on corporate world". The Times of India (Mumbai), March 8, 2009.

C) Research papers:

1. Mr. Irshad. I and Bhat. Apresented and published the paper "The vitality and role of Self Help Group (SHGS) in women upliftment: Special reference to Kashmir" in International Journal of Research – Granthaalyah. Vol.3 Issue.8 in the year 2015.

D) Citing Thesis or dissertation:

1. Bangalore: Indian Institute of Science, 1986.

E) Publications of Government and public organisations:

1. World Bank, Rural development sector policy paper, Washington DC, 1975. P – 26.

F) Internet and websites:

https://mavimindia.org/

7.6 BIBLIOGRAPHY:

7.6.1 Meaning

This is the first terminal item presented at the end of the report. The bibliography contains the list of books, articles, and magazines etc which were referred by reporter while preparing research report. The list is presented at the end of the report. The list must be in alphabetical order. Hence, it becomes easy to find and identify particular book or articles. Alternatively, the list of names could be grouped like:

- Books	- Magazine	- Links
- Articles	- Research papers	- Newspapers

Bibliography is different from footnotes. Footnotes are placed at the end of the page at the bottom. Whereas, on the other hand bibliography are placed at the end of the report. In bibliography name of author is written later and surname is written first. Eg: Drucker, Peter. E

The functions of bibliography are different from footnotes. Bibliography gives the identification detail, as a whole. Footnotes give the complete detail about from where the information has been taken. Bibliography does not displace exact place i.e. page number. A bibliography containing less than 20 lines need not be subdivided into categories. They could be written one below the:

A) Books:

1. Bulmer Martin, Sociological Research Methods, London, 1977, Macmillan.
2. Brislin, Cross – Culture Research Methods, New York, John Wiley and sons, 1973.

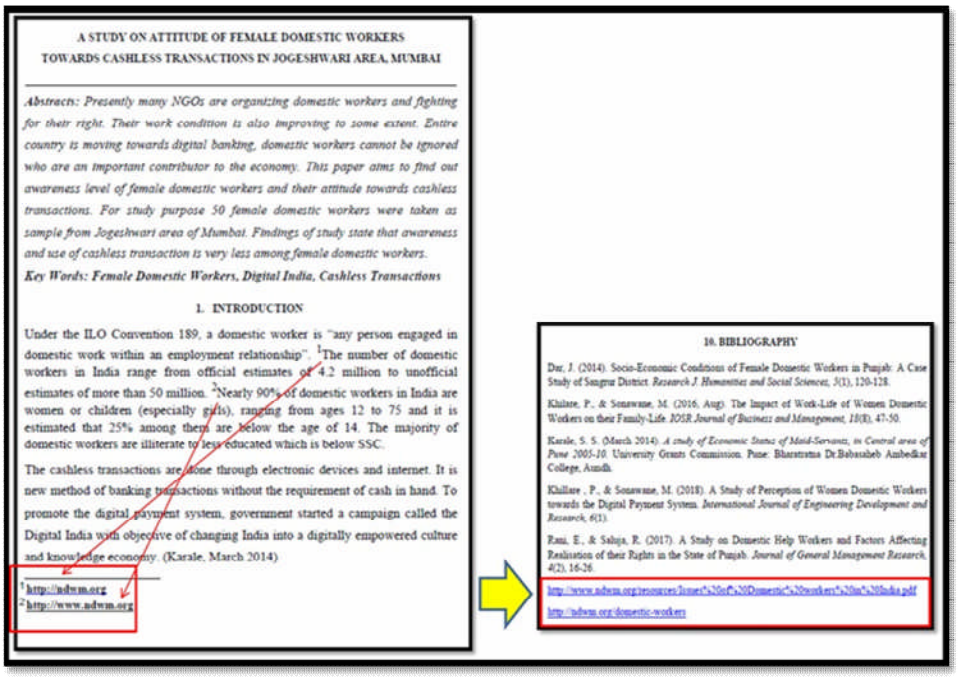
B) Reports

1. World Bank, World Development Report 1987, Washington.
2. United Nations, 1984.
3. UNCTAD, The least Development countries, 1984.

C) Journals

1. Das.D, Das.Band Mitra. S (2017). Impact of Women Participation in SHGs for Their Empowerment and Livelihood in Belbari Block of West Tripura District, Tripura, India. *International Journal of Research in Geography (IJRG). Volume 3, Issue 3.* PP 60-68.

2. Kamalanathan. K. (2016). Women empowerment and microfinance – A study on Self Help Groups in Thane district of Maharashtra. *Scholarly Research Journal for Interdisciplinary Studies. Vol 4.25. PP 2449 – 2457.*



A STUDY ON ATTITUDE OF FEMALE DOMESTIC WORKERS TOWARDS CASHLESS TRANSACTIONS IN JOGESHWARI AREA, MUMBAI

Abstracts: Presently many NGOs are organizing domestic workers and fighting for their right. Their work condition is also improving to some extent. Entire country is moving towards digital banking, domestic workers cannot be ignored who are an important contributor to the economy. This paper aims to find out awareness level of female domestic workers and their attitude towards cashless transactions. For study purpose 50 female domestic workers were taken as sample from Jogeshwari area of Mumbai. Findings of study state that awareness and use of cashless transaction is very less among female domestic workers.

Key Words: Female Domestic Workers, Digital India, Cashless Transactions

1. INTRODUCTION

Under the ILO Convention 189, a domestic worker is "any person engaged in domestic work within an employment relationship". [1]The number of domestic workers in India range from official estimates of 4.2 million to unofficial estimates of more than 50 million. [2]Nearly 90% of domestic workers in India are women or children (especially girls), ranging from ages 12 to 75 and it is estimated that 25% among them are below the age of 14. The majority of domestic workers are illiterate to less educated which is below SSC.

The cashless transactions are done through electronic devices and internet. It is new method of banking transactions without the requirement of cash in hand. To promote the digital payment system, government started a campaign called the Digital India with objective of changing India into a digitally empowered culture and knowledge economy. (Karale, March 2014)

[1] http://ndwm.org
[2] http://www.ndwm.org

10. BIBLIOGRAPHY

Dar, J. (2014). Socio-Economic Conditions of Female Domestic Workers in Punjab: A Case Study of Sangrur District. *Research J. Humanities and Social Sciences, 5*(1), 120-128.

Khilare, P., & Sonawane, M. (2016, Aug). The Impact of Work-Life of Women Domestic Workers on their Family-Life. *IOSR Journal of Business and Management, 18*(8), 47-50.

Karale, S. S. (March 2014). *A study of Economic Status of Maid-Servants, in Central area of Pune 2005-10.* University Grants Commission. Pune: Bharatratna Dr.Babasaheb Ambedkar College, Aundh.

Khillare , P., & Sonawane, M. (2018). A Study of Perception of Women Domestic Workers towards the Digital Payment System. *International Journal of Engineering Development and Research, 6*(1).

Rani, E., & Saluja, R. (2017). A Study on Domestic Help Workers and Factors Affecting Realisation of their Rights in the State of Punjab. *Journal of General Management Research, 4*(2), 16-26.

http://www.ndwm.org/resources/Issue%20of%20Domestic%20workers%20in%20India.pdf

http://ndwm.org/domestic-workers

Difference between Footnote and Bibliography

7.7 SUMMARY

This chapter highlights the process and essentials of compiling all the research oriented data and putting it in the right order and making a complete research report of it. The report can be divided into several parts and that is explained in a process format in detailed manner. Writing a research report is an art and thus every researcher needs to possess those skills and techniques. The difference between Bibliography and References has been explained as they play a vital role in every research report. There are different ways to cite the research work that has been reviewed by researchers and the same has been explained with suitable examples.

7.8 EXERCISE

Practice questions:

1. Explain the layout of research report in detail.
2. Briefly explain the essentials of research report.
3. Describe the importance of research report.
4. What is references? Explain different styles of references.
5. Elaborate the difference between footnote and bibliography with suitable examples.

Fill in the blanks

1. Table of content is also known as _____.
2. _____ consist of a brief summary framed by researcher about the past research or studies done by other researchers.
3. Questionnaire is used to collect _____ data from respondents.
4. A _____is a way of giving credit to individuals for their creative and intellectual works that is been utilized by a researcher for the work.

Answers:1. Index 2. Review of literature 3. Primary 4. Citation

True or False:

1. Footnote and Bibliography are different. – True
2. Review of literature is based on secondary data – True
3. Chapter: Analysis and interpretation consist of tables and graphs explaining primary data – True
4. Bibliography is always mentioned in the beginning of the report – False

7.9 REFERENCE

1. C. R. Kothari. Research Methodology: Methods & techniques. 2nd ed.
2. P.S.S. Sunder Rao, J. Richard.Introduction to Biostatistics and Research Methodology. 4th ed.

- https://labwrite.ncsu.edu/res/res-citsandrefs.html
- https://pitt.libguides.com/citationhelp/apa7
- https://www.slideshare.net/ShrutiMishra19/ppt-on-report-writing
- http://www.jiwaji.edu/pdf/ecourse/economics/Research%20Report.pdf

8

MODERN PRACTICES IN RESEARCH

Unit Structure

8.0 OBJECTIVE

- To understand modern methodologies used in research.
- To make out the role of companies in research.
- To recognize ethical norms in research.
- To comprehend the plagiarism checks in research.

8.1 INTRODUCTION

The role of research in several fields whether related to business or to the economy as a whole, has greatly increased in modern times. The increasingly complex nature of business and government has focused attention on the use of research in solving operational problems. Research, as an aid to economic policy, has gained added importance, both for government and business.

Usage of technology and different software has made the work of researchers more easy and interesting. Moreover, the data and studies of different researcher can be easily reviewed and re worked on it.

For instance, government's budgets rest in part on an analysis of the needs and desires of the people and on the availability of revenues to meet these needs. The cost of needs has to be equated to probable revenues and this is a field where research is most needed. Through research we can devise alternative policies and can as well examine the consequences of each of these alternatives.

8.2 ROLE OF COMPUTERS IN RESEARCH

8.2.1 MEANING

The computers are indispensable throughout the research process. The role of computer becomes more important when the research is on a large sample. Data can be stored in computers for immediate use or can be stored in auxiliary memories like floppy discs, compact discs, universal serial buses (pen drives) or memory cards, so that the same can be retrieved and easily accessible from everywhere. The usage of computer has made the job of researcher easier due to the speed in working and high level of accuracy level. The new software has made the things easier to understand and implement.

The computers assist the researcher throughout different phases of research process. There are five major phases of the research process. They are:

A. Conceptual phase
B. Design and planning phase
C. Empirical phase
D. Analytic phase
E. Dissemination phase

A. ROLE OF COMPUTER IN CONCEPTUAL PHASE:

The conceptual phase consists of formulation of research problem, review of literature, theoretical frame work and formulation of hypothesis. Computers help for searching the literatures (for review of literature) and bibliographic references stored in the electronic databases of the worldwide webs. The articles published in international journals can also be accessed through computers. There are different portals available where the researches of earlier scholars are uploaded for reference purpose. It can thus be used for storing relevant published articles to be retrieved whenever needed. This has the advantage over searching the literatures in the form of books, journals and other newsletters at the libraries which consume considerable amount of time and effort. The MS- Word also helps the researcher to put references in the required style.

B. ROLE OF COMPUTERS IN DESIGN AND PLANNING PHASE:

Design and planning phase consist of research design, population, research variables, sampling plan, reviewing research plan and pilot study. Role of Computers for Sample Size Calculation: Several software are available to calculate the sample size required for a proposed study. NCSS-PASS-GESS is such software. The standard deviation of the data from the pilot study is required for the sample size calculation. Usage of Ms-Excel can help the researcher to sort, analyse and present the data as per requirement. Basic statistical techniques like chi-square, correlation, etc and also be implemented with the add on features of Megastat.

C. ROLE OF COMPUTERS IN EMPIRICAL PHASE:

Empirical phase consist of collecting and preparing the data for analysis. The data obtained from the respondents are stored in computers as word files or excel spread sheets. This has the advantage of making necessary corrections or editing the whole layout of the tables if needed, which is impossible or time consuming incase, of writing in papers. Thus, computers help in data entry, data editing, data management including follow up actions etc. Computers also allow for greater flexibility in recording the data while they are collected as well as greater ease during the analysis of these data. In research studies, the preparation and inputting data is the most labor-intensive and time-consuming aspect of the work. Typically, the data will be initially recorded on a questionnaire or record form suitable for its acceptance by the computer. To do this the researcher in conjunction with the statistician and the programmer, will convert the data into Microsoft word file or excel spreadsheet. These spreadsheets can be directly opened with statistical software for analysis.

D. ROLE OF COMPUTERS IN DATA ANALYSIS:

This phase consists of statistical analysis of the data and interpretation of results. Many software are now available to perform the 'mathematical part of the research process i.e. the calculations using various statistical methods. The software like SPSS, Stata, JMP, SAS etc are freely available to analyse the data for different statistical techniques.

E. ROLE OF COMPUTERS IN RESEARCH DISSEMINATION:

This phase is the publication of the research study. It helps the researcher to compile the data and put it into one format and book so that a complete thesis can be formed. The research article is typed in word format and converted top or table data format (PDF) and stored and/or published in the World Wide Web.

To conclude, computers are useful tools that make the research process easier and faster with accuracy and greater reliability and fewer errors. The programmer or the computer operator should have a thorough knowledge about the abilities and limitations of the software used for better use of computers.

8.3 ETHICAL NORMS IN RESEARCH:

8.3.1 MEANING

The term research ethics refers to a wide variety of values, norms, and institutional arrangements that help constitute and regulate scientific activities. Research ethics is a codification of scientific morality in practice.

Ethics are the moral principles that a person must follow, irrespective of the place or time. Behaving ethically involves doing the right thing at the right time. Research ethics focus on the moral principles

that researchers must follow in their respective fields of research. The researcher should not undertake research misconduct such as:

- **Fabrication** - making up data or results and recording or reporting them.
- **Falsification** - manipulating research materials, or changing or omitting data or results such that the research is not accurately represented in the research record.
- **Plagiarism** - the appropriation of another person's ideas, processes, results, or words without giving appropriate credit.

Research ethics provides guidelines for the responsible conduct of research. In addition, it educates and monitors researcher / scholars conducting research to ensure a high ethical standard. The following is a general summary of some ethical principles:

1) **Honesty:** The researcher should honestly report data, results, methods and procedures, and publication status. He/she should not fabricate, falsify, or misrepresent data.

2) **Objectivity:** The researcher should strive to avoid bias in experimental design, data analysis, data interpretation grant writing, expert testimony, and other aspects of research.

3) **Carefulness:** The researcher should avoid careless errors and negligence. He/she should carefully and critically examine the work. He/she should keep good records of research activities.

4) **Openness:** The researcher should share data, results, ideas, tools, resources. Wherever, required references must be mentioned to give credit to the original work of the researcher. He/she should be open to criticism and new ideas.

5) **Respect for Intellectual Property:** The researcher should honour patents, copyrights, and other forms of intellectual property. He/she should not use unpublished data, methods, or results without permission. Researcher should give credit to the other author if his content is used and should not plagiarize.

6) **Confidentiality:** The researcher should maintain the confidentiality about the identity of respondent who respond to the research by filling up questionnaire or giving interview. As well as other records which are to be kept secret should not be disclosed to any other person.

7) **Responsible Publication:** The researcher should publish in order to advance research and scholarship, not to advance just your own career. He/she should avoid wasteful and duplicative publication.

8) **Respect for Colleagues:** The researcher should respect his/her colleagues and treat them fairly. It is necessary to maintain their motivation in carrying on research activity.

9) **Social Responsibility:** The research should strive to promote social good and prevent or mitigate social harms through research. The research should contribute to the welfare of the society.

10) **Norms and values of research:** Research is a systematic and socially organised activity governed by various specific and values. In the humanities and social sciences, involvement and interpretation are often integral parts of the research process. Different academic approaches and theoretical positions may also allow for different, but nonetheless reasonable, interpretations of the same material.

11) **Respect for third parties**: Researchers should consider and anticipate effects on third parties that are not directly included in the research. Researchers should take account of the possible negative consequences for third parties. This is particularly important when vulnerable individuals, like children and minors, are indirectly involved in the research.

8.4 PLAGIARISM IN RESEARCH:

8.4.1 DEFINITION

According to Oxford Dictionary of English, plagiarism is "the practice of taking someone else's work or ideas and passing them off as one's own." Plagiarism is presenting someone else's work or ideas as your own, with or without their consent, by incorporating it into your work without full acknowledgement. All published and unpublished material, whether in manuscript, printed or electronic form, is covered under this definition. Plagiarism may be intentional or reckless, or unintentional.

8.4.2 TYPES OF PLAGIARISM

Type of plagiarism	Definition	Severity
Global plagiarism	Presenting an entire text by someone else as your own work.	Severe
Paraphrasing plagiarism	Rephrasing someone else's ideas without citation.	Serious
Verbatim plagiarism	Directly copying a passage of text without citation.	Serious
Mosaic plagiarism	Combining text and ideas from different sources without citation.	Serious
Self-plagiarism	Reusing passages and ideas from your own previously submitted work.	Moderate
Incorrect citation	Failing to give all the necessary information in your source citation.	Moderate

Source: https://www.scribbr.com/plagiarism/types-of-plagiarism/

1. **Global Plagiarism:** It means copying entire work of someone and quoting as if you're your work. E.g.: If you get someone else to write

an essay or assignment for you, or if you find a text online and submit it as your own work, you are committing plagiarism. Because it involves deliberately and directly lying about the authorship of a work, this is one of the most serious types of plagiarism, and it can have severe consequences.

2. **Paraphrasing Plagiarism:** Paraphrasing means rephrasing a piece of text in your own words. Paraphrasing without citation is the most common type of plagiarism.

   Paraphrasing itself is not plagiarism so long as you properly cite your sources. However, paraphrasing becomes plagiarism when you read a source and then rewrite its key points as if they were your own ideas. Additionally, if you translate a piece of text from another language, you need correctly cite the original source. A translation without a source is still plagiarism, as the idea is of someone else.

3. **Verbatim Plagiarism:** Verbatim plagiarism when you directly copy text from a source and paste it into your own document without acknowledging the creator. If the structure and the majority of the words are the same as in the original, then it is verbatim plagiarism, even if you delete or change a couple of words here and there.If you want to use an author's exact words, you need to quote the original source by putting the copied text in quotation marks and including an in-text citation.

4. **Mosaic Plagiarism:** Mosaic plagiarism means copying phrases, passages and ideas from different sources and putting them together to create a new text. This includes slightly rephrasing passages while keeping many of the same words and structure as the original. This type of plagiarism requires a little more effort and is more insidious than just copying and pasting from a source, but plagiarism checkers like Turnitin can still easily detect it.

5. **Self Plagiarism:** Self-plagiarism means reusing work that you've previously submitted. Even though it's your own work, it's considered dishonest to present a paper or a piece of data as brand new when you've already gotten credit for the work.

   There are a couple of different versions of self-plagiarism. The more serious is to turn in a paper you already submitted for a grade to another class. Unless you have explicit permission to do so, this is always considered self-plagiarism.Self-plagiarism can also occur when you use ideas, phrases or data from your previous assignments. As with paraphrasing, reworking old ideas and passages is not inherently plagiarism, but you should cite your previous work to make the origins clear.

6. **Incorrect citation:** The key to avoiding plagiarism is citing your sources. You need to correctly format your citations according to the

rules of the citation style you are following. If you don't include all the necessary information or you put it in the wrong place, you could be committing plagiarism. Most styles require in-text citations plus a reference list or bibliography at the end of your paper, where you give full details of every source you cited.

To avoid plagiarism, researcher must give credit whenever he/she:

- Directly quote another person's written or spoken words. Be sure to enclose these words and/or sentences in quotations marks!
- Paraphrase another person's spoken or written words. Paraphrase means to re-write in your own words; merely reordering or substituting words is still considered plagiarism!
- Use theories, ideas, opinions, research, etc. that are not your own.
- Use historical, statistical, or scientific facts or data that are not your own.

8.5 SUMMARY

- The methodology has changed drastically over a period of time and thus this unit consists of modern techniques adopted to make a research project more genuine and useful for readers. Thus a detailed explanation of Plagiarism has been given. Different types of Plagiarism are highlighted and dos and don'ts to avoid Plagiarism have also been mentioned. The roles of computers have increased at every level while conducting research and the same has been explained in detail. Different software can be used to analyse the data, create charts and also to manage the data in proper sequence. Every researcher needs to follow certain ethics while conducting research so that research is genuine and concrete. Thus, ethical norms in research have also been captured in the unit.

8.6 EXERCISE

Practice questions:

1. Explain the role of computers in research.
2. What are ethics followed in research.
3. Write a note on Plagiarism
4. Describe different types of Plagiarism

Fill in the blanks

1. _____refers to a wide variety of values, norms, and institutional arrangements that help constitute and regulate scientific activities.
2. _____ manipulating research materials, or changing or omitting data or results.

3. SPSS software helps to _____ the research data.
4. _____ is the practice of taking someone else's work or ideas and passing them off as one's own.

Answers:1. Ethics 2. Falsification 3.Analyse 4. Plagiarism

True or False:

1. Plagiarism check is not at all important – False
2. SPSS software can be used for analysis of the data – True
3. Usage of Ms-Power Point can help the researcher to sort, analyse and present the data as per requirement. - False
4. Usage of computer has made the job of researcher very easy. – True

8.7 REFERENCES:

1. C. R. Kothari. Research Methodology: Methods & techniques. 2nd ed.
2. P.S.S. Sunder Rao, J. Richard.Introduction to Biostatistics and Research Methodology. 4th ed.

- https://labwrite.ncsu.edu/res/res-citsandrefs.html
- https://www.slideshare.net/saravananmsw/role-of-computers-in-research
- https://www.forskningsetikk.no/en/guidelines/social-sciences-humanities-law-and-theology/guidelines-for-research-ethics-in-the-social-sciences-humanities-law-and-theology/
- https://www.scribbr.com/plagiarism/types-of-plagiarism/

❖❖❖❖