



**GATE INSTITUTE OF TECHNOLOGY AND MANAGEMENT SCIENCES,
CHIGURUWADA**

MBA – FIRST SEMESTER

105-BUSINESS RESEARCH METHODS



MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

SYLLABUS

MASTER OF BUSINESS ADMINISTRATION (M.B.A)

105 –Business Research Methods

UNIT - I

Fundamentals of Research: Meaning - Scope and Significance - Types of Research - Research Process - Characteristics of Good Research - Scientific Method - Problems in Research - Identifying Research Problem: Concepts, Constructs, and Theoretical Framework.

UNIT - II

Formulation of Hypotheses: Meaning – Types of Hypotheses – Formulation of Hypotheses, Research Design: Types – Features of Good Design - Measurement Errors Techniques of Measurement - Scaling Techniques: Types of Scales - Scale Construction Techniques.

UNIT - III

Sampling Design: Meaning - Steps in Sampling - Criteria for Good Sample Design - Types of Sample Designs - Probability and Non-probability Samples. Data Collection: Types of Data - Sources - Tools for Data Collection Methods of Data Collection – Constructing Questionnaire - Pilot Study - Data Processing: Coding - Editing - Tabulation of Data - Data Analysis.

UNIT - IV

Statistical tools for Analysis: Measures of Central Tendency – Mean, Median, Mode – Measures of Dispersion– Standard Deviation and Variance -Assumptions about Parametric and Non-Parametric Tests. Parametric Tests: T test, Z test, One-way and Two-way ANOVA, Simple Linear Regression and Correlation. Non-Parametric Tests: U Test, Kruskal-Wallis, Sign test and Chi-square test.

UNIT - V

Interpretation & Report Writing: Meaning – Techniques of Interpretation- Report Writing: Significance –Report Writing: Steps in Report Writing – Layout of Report – Types of Reports - Oral Presentation- Executive Summary - Mechanics of Writing Research Report - Precautions for Writing a report- Norms for using Tables, Charts, and Diagrams - Appendix: norms for using Index and Bibliography.

UNIT-I

FUNDAMENTALS OF FRESEARCH

MEANING OF RESEARCH

- Research refers to a search for knowledge. Research is an art of scientific investigation.
- In the modern complex world every society today is faced with serious social, economic & political problems. These problems need systematic, intelligent and Practical solutions. Problem solving is technical process. It requires the accumulation of new knowledge.
- Research provides the means for accumulating knowledge & wisdom. In other words, research is a systematic effort of gathering analysis & interpretation of problems confronted by humanity. It is a thinking process and scientific method of studying a problem and finding solution. It is an in depth analysis based on reflective thinking.

DEFINITION OF RESEARCH

According to Redman and Mory, Research as a,” Systematized effort to gain new knowledge”. Some people consider research as a movement, a movement from the known to the unknown.

According to Clifford woody, Research comprises defining and redefining problems, formulating hypothesis or suggested solutions collecting, organising and evaluating data, making deductions and reaching conclusions; to determine whether they fit the formulating hypothesis.

According to William Emory Research as “any organized enquiry designed and carried out to provide information for solving a problem”

The new Oxford English Dictionary defines research is “the scientific investigation into and study of material, sources etc. in order to establish facts and the reach new conclusions”.

Advanced Leane’s Dictionary, “A careful investigation or inquiry especially through search for new facts in any branch of knowledge”

BUSINESS RESEARCH

Business research is a process of acquiring detailed information of all the areas of business and using such information in maximizing the sales and profit of the business. Such a study helps companies determine which product/service is most profitable or in demand. In simple

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

words, it can be stated as the acquisition of information or knowledge for professional or commercial purpose to determine opportunities and goals for a business.

Business research can be done for anything and everything. In general, when people speak about business research it means asking research questions to know where the money can be spent to increase sales, profits or market share. Such research is critical to make wise and informed decisions.

“The application of the scientific method in searching for the truth about business phenomena. These activities include defining business opportunities and problems, generating and evaluating ideas, monitoring performance, and understanding the business process.”

For example: A mobile company wants to launch a new model in the market. But they are not aware of what are the dimensions of a mobile that are in most demand. Hence, the company conducts a business research using various methods to gather information and the same is then evaluated and conclusions are drawn, as to what dimensions are most in-demand, This will enable the researcher to make wise decisions to position his phone at the right price in the market and hence acquire a larger market share.

 QuestionPro



SCOPE OF BUSINESS RESEARCH

- ❖ **Production Management:** The research performs an important function in product development, diversification, introducing a new product, product improvement, process technologies, choosing a site, new investment etc.
- ❖ **Personnel Management/ HR:** Research works well for job redesign, organization restructuring, development of motivational strategies and organizational development.
- ❖ **Marketing Management:** Research performs an important part in choice and size of target market, the consumer behavior with regards to attitudes, life style, and influences of the target market. It is the primary tool in determining price policy, selection of channel of distribution and development of sales strategies, product mix, promotional strategies, etc.
- ❖ **Financial Management:** Research can be useful for portfolio management, distribution of dividend, capital raising, hedging and looking after fluctuations in foreign currency and product cycles.
- ❖ **Materials Management:** It is utilized in choosing the supplier, making the decisions relevant to make or buy as well as in selecting negotiation strategies.
- ❖ **General Management:** It contributes greatly in developing the standards, objectives, long-term goals, and growth strategies.

SIGNIFICANCE OF BUSINESS RESEARCH

1. Better Decision-Making

Research provides factual data rather than assumptions, helping managers choose the best strategies, reduce uncertainty, and solve business problems effectively.

2. Identifying Opportunities and Threats

Through market analysis, consumer studies, and trend forecasting, businesses can spot new opportunities and detect potential risks early.

3. Understanding Customers

Research helps companies understand customer needs, preferences, and behaviors. This allows them to design better products, improve services, and increase customer satisfaction.

4. Improving Efficiency and Performance

By studying internal operations, research uncovers inefficiencies and suggests improvements that reduce costs and boost productivity.

5. Supporting Innovation

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Research stimulates new ideas for products, services, and business models. It helps companies stay competitive in a changing market.

6. Evaluating Business Strategies

Organizations use research to measure the impact of marketing campaigns, financial decisions, and operational strategies to see what works and what doesn't.

7. Reducing Risk

Research provides information that helps businesses avoid expensive mistakes, plan for the future, and make safer investments.

8. Enhancing Competitive Advantage

Companies that conduct regular research stay updated on industry trends, competitor actions, and market demands, giving them an edge in competition.

TYPES OF RESEARCH

I. Qualitative Research

Qualitative Research is about gathering and studying information that isn't numbers, like videos, words, or sounds, to grasp ideas, opinions, or experiences. It helps us dig deep into issues or come up with fresh research ideas. Unlike quantitative research, which deals with numbers, qualitative research focuses on understanding things more deeply rather than crushing statistics.

1. Interviews

Interviews focus on a one-on-one dialogue between a researcher and a participant, aiming to delve into profound insights, opinions, and experiences concerning a specific subject.

Features:

- **Open Exploration:** Interviews allow for open-ended and examining inquiries, enabling the exploration of intricate topics in depth.
- **Non-verbal Insight:** They provide the opportunity to observe non-verbal hints and establish rapport with the participant, fostering a deeper understanding.
- **Clarification Opportunity:** Additionally, interviews empower the researcher to seek further context and explain responses as necessary, ensuring clarity and depth in data collection.

Advantages:

- **Rich Data Yield:** Interviews get comprehensive and detailed data from the participants' viewpoints, offering valuable insights into their perspectives.

- **Personalized Interaction:** They offer a personalized and interactive research methodology, allowing for tailored inquiries based on participant responses.
- **Flexibility:** Interviews provide flexibility to adapt the questioning and focus areas based on the participant's feedback, enhancing the depth of understanding.

Disadvantages:

- **Resource Intensiveness:** Conducting and analyzing interviews can be time-consuming and resource-intensive, requiring substantial investment.
- **Potential for Bias:** There's a risk of interviewer bias or participant reactivity influencing the data, which may compromise its objectivity.
- **Generalization Challenges:** Findings from interviews may be challenging to generalize due to the limited sample size, potentially limiting broader applicability.

For example, researchers from the Indian Institute of Management in Ahmedabad engaged in in-depth interviews with rural entrepreneurs in Gujarat to separate the motivations and hurdles underlying their entrepreneurial aspirations.

2. Focus Groups

Focus groups involve small group discussions, typically comprising 6–12 participants, facilitated by a researcher to explore attitudes, perceptions, and behaviors around a specific topic.

Features:

- **Interactive Discussions:** Focus Groups encourage group interaction and dynamic discussions among participants, supporting a collaborative exploration of the topic.
- **Observational Insight:** They allow the researcher to observe non-verbal clues and group dynamics, providing additional layers of understanding.
- **Effortless Analysis:** Additionally, focus groups provide the flexibility to analyze and follow up on interesting insights, enhancing the depth of exploration.

Advantages:

- **Contextual Data Generation:** Focus Groups generate rich, contextual data through group interactions, offering insights into shared perspectives.
- **Diverse Perspectives:** They enable the researcher to gather diverse perspectives on a topic, enriching the understanding of complex phenomena.
- **Cost-effectiveness:** Focus Groups offer a cost-effective and efficient way to collect qualitative data compared to other methods, optimizing resource utilization.

Disadvantages:

- **Influence of Group Dynamics:** There's a risk that group dynamics may influence individual responses, potentially leading to conformity bias or dominant participant effects.
- **Generalization Challenges:** Findings from focus groups may be difficult to generalize due to the small sample size and the specific context of the discussion.
- **Facilitation Complexity:** Skilled facilitation is required to manage group dynamics and ensure balanced participation, adding a layer of complexity to the process.

For example, the National Institute of Design in India conducted focus group discussions with urban consumers to understand their preferences and attitudes towards sustainable fashion products.

3. Ethnographic Research

Ethnographic Research involves the in-depth study of a cultural group or community, where the researcher immerses themselves in the natural setting to observe and understand the group's behaviors, customs, and interactions.

Features:

- **Extended Engagement:** Ethnographic Research focuses on the researcher's prolonged immersion and observation within the research setting, allowing for in-depth understanding.
- **Cultural Context Focus:** It focuses on understanding the cultural context and meanings behind behaviors and practices, uncovering deeper layers of understanding.
- **Multi-method Approach:** Additionally, ethnographic research utilizes a range of data collection methods, such as participant observation, interviews, and document analysis, ensuring comprehensive data collection.

Advantages:

- **Holistic Understanding:** Ethnographic Research provides a holistic and contextual understanding of the subject, capturing the complexities of human behavior and culture.
- **Rich Data Capture:** It generates rich, detailed data that captures the variations of human behavior and cultural practices, offering insights that may not be accessible through other methods.
- **Insight Unveiling:** Ethnographic Research allows the researcher to uncover hidden insights and challenge assumptions, shedding light on complex societal phenomena.

Disadvantages:

- **Resource Intensiveness:** Conducting ethnographic research requires significant time and resources for the researcher to immerse themselves in the field, which can be challenging to sustain.
- **Potential for Bias:** There's a risk of researcher bias and subjectivity due to the close involvement with the research subject, which may influence data interpretation.
- **Generalization Challenges:** Findings from ethnographic research may be difficult to generalize due to the unique cultural context of the study, limiting broader applicability.

For example, anthropologists from the Tata Institute of Social Sciences conducted an ethnographic study of the Warli tribal community in Maharashtra, India, to understand their traditional art, rituals, and livelihood practices.

4. Website Visitor Research

Website Visitor Research involves the study of user behavior, preferences, and interactions on a website or digital platform to understand the user experience and inform design and content decisions.

Features:

- **Diverse Data Collection Methods:** Website Visitor Research utilizes various data collection methods, such as web analytics, user testing, and online surveys, to gather comprehensive insights.
- **User Journey Understanding:** It focuses on understanding the user's journey, pain points, and decision-making processes, providing valuable information for website optimization.
- **Effectiveness Evaluation:** Website Visitor Research evaluates the effectiveness of the website's design, content, and functionality in engaging users and achieving desired outcomes.

Advantages:

- **Data-driven Insights:** Website Visitor Research generates data-driven insights to optimize the website for better user engagement and conversion, boosting overall performance.
- **User Feedback Utilization:** It enables the identification of user pain points and opportunities for improvement, guiding iterative design and content enhancements.
- **Cost-effectiveness:** Website Visitor Research provides a cost-effective and scalable way to gather user feedback and behavior data compared to traditional methods.

Disadvantages:

- **Reliance on Self-reported Data:** There's a reliance on self-reported data from surveys or user testing, which may be subject to bias and inaccuracies.
- **Causality Establishment Difficulty:** Difficulty exists in establishing causality between website elements and user behavior, requiring careful interpretation of data.
- **Researcher Presence Influence:** There's a potential for changes in user behavior due to the presence of the researcher, impacting the validity of findings.

For example, the e-commerce platform Flipkart conducted extensive website visitor research, including user testing and analytics, to understand the browsing and purchasing behavior of its customers and optimize the platform's user experience.

5. Observational Studies

Observational Studies involve the systematic observation and recording of human behavior, activities, and interactions in their natural environment without direct intervention from the researcher.

Features:

- **Real-world Context:** Observational Studies focus on understanding the context and natural behaviors of the research subjects, providing insights into everyday situations.
- **Data Collection Techniques:** Techniques such as direct observation, video recording, and field notes are utilized to gather data on human behavior and interactions.
- **Non-verbal Behavior Observation:** They allow the researcher to gather data on non-verbal hints and spontaneous behaviors, providing a comprehensive view of human actions.

Advantages:

- **Insights into Real-World Behaviors:** Observational Studies provide insights into real-world behaviors and interactions that may not be captured through self-reported data.
- **Uncovering Hidden Patterns:** They enable the researcher to uncover hidden patterns and differences in human behavior, leading to a deeper understanding of the research subject.
- **Exploration of Complex Topics:** Observational Studies permit the exploration of complex, dynamic, or sensitive research topics in their natural setting.

Disadvantages:

- **Observer Effect:** There is a risk that the researcher's presence may influence the observed behavior, potentially altering the natural state of the environment.

- **Generalization Challenges:** Findings from observational studies may be difficult to generalize due to the small sample size and specific context of the observation.
- **Resource Intensiveness:** Conducting comprehensive observational studies can be time-consuming and resource-intensive, requiring careful planning and execution.

For example, researchers from the National Institute of Design conducted observational studies in rural Indian communities to understand the daily routines, social interactions, and design needs of the residents to inform the development of appropriate technologies and products.

6. Case Study

A case study involves an in-depth investigation of a single or multiple cases, such as individuals, firms, events, or phenomena, aiming to gain a comprehensive understanding of the subject.

Features:

- **Specific Context Focus:** Case Studies focus on a specific, real-world case or situation in its natural context, allowing for detailed examination.
- **Multiple Data Sources:** They utilize multiple data sources, including interviews, observations, and document analysis, to provide a comprehensive view of the case.
- **Holistic Understanding:** Additionally, case studies aim to provide a detailed, holistic, and contextual understanding of the case, capturing various dimensions of the subject.

Advantages:

- **Depth of Understanding:** Case Studies allow for a deep, refined understanding of complex phenomena, offering insights that may not be achievable through other methods.
- **Theory Development Support:** They provide insights that can inform theory development and practical applications, contributing to knowledge advancement.
- **Unique Case Exploration:** Case Studies enable the exploration of unique or rare cases that may not be accessible through other research methods, enriching the understanding of diverse contexts.

Disadvantages:

- **Limited Generalizability:** Findings from case studies may not be generalizable to other contexts or situations, limiting their broader applicability.
- **Potential for Bias:** There's a risk of researcher bias in the selection and interpretation of the case, which may influence the findings.

- **Resource Intensiveness:** Conducting a comprehensive case study can be time-consuming and resource-intensive, requiring substantial investment.

For instance, the Indian School of Business conducted a case study on the successful turnaround of Tata Motors, analyzing the company's strategies, leadership, and organizational changes that led to its revival.

II. Quantitative Research

Quantitative Research constitutes a systematic approach aimed at gathering and analyzing numerical data to quantify social phenomena and generate empirical findings. By employing statistical and mathematical techniques, this method rigorously tests theories and hypotheses concerning individuals' attitudes and behaviors. It adopts an objective and formal framework, emphasizing the precision and reliability of its outcomes.

1. Survey Research

Survey Research involves collecting data from a targeted group through surveys, questionnaires, or polls to gather insights and opinions on a specific topic.

Features:

- **Structured Questioning:** It can efficiently gather quantitative data by utilizing structured questions.
- **Multi-platform Availability:** Surveys can be conducted online, via phone, or in person, offering flexibility in data collection methods.
- **Large Sample Sizes:** This method allows for the inclusion of large sample sizes, ensuring statistically significant results.

Advantages:

- **Snapshot of Population:** Surveys provide a snapshot of opinions and behaviors within a population.
- **Diverse Participant Pool:** They enable data collection from a diverse group of participants, enhancing the breadth of insights.
- **Quick Data Collection:** Surveys offer swift data collection and analysis, facilitating timely decision-making.

Disadvantages:

- **Response Bias Risk:** Surveys are prone to potential response bias or inaccuracies due to self-reporting.
- **Limited Depth:** They offer a limited depth of insights compared to qualitative methods.

- **Response Rate Challenges:** Surveys face challenges in ensuring high response rates and representative samples.

For instance, Nielsen India surveyed to gather consumer feedback on a new product launch in the Indian market.

2. Correlational Research

Correlational Research will analyze the relationship between two or more variables to detect if changes in an individual variable are associated with changes in another.

Features:

- **Pattern Identification:** Correlational Research can detect patterns and associations between variables.
- **Statistical Analysis:** It involves using statistical analysis to measure the strength and direction of relationships.
- **Non-causal Inference:** This method can infer associations between variables but does not establish causation.

Advantages:

- **Relationship Insights:** Correlational Research provides insights into relationships between variables without manipulation.
- **Outcome Prediction:** It allows for the prediction of outcomes based on correlations observed.
- **Preliminary Data:** Correlational Studies offer valuable preliminary data for further research or hypothesis testing.

Disadvantages:

- **Causation Ambiguity:** Correlational Research cannot establish causation or determine the direction of effects between variables.
- **Third Variables Influence:** It is vulnerable to the influence of third variables or confounding factors that may affect the observed correlations.
- **Mechanism Explanation Limitation:** Correlational Studies are limited in explaining the underlying mechanisms driving the relationships observed.

For instance, researchers at IIT Delhi conducted a correlational study to explore the relationship between air pollution levels and respiratory illnesses in urban areas.

3. Causal-Comparative Research

Causal-comparative Research investigates the cause-and-effect relationship between variables by comparing groups that differ on a specific factor.

Features:

- **Impact Exploration:** Causal-comparative Research focuses on exploring the impact of an independent variable on a dependent variable.
- **Natural Group Comparison:** It involves comparing groups that naturally differ in the variable of interest.
- **Causal Relationship Determination:** This method seeks to determine the causal relationship between variables.

Advantages:

- **Cause-and-effect Examination:** Causal-comparative Research allows for the examination of cause-and-effect relationships in non-experimental settings.
- **Effects Insights:** It provides insights into the effects of variables without direct manipulation.
- **Information Value:** Causal-comparative Studies offer valuable information for understanding the impact of variables on outcomes.

Disadvantages:

- **Causality Ambiguity:** Causal-comparative Research cannot establish a direct cause-and-effect relationship due to the lack of experimental control.
- **Confounding Variables Influence:** It is vulnerable to invalid variables influencing the observed relationships between variables.
- **Effects Directionality Limitation:** Causal-comparative Studies are limited in establishing the directionality of effects between variables.

For instance, researchers at IIM Bangalore conducted a causal-comparative study to investigate the impact of leadership styles on employee motivation in Indian organizations.

4. Experimental Research

Experimental Research involves manipulating an independent variable to observe its effect on a dependent variable under controlled conditions to establish cause-and-effect relationships.

Features:

- **Participant Randomization:** Experimental Research involves random assignment of participants to experimental and control groups.
- **Variable Manipulation:** It allows for the manipulation of variables to establish causation.

- **Extraneous Variables Control:** Rigorous control over extraneous variables ensures internal validity.

Advantages:

- **Causation Establishment:** Experimental research establishes cause-and-effect relationships between variables.
- **High Internal Validity:** It provides high internal validity through experimental control.
- **Findings Replication:** It allows for the replication of studies to verify findings.

Disadvantages:

- **External Validity Limitation:** Experimental Research may lack external validity due to artificial experimental conditions.
- **Resource and Time Intensiveness:** This requires resources, time, and ethical considerations for conducting experiments.
- **Ethical Constraints:** Ethical constraints may limit the manipulation of certain variables in experimental settings.

For example, researchers at AIIMS Delhi conducted an experimental study to investigate the effectiveness of a new drug treatment on a specific medical condition in Indian patients.

5. Literature Research

Literature Research involves reviewing existing literature, documents, and sources to gather information, analyze findings, and synthesize knowledge on a specific topic.

Features:

- **Synthesis and Analysis:** Literature Research focuses on synthesizing and analyzing existing research and scholarly works.
- **Systematic Review:** It involves a systematic review of literature to identify trends, gaps, and insights.
- **Comprehensive Overview:** Literature Research provides a comprehensive overview of existing knowledge on a topic.

Advantages:

- **Broad Understanding:** It offers a broad understanding of the current state of knowledge on a topic.
- **Gap Identification:** It helps identify gaps in research and areas for further investigation.

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

- **Foundation Establishment:** Literature Research provides a foundation for developing hypotheses and research questions.

Disadvantages:

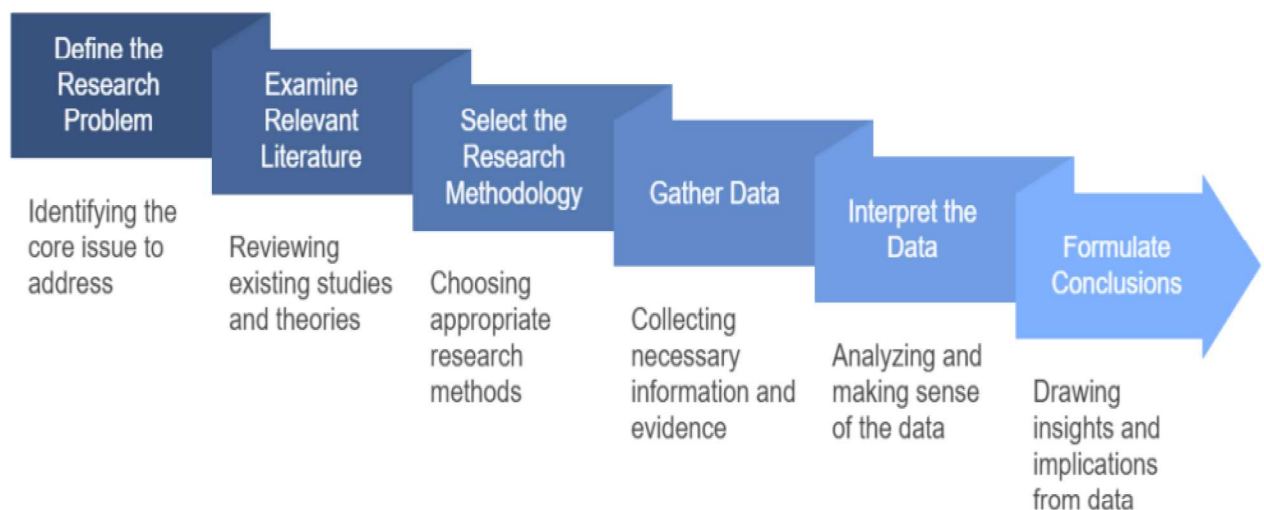
- **Temporal Limitation:** Literature Research is limited to existing literature and may not capture the most recent developments.
- **Selection Bias:** It is vulnerable to bias in the selection and interpretation of literature.
- **Comprehensiveness Challenges:** Literature Research faces challenges in ensuring the comprehensiveness and relevance of the reviewed literature.

For instance, researchers at IISc Bangalore conducted a literature review to explore the impact of climate change on agricultural practices in India, synthesizing findings from various studies and reports.

RESEARCH PROCESS

The **business research process** is a systematic approach that guides businesses in solving problems, making informed decisions, and exploring new opportunities. This process is often structured in seven key steps that ensure data is collected, analyzed, and used effectively. Let's break down each step in the process and explore the specific strategies businesses can use to achieve success at every stage.

Business Research Process



1. Define the Problem

The first and most crucial step in any business research process is identifying the problem or opportunity that needs to be addressed. This involves recognizing a gap or challenge within the business, such as customer dissatisfaction, declining sales, or the need for market expansion. Clear problem identification ensures that the research will be focused and relevant to the needs of the business.

What to do:

- **Clarify the Research Objective:** Begin by defining what you hope to achieve with the research. Are you solving a specific problem (e.g., low customer retention) or looking for new opportunities (e.g., identifying underserved markets)?
- **Set Clear Research Questions:** Frame the research around specific questions that need answering. For example, “What factors contribute to our declining sales?” or “What unmet needs do customers have in our target market?”
- **Consult Stakeholders:** Engage key stakeholders such as managers, product teams, and customers to ensure the problem is well-defined and everyone’s perspectives are considered.
- **Use a SWOT Analysis:** Perform a **SWOT analysis** (Strengths, Weaknesses, Opportunities, and Threats) to fully understand the problem from all angles.

Strategies to use:

- **Problem Definition Workshops:** Organize brainstorming sessions with relevant stakeholders.
- **Customer Feedback:** Use surveys or focus groups to directly understand the issues customers face.
- **Competitive Analysis:** Study competitors to see if they have already solved similar problems or if they present opportunities you can capitalize on.

2. Examine Relevant Literature

Once the problem is clearly defined, it’s essential to review existing literature. This means looking at past research, reports, articles, and case studies that relate to your topic. By reviewing what has already been discovered, businesses can gain valuable insights, identify gaps, and avoid duplicating efforts.

What to do:

Search Academic and Industry Research: Use resources like academic journals, industry reports, and market analysis to gather background information on your research topic.

Identify Gaps in Knowledge: Assess the existing research to find areas that haven't been explored in detail or where your business's unique circumstances might differ from existing studies.

Create a Literature Review Summary: Summarize key findings from the existing literature and use them to inform your own research objectives.

Strategies to use:

- **Use Online Databases:** Platforms like Google Scholar, JSTOR, or industry-specific databases are great for finding relevant literature.
- **Review Market Research Reports:** Many market research firms publish free executive summaries or offer paid reports with in-depth information.
- **Competitive Intelligence:** Look at what your competitors have published or what insights their case studies provide.

3. Select the Research Methodology

The research design is the blueprint for how the research will be conducted. It involves deciding which research methods and strategies will be used to collect data and ensure that the findings are reliable and valid. This stage determines whether the research will be qualitative, quantitative, or a combination of both.

What to do:

- **Decide on the Research Type:** Determine whether your research will be **qualitative** (focused on understanding opinions, experiences, or motivations) or **quantitative** (focused on measurable data, statistics, and numerical analysis).
- **Select Research Methods:** Choose the appropriate research methods, such as **surveys, interviews, focus groups, experiments, or case studies**.
- **Design Sampling Strategy:** Identify the target audience and how you will select participants. This can involve random sampling, stratified sampling, or convenience sampling based on the research's needs.
- **Set Variables and Parameters:** Define the variables you will examine (e.g., customer satisfaction, employee engagement) and how they will be measured.

Strategies to use:

- **Use Mixed Methods:** In many cases, combining qualitative and quantitative research methods (**mixed-methods design**) can provide a more comprehensive understanding of the problem.

- **Pilot Testing:** Run a small pilot of your chosen research design to iron out any potential issues with the process, such as unclear survey questions or unreliable data collection methods.

4. Gather Data

Data collection is the process of gathering relevant information based on your research design. This can be done through various methods depending on the type of research and the research questions. The quality of data collected directly impacts the quality of the conclusions drawn.

What to do:

- **Choose Your Data Collection Tools:** Depending on your design, you may need tools such as online survey platforms (e.g., SurveyMonkey, Google Forms), interview scripts, or observation checklists.
- **Implement Data Collection Methods:** Execute the plan by gathering data through the selected methods, whether it's distributing surveys, conducting interviews, or observing consumer behavior.
- **Ensure Ethical Standards:** Always follow ethical guidelines, including gaining consent from participants, ensuring anonymity, and protecting privacy.
- **Monitor the Process:** Keep track of how data is being collected to ensure consistency and address any issues that arise (e.g., low response rates or biased sample groups).

Strategies to use:

- **Online Surveys:** Utilize platforms that allow easy data collection and analysis.
- **Focus Groups and Interviews:** For more in-depth qualitative insights, consider **semi-structured interviews** or **focus group discussions**.
- **Secondary Data:** In some cases, secondary data (existing reports, public databases) can be used to supplement primary data collection.

5. Interpret the Data

Once data is collected, the next step is to analyze it. This step involves processing the raw data, applying statistical or qualitative analysis methods, and interpreting the results. Proper analysis allows businesses to uncover trends, patterns, and insights that will drive decisions.

What to do:

- **Organize the Data:** Ensure that the data is structured in a format that allows easy analysis (e.g., spreadsheets for quantitative data or coding for qualitative responses).

- **Select Analysis Techniques:** For quantitative data, use statistical tools like regression analysis, correlation analysis, or chi-square tests. For qualitative data, employ thematic analysis or content analysis to identify key themes and insights.
- **Validate Findings:** Cross-check the data with any existing literature or benchmarks to ensure that the analysis is accurate and valid.
- **Interpret the Results:** Look for patterns, correlations, or trends that answer your original research questions.

Strategies to use:

- **Use Analytical Tools:** Data analysis tools can help in precise analysis in a shorter period of time. Software like Excel, Microsoft Power BI, and Tableau can help.
- **Statistical Significance Testing:** Ensure that the results are statistically significant before drawing conclusions.
- **Data Visualization:** Use charts, graphs, and infographics to present findings in an easily digestible format.

6. Formulate Conclusions

Drawing conclusions involves interpreting the data's implications for the business. At this stage, the researcher synthesizes the findings and translates them into actionable insights. It's important to link the conclusions back to the original problem and research questions.

What to do:

- **Summarize Key Findings:** Highlight the most significant findings that directly answer the research questions.
- **Make Connections:** Relate the conclusions back to the business problem, objectives, and existing literature.
- **Assess Limitations:** Identify any limitations in the research (e.g., sample size, potential biases) and how they may affect the conclusions.
- **Recommend Actions:** Provide practical recommendations based on the research findings, guiding decision-makers on the next steps.

Strategies to use:

- **Use Clear and Concise Language:** Ensure that conclusions are easy to understand and focused on actionable outcomes.
- **Link to Business Objectives:** Make sure that the conclusions are tied to the overall business goals and strategies.

7. Present the Findings

The final step of the **business research process** is reporting the results. This step involves presenting the findings in a structured format, such as a research report or presentation, to stakeholders. Clear communication ensures that the research has a lasting impact and drives informed decision-making.

What to do:

- **Prepare a Research Report:** Write a detailed report that includes an introduction, methodology, results, and conclusions. Ensure that the report is well-organized and easily understandable.
- **Create Visual Presentations:** For stakeholders who may prefer quick insights, prepare a summary presentation using visuals like charts, graphs, and tables.
- **Present Findings to Stakeholders:** Hold meetings or presentations to communicate the results to key decision-makers and relevant teams.
- **Follow-Up:** After presenting the results, follow up with stakeholders to discuss any questions or implementation of the recommendations.

Strategies to use:

- **Executive Summary:** Start the report with an executive summary that highlights the key findings and recommendations.
- **Clear Structure:** Organize the report into clearly defined sections—introduction, methodology, analysis, findings, and conclusion—to enhance readability.

CHARACTERISTICS OF GOOD RESEARCH

1. Systematic Approach

- Research must follow a structured, methodical approach to investigate the problem.

2. Logical

- The research process should be guided by clear reasoning, ensuring conclusions follow logically from the data.

3. Empirical

- Good research relies on observed and measured phenomena. Data collection should be based on actual experience or experiments.

4. Replicable

- The research should be designed so others can repeat the study and verify results.

5. **Valid and Reliable**

- The research methods and measurements should accurately capture the phenomena under study (validity) and produce consistent results (reliability).

6. **Objective**

- It should be free from bias or personal feelings of the researcher.

7. **Clear and Precise**

- Research problems, hypotheses, and methods should be clearly defined and focused.

8. **Innovative and Original**

- It should seek to contribute new knowledge or insights.

9. **Controlled**

- It involves controlling variables to establish cause-effect relationships.

10. **Extensive Literature Review**

- It builds upon existing knowledge and identifies gaps that the research aims to fill.

11. **Ethical**

- Research must adhere to ethical guidelines, including honesty, confidentiality, and informed consent.

12. **Generalizable**

- Findings should be applicable to other contexts beyond the study sample. These characteristics help ensure research is credible, trustworthy, and useful.

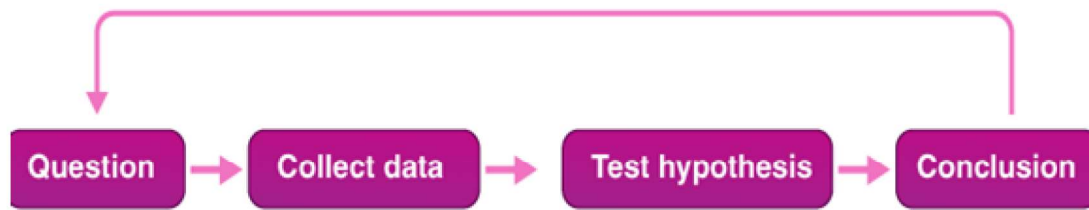
SCIENTIFIC METHODS

What is Scientific Method?

The Scientific method is a process with the help of which scientists try to investigate, verify, or construct an accurate and reliable version of any natural phenomena. They are done by creating an objective framework for the purpose of scientific inquiry and analysing the results scientifically to come to a conclusion that either supports or contradicts the observation made at the beginning.

Scientific Method Steps

The aim of all scientific methods is the same, that is, to analyse the observation made at the beginning. Still, various steps are adopted per the requirement of any given observation. However, there is a generally accepted sequence of steps in scientific methods.



1. **Observation and formulation of a question:** This is the first step of a scientific method. To start one, an observation has to be made into any observable aspect or phenomena of the universe, and a question needs to be asked about that aspect. For example, you can ask, “Why is the sky black at night? or “Why is air invisible?”
2. **Data Collection and Hypothesis:** The next step involved in the scientific method is to collect all related data and formulate a hypothesis based on the observation. The hypothesis could be the cause of the phenomena, its effect, or its relation to any other phenomena.
1. **Testing the hypothesis:** After the hypothesis is made, it needs to be tested scientifically. Scientists do this by conducting experiments. The aim of these experiments is to determine whether the hypothesis agrees with or contradicts the observations made in the real world. The confidence in the hypothesis increases or decreases based on the result of the experiments.
4. **Analysis and Conclusion:** This step involves the use of proper mathematical and other scientific procedures to determine the results of the experiment. Based on the analysis, the future course of action can be determined. If the data found in the analysis is consistent with the hypothesis, it is accepted. If not, then it is rejected or modified and analysed again.

It must be remembered that a hypothesis cannot be proved or disproved by doing one experiment. It needs to be done repeatedly until there are no discrepancies in the data and the result. When there are no discrepancies and the hypothesis is proved, it is accepted as a ‘theory’.

Scientific Method Examples

Following is an example of the scientific method:

Growing bean plants:

1. **What is the purpose:** The main purpose of this experiment is to know where the bean plant should be kept inside or outside to check the growth rate and also set the time frame as four weeks.
2. **Construction of hypothesis:** The hypothesis used is that the bean plant can grow anywhere if the scientific methods are used.
3. **Executing the hypothesis and collecting the data:** Four bean plants are planted in identical pots using the same soil. Two are placed inside, and the other two are placed outside. Parameters like the amount of exposure to sunlight, and amount of water all are the same. After the completion of four weeks, all four plant sizes are measured.
4. **Analyse the data:** While analysing the data, the average height of plants should be taken into account from both places to determine which environment is more suitable for growing the bean plants.
5. **Conclusion:** The conclusion is drawn after analyzing the data.
6. **Results:** Results can be reported in the form of a tabular form.

PROBLEMS IN RESEARCH

1. Time Management

Research is a time-consuming process that requires careful planning and scheduling. Researchers often face challenges in managing their time effectively, especially when juggling multiple research projects or other commitments. To overcome this challenge, researchers can prioritize their tasks, create a schedule, and break down their work into smaller, manageable tasks.

2. Funding

Funding is a critical aspect of research, as it provides the necessary resources to conduct the study. However, securing funding can be a significant challenge for researchers, especially those in the early stages of their careers. To overcome this challenge, researchers can explore various funding sources such as grants, fellowships, and scholarships.

3. Data Collection

Collecting reliable and valid data is a crucial aspect of research. However, researchers often face challenges in collecting data that is representative and unbiased. To overcome this challenge, researchers can use various data collection methods such as surveys, interviews, and observations, and ensure that their research instruments are standardized and valid.

4. **Data Analysis** Data analysis is a crucial aspect of research that involves organizing, interpreting, and drawing conclusions from the collected data. However, researchers often

face challenges in selecting the appropriate statistical techniques and analyzing the data accurately. To overcome this challenge, researchers can seek assistance from experts, attend workshops and training programs, and use appropriate statistical software.

5. Ethical Considerations

Research should be conducted in an ethical manner that protects the rights and well-being of the participants. Researchers often face challenges in ensuring that their research is ethical, especially when dealing with sensitive topics or vulnerable populations. To overcome this challenge, researchers can seek guidance from ethical review boards, obtain informed consent from the participants, and ensure confidentiality and privacy.

IDENTIFYING RESEARCH PROBLEM

What is a Research Problem?

A research problem is an issue or concern that an investigator presents and justifies in a research study. A problem that someone would like to research



Meaning of Research Problem

A research problem is a specific issue, difficulty, or gap in knowledge that a researcher aims to study and find a solution for. Identifying the right problem is the first and most important step in the research process because it guides the direction and purpose of the entire study.

Example: Decline in employee productivity, low customer satisfaction, or fall in sales.

Before you conceive of your project, you need to ask yourself “What is a research problem?” A research problem definition can be broadly put forward as the primary statement of a knowledge gap or a fundamental challenge in a field, which forms the foundation for research. Conversely, the findings from a research investigation provide solutions to the problem.

A research problem guides the selection of approaches and methodologies, data collection, and interpretation of results to find answers or solutions. A well-defined problem determines the generation of valuable insights and contributions to the broader intellectual discourse.

Characteristics of a Research Problem

- **Novel:** An ideal research problem introduces a fresh perspective, offering something new to the existing body of knowledge. It should contribute original insights and address unresolved matters or essential knowledge.
- **Significant:** A problem should hold significance in terms of its potential impact on theory, practice, policy, or the understanding of a particular phenomenon. It should be relevant to the field of study, addressing a gap in knowledge, a practical concern, or a theoretical dilemma that holds significance.
- **Feasible:** A practical research problem allows for the formulation of hypotheses and the design of research methodologies. A feasible research problem is one that can realistically be investigated given the available resources, time, and expertise. It should not be too broad or too narrow to explore effectively, and should be measurable in terms of its variables and outcomes. It should be amenable to investigation through empirical research methods, such as data collection and analysis, to arrive at meaningful conclusions. A practical research problem considers budgetary and time constraints, as well as limitations of the problem. These limitations may arise due to constraints in methodology, resources, or the complexity of the problem.
- **Clear and specific:** A well-defined research problem is clear and specific, leaving no room for ambiguity; it should be easily understandable and precisely articulated. Ensuring specificity in the problem ensures that it is focused, addresses a distinct aspect of the broader topic and is not vague.
- **Rooted in evidence:** A good research problem leans on trustworthy evidence and data, while dismissing unverifiable information. It must also consider ethical guidelines, ensuring the well-being and rights of any individuals or groups involved in the study.

Types of Research Problems

Across fields and disciplines, there are different types of research problems. We can broadly categorize them into three types.

1. Theoretical research problems

Theoretical research problems deal with conceptual and intellectual inquiries that may not involve empirical data collection but instead seek to advance our understanding of

complex concepts, theories, and phenomena within their respective disciplines. For example, in the social sciences, research problems may be casuist (relating to the determination of right and wrong in questions of conduct or conscience), difference (comparing or contrasting two or more phenomena), descriptive (aims to describe a situation or state), or relational (investigating characteristics that are related in some way).

Here are some theoretical research problem examples:

- Ethical frameworks that can provide coherent justifications for artificial intelligence and machine learning algorithms, especially in contexts involving autonomous decision-making and moral agency.
- Determining how mathematical models can elucidate the gradual development of complex traits, such as intricate anatomical structures or elaborate behaviors, through successive generations.

2. **Applied research problems**

Applied or practical research problems focus on addressing real-world challenges and generating practical solutions to improve various aspects of society, technology, health, and the environment.

Here are some applied research problem examples:

- Studying the use of precision agriculture techniques to optimize crop yield and minimize resource waste.
- Designing a more energy-efficient and sustainable transportation system for a city to reduce carbon emissions.

3. **Action research problems**

Action research problems aim to create positive change within specific contexts by involving stakeholders, implementing interventions, and evaluating outcomes in a collaborative manner.

Here are some action research problem examples:

- Partnering with healthcare professionals to identify barriers to patient adherence to medication regimens and devising interventions to address them.
- Collaborating with a nonprofit organization to evaluate the effectiveness of their programs aimed at providing job training for underserved populations.

How to Define a Research Problem

You might now ask “How to define a research problem?” These are the general steps to follow:

1. **Look for a broad problem area:** Identify under-explored aspects or areas of concern, or a controversy in your topic of interest. Evaluate the significance of addressing the problem in terms of its potential contribution to the field, practical applications, or theoretical insights.
2. **Learn more about the problem:** Read the literature, starting from historical aspects to the current status and latest updates. Rely on reputable evidence and data. Be sure to consult researchers who work in the relevant field, mentors, and peers. Do not ignore the gray literature on the subject.
3. **Identify the relevant variables and how they are related:** Consider which variables are most important to the study and will help answer the research question. Once this is done, you will need to determine the relationships between these variables and how these relationships affect the research problem.
4. **Think of practical aspects:** Deliberate on ways that your study can be practical and feasible in terms of time and resources. Discuss practical aspects with researchers in the field and be open to revising the problem based on feedback. Refine the scope of the research problem to make it manageable and specific; consider the resources available, time constraints, and feasibility.
5. **Formulate the problem statement:** Craft a concise problem statement that outlines the specific issue, its relevance, and why it needs further investigation.
6. **Stick to plans, but be flexible:** When defining the problem, plan ahead but adhere to your budget and timeline. At the same time, consider all possibilities and ensure that the problem and question can be modified if needed.

Key Takeaways

- A research problem concerns an area of interest, a situation necessitating improvement, an obstacle requiring eradication, or a challenge in theory or practical applications.
- The importance of research problem is that it guides the research and helps advance human understanding and the development of practical solutions.

- Research problem definition begins with identifying a broad problem area, followed by learning more about the problem, identifying the variables and how they are related, considering practical aspects, and finally developing the problem statement.
- Different types of research problems include theoretical, applied, and action research problems, and these depend on the discipline and nature of the study.
- An ideal problem is original, important, feasible, specific, and based on evidence.

Steps in Identifying a Research Problem

1. Observation of a Situation:

- Identify an issue or gap from personal experience, organization, or past studies.

2. Review of Literature:

- Study previous research to understand what is already known and where the gaps exist.

3. Discussions with Experts:

- Seek opinions from experienced researchers or professionals in the field.

4. Narrowing the Topic:

- Define the problem clearly and make it specific, measurable, and researchable.

CONCEPTS

- A concept is a general idea or understanding about a phenomenon.
- It represents the basic building blocks of a theory or research study.
- Concepts help define what is being studied and add meaning to observations.

Examples: “Motivation,” “Customer satisfaction,” “Brand loyalty,” “Profitability.”

CONSTRUCTS

- A construct is a specific, measurable form of a concept created for research purposes.
- It cannot be directly observed but can be measured through indicators or variables.
- Constructs make abstract ideas measurable.

Example:

- Concept: Job Satisfaction measured through factors like pay, reorganisation, and work environment.

THEORETICAL FRAMEWORK

A theoretical framework provides the structure and direction for a research problem by outlining the concepts, theories, and assumptions that will be used to guide the study and

interpret the findings. It acts as a blueprint, explaining the relationships between variables and the rationale behind the research. The framework is derived from existing theories and helps researchers focus the study by identifying key concepts and narrowing the scope of the data collected.

How a theoretical framework addresses a research problem

- **Provides a foundation:** It establishes the theoretical basis for the study, justifying why the research problem is important to investigate.
- **Defines the scope:** It helps identify and define the key variables and concepts, guiding the researcher to focus on specific relationships and phenomena.
- **Guides research design:** The framework informs the development of research questions, hypotheses, and methodology, ensuring the study is coherent and logical.
- **Facilitates data interpretation:** It provides a lens through which to analyze and interpret the collected data, helping to understand the meaning behind the findings.
- **Addresses gaps in knowledge:** By grounding the research in existing theory, it can help identify gaps in current knowledge, which the study then aims to fill.
- **Organizes complex phenomena:** For complex topics, a theoretical framework helps organize the research questions, hypotheses, and findings into a structured and understandable format.

Conclusion

Identifying a research problem involves defining what to study and why. Concepts provide the basic ideas, constructs make them measurable, and the theoretical framework.

UNIT-II
FORMULATION OF HYPOTHESIS

WHAT IS A HYPOTHESIS?



A hypothesis is a precise, testable statement of what the researcher predicts will be the outcome of the study. It is stated at the start of the study.

This usually involves proposing a possible relationship between **two variables**: the independent variable (what the researcher changes) and the dependent variable (what the research measures).

A fundamental requirement of a hypothesis is that it can be tested against reality, and can then be supported or rejected.

Glossary of terms | © www.WorksheetsPlanet.com | All rights reserved



MEANING OF HYPOTHESIS

A hypothesis is an assumption that is made based on some evidence. This is the initial point of any investigation that translates the research questions into predictions. It includes components like variables, population and the relation between the variables. A research hypothesis is a hypothesis that is used to test the relationship between two or more variables.

DEFINITION

According to John W “It is a shrewd guess or inference that is formulated and provisionally adopted to explain observed facts or conditions and to guide in further investigation.”

According to James E. Greighton “It is a tentative supposition or provisional guess which seems to explain the situation under observation.”

“A hypothesis can be defined as a tentative explanation of the research problem, a possible outcome of the research, or an educated guess about the research outcome”.

According to Goode and Hatt “A hypothesis is a proposition which can be put to test to determine its validity”.

“Hypotheses are single tentative guesses, good hunches – assumed for use in devising theory or planning experiments intended to be given a direct experimental test when possible”.

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

NATURE OF HYPOTHESIS

The hypothesis is a clear statement of what is intended to be investigated. It should be specified before research is conducted and openly stated in reporting the results.

This allows to:

- Identify the research objectives.
- Identify the key abstract concepts involved in the research.
- Identify its relationship to both the problem statement and the literature review.
- A problem cannot be scientifically solved unless it is reduced to hypothesis form.
- It is a powerful tool of advancement of knowledge, consistent with existing knowledge and conducive to further enquiry.
- It can be tested – verifiable or falsifiable.
- Hypotheses are not moral or ethical questions.
- It is neither too specific nor too general.
- It is a prediction of consequences.
- It is considered valuable even if proven false.

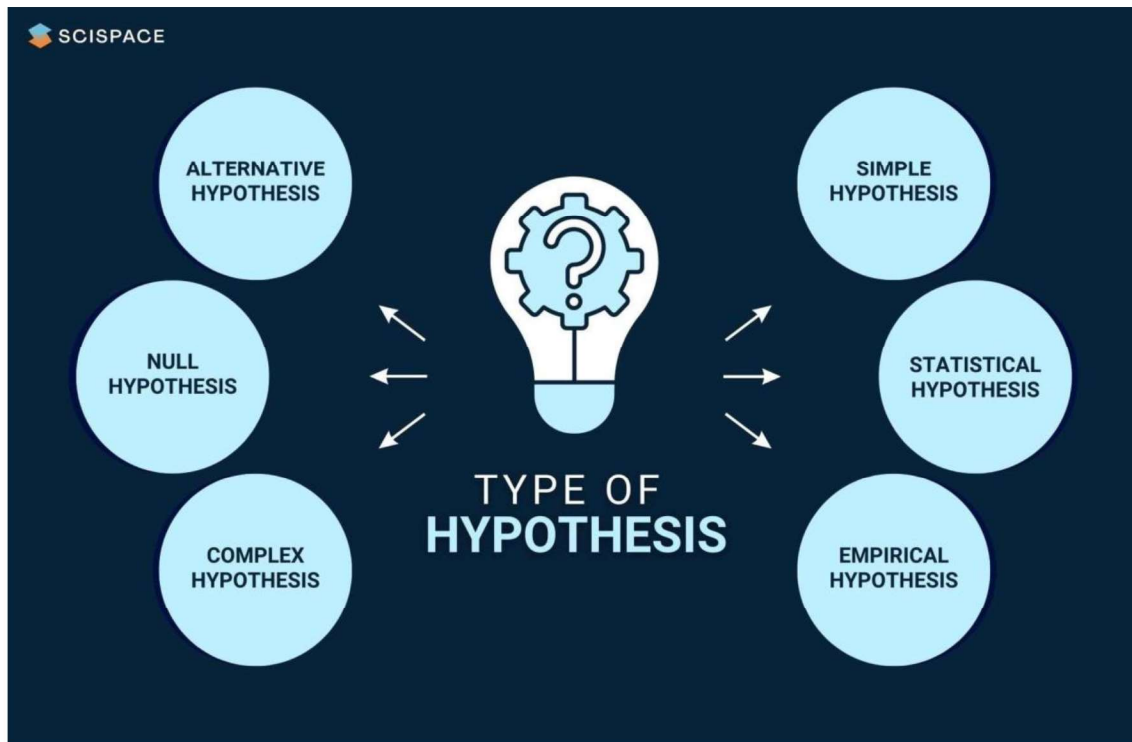
CHARACTERISTICS OF A GOOD HYPOTHESIS

Writing a hypothesis is essential as it can make or break your research for you. That includes your chances of getting published in a journal. So when you're designing one, keep an eye out for these pointers:

- A research hypothesis has to be simple yet clear to look justifiable enough.
- It has to be testable — your research would be rendered pointless if too far-fetched into reality or limited by technology.
- It has to be precise about the results — what you are trying to do and achieve through it should come out in your hypothesis.
- A research hypothesis should be self-explanatory, leaving no doubt in the reader's mind.

- If you are developing a relational hypothesis, you need to include the variables and establish an appropriate relationship among them.
- A hypothesis must keep and reflect the scope for further investigations and experiments.

TYPES OF HYPOTHESIS



Some would stand by the notion that there are only two types of hypotheses: a Null hypothesis and an Alternative hypothesis. While that may have some truth to it, it would be better to fully distinguish the most common forms as these terms come up so often, which might leave you out of context.

Apart from Null and Alternative, there are Complex, Simple, Directional, Non-Directional, Statistical, and Associative and casual hypotheses. They don't necessarily have to be exclusive, as one hypothesis can tick many boxes, but knowing the distinctions between them will make it easier for you to construct your own.

1. Null hypothesis (H₀):

A null hypothesis proposes no relationship between two variables. Denoted by H₀, it is a negative statement like “Attending physiotherapy sessions does not affect athletes' on-field performance.” Here, the author claims physiotherapy sessions have no effect on on-field performances. Even if there is, it's only a coincidence.

2. Alternative hypothesis (H1):

Considered to be the opposite of a null hypothesis, an alternative hypothesis is denoted as H1 or Ha. It explicitly states that the dependent variable affects the independent variable. A good alternative hypothesis example is “Attending physiotherapy sessions improves athletes' on-field performance.” or “Water evaporates at 100°C.” The alternative hypothesis further branches into directional and non-directional.

Directional hypothesis: A hypothesis that states the result would be either positive or negative is called directional hypothesis. It accompanies H1 with either the ‘<’ or ‘>’ sign.

Non-directional hypothesis: A non-directional hypothesis only claims an effect on the dependent variable. It does not clarify whether the result would be positive or negative. The sign for a non-directional hypothesis is ‘≠.’

3. Simple hypothesis:

A simple hypothesis is a statement made to reflect the relation between exactly two variables. One independent and one dependent. Consider the example, “Smoking is a prominent cause of lung cancer.” The dependent variable, lung cancer, is dependent on the independent variable, smoking.

4. Complex hypothesis:

In contrast to a simple hypothesis, a complex hypothesis implies the relationship between multiple independent and dependent variables. For instance, “Individuals who eat more fruits tend to have higher immunity, lesser cholesterol, and high metabolism.” The independent variable is eating more fruits, while the dependent variables are higher immunity, lesser cholesterol, and high metabolism.

5. Associative and casual hypothesis:

Associative and casual hypotheses don't exhibit how many variables there will be. They define the relationship between the variables. In an associative hypothesis, changing any one variable, dependent or independent, affects others. In a casual hypothesis, the independent variable directly affects the dependent.

6. Empirical hypothesis:

Also referred to as the working hypothesis, an empirical hypothesis claims a theory's validation via experiments and observation. This way, the statement appears justifiable and different from a wild guess.

Say, the hypothesis is “Women who take iron tablets face a lesser risk of anemia than those who take vitamin B12.” This is an example of an empirical hypothesis where the

researcher the statement after assessing a group of women who take iron tablets and charting the findings.

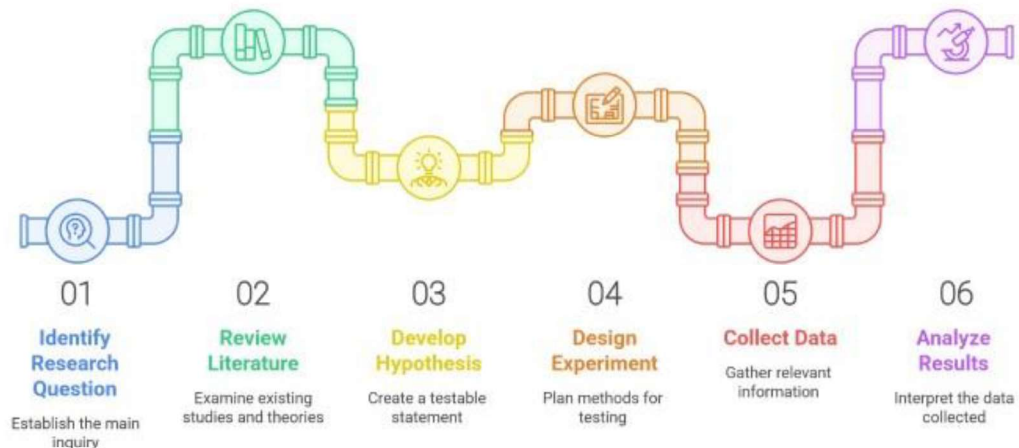
7. Statistical hypothesis:

The point of a statistical hypothesis is to test an already existing hypothesis by studying a population sample. Hypothesis like “44% of the Indian population belong in the age group of 22-27.” leverage evidence to prove or disprove a particular statement.

FORMULATION OF HYPOTHESIS



Hypothesis Formulation Process



<https://www.6sigma.us>

Step 1: Identify the Research Question

- Start by selecting a topic of interest within psychology. It could be an area of human behaviour, cognition, or emotion.
- Define the specific aspect or problem you want to investigate. The research question should be clear and concise.

Step 2: Review Existing Literature

- A thorough literature review helps you understand the current state of knowledge on your chosen topic. It allows you to identify what is already known and what gaps exist in the literature.
- Identifying gaps in existing research can inspire your hypothesis. For example, you may notice that there is limited research on the relationship between social media usage and self-esteem among adolescents. This gap could lead you to formulate a hypothesis related to this topic.

Step 3: Choose Variables

- Variables are the factors or characteristics that can change or vary in a research study. In psychology, variables can include behaviors, attitudes, emotions, and more.
- Select the independent and dependent variables that are relevant to your research question. The independent variable is the one you manipulate or examine to see its effect on the dependent variable.

Step 4: Formulate the Hypothesis

- Write a clear and specific statement predicting the relationship between the chosen variables. Your hypothesis should be concise and straightforward, leaving no room for ambiguity.
- Ensure your hypothesis is testable and falsifiable. In other words, it should be possible to design a study that can either support or refute your hypothesis based on empirical evidence.
- Using the example of social media usage and self-esteem among adolescents, your hypothesis might be: "Increased daily use of social media is negatively correlated with self-esteem levels in adolescents."

Step 5: Determine the Type of Hypothesis

- Consider the nature of your research and whether you are testing for a relationship, absence of a relationship, or an alternative outcome.

- If your goal is to explore a potential relationship between variables, you would formulate a research hypothesis (H1). If you want to investigate whether there is no significant relationship, you would formulate a null hypothesis (H0). The alternative hypothesis (Ha) would be used when you are proposing a specific alternative to the null hypothesis.
- In our example, if you expect that social media usage and self-esteem are related, your research hypothesis (H1) would be: "Increased daily use of social media is negatively correlated with self-esteem levels in adolescents." The null hypothesis (H0) would state that there is no significant correlation between these variables.

Step 6: Refine and Revise

- It's essential to refine and revise your hypothesis based on feedback and further research. A well-crafted hypothesis evolves as your understanding deepens.
- Throughout your research process, you may encounter new information or data that necessitates adjustments to your hypothesis. Being open to revision ensures that your hypothesis remains aligned with your research findings.

RESEARCH DESIGN

A research design is the plan or framework used to conduct a research study. It involves outlining the overall approach and methods that will be used to collect and analyze data in order to answer research questions or test hypotheses. A well-designed research study should have a clear and well-defined research question, a detailed plan for collecting data, and a method for analyzing and interpreting the results. A well-thought-out research design addresses all these features.

What is research design:

Research design is the framework of research methods and techniques chosen by a researcher to conduct a study. The design allows researchers to sharpen the research methods suitable for the subject matter and set up their studies for success.

Creating a research topic explains the type of research (experimental, survey research, correlational, semi-experimental, review) and its sub-type (experimental design, research problem, descriptive case-study).

There are three main types of designs for research:

- Data collection
- Measurement
- Data Analysis

TYPES OF RESEARCH DESIGN

A research design is essential to systematically investigate, understand, and interpret phenomena of interest. Let's look at different types of research design and research design examples.

Types of Research Designs:

1. Quantitative Research Design:

- a. Descriptive Research Design**
- b. Correlational Research Design**
- c. Experimental Research Design**
- d. Quasi-Experimental Research Design**

2. Qualitative Research Design:

- a. Phenomenological Research Design**
- b. Grounded Theory Research Design**
- c. Ethnographic Research Design**
- d. Case Study Research Design**

Broadly, research design types can be divided into qualitative and quantitative research.

Qualitative research is subjective and exploratory. It determines relationships between collected data and observations. It is usually carried out through interviews with open-ended questions, observations that are described in words, etc.

Quantitative research is objective and employs statistical approaches. It establishes the cause-and-effect relationship among variables using different statistical and computational methods. This type of research is usually done using surveys and experiments.

Qualitative research vs. Quantitative research

Qualitative research	Quantitative research
Deals with subjective aspects, e.g., experiences, beliefs, perspectives, and	Measures different types of variables and describes frequencies, averages, correlations,

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

concepts.	etc.
Deals with non-numerical data, such as words, images, and observations.	Tests hypotheses about relationships between variables. Results are presented numerically and statistically.
In qualitative research design, data are collected via direct observations, interviews, focus groups, and naturally occurring data. Methods for conducting qualitative research are grounded theory, thematic analysis, and discourse analysis.	Quantitative research design is empirical. Data collection methods involved are experiments, surveys, and observations expressed in numbers. The research design categories under this are descriptive, experimental, correlational, diagnostic, and explanatory.
Data analysis involves interpretation and narrative analysis.	Data analysis involves statistical analysis and hypothesis testing.
The reasoning used to synthesize data is inductive.	The reasoning used to synthesize data is deductive.
Typically used in fields such as sociology, linguistics, and anthropology.	Typically used in fields such as economics, ecology, statistics, and medicine.
Example: Focus group discussions with women farmers about climate change perception.	Example: Testing the effectiveness of a new treatment for insomnia.

Qualitative research design types

The following will familiarize you with the research design categories in qualitative research:

- **Grounded theory:** This design is used to investigate research questions that have not previously been studied in depth. Also referred to as exploratory design, it creates

sequential guidelines, offers strategies for inquiry, and makes data collection and analysis more efficient in qualitative research.

Example: A researcher wants to study how people adopt a certain app. The researcher collects data through interviews and then analyzes the data to look for patterns. These patterns are used to develop a theory about how people adopt that app.

- **Thematic analysis:** This design is used to compare the data collected in past research to find similar themes in qualitative research.

Example: A researcher examines an interview transcript to identify common themes, say, topics or patterns emerging repeatedly.

- **Discourse analysis:** This research design deals with language or social contexts used in data gathering in qualitative research.

Example: Identifying ideological frameworks and viewpoints of writers of a series of policies

Quantitative research design types

Note the following research design categories in quantitative research:

- **Descriptive research design:** This quantitative research design is applied where the aim is to identify characteristics, frequencies, trends, and categories. It may not often begin with a hypothesis. The basis of this research type is a description of an identified variable. This research design type describes the “what,” “when,” “where,” or “how” of phenomena (but not the “why”).

Example: A study on the different income levels of people who use nutritional supplements regularly.

- **Correlational research design:** Correlation reflects the strength and/or direction of the relationship among variables. The direction of a correlation can be positive or negative. Correlational research design helps researchers establish a relationship between two variables without the researcher controlling any of them.

Example: An example of correlational research design could be studying the correlation between time spent watching crime shows and aggressive behavior in teenagers.

- **Diagnostic research design:** In diagnostic design, the researcher aims to understand the underlying cause of a specific topic or phenomenon (usually an area of

improvement) and find the most effective solution. In simpler terms, a researcher seeks an accurate “diagnosis” of a problem and identifies a solution.

Example: A researcher analyzing customer feedback and reviews to identify areas where an app can be improved.

- **Explanatory research design:** In explanatory research design, a researcher uses their ideas and thoughts on a topic to explore their theories in more depth. This design is used to explore a phenomenon when limited information is available. It can help increase current understanding of unexplored aspects of a subject. It is thus a kind of “starting point” for future research.

Example: Formulating hypotheses to guide future studies on delaying school start times for better mental health in teenagers.

- **Causal research design:** This can be considered a type of explanatory research. Causal research design seeks to define a cause and effect in its data. The researcher does not use a randomly chosen control group but naturally or pre-existing groupings. Importantly, the researcher does not manipulate the independent variable.

Example: Comparing school dropout levels and possible bullying events.

- **Experimental research design:** This research design is used to study causal relationships. One or more independent variables are manipulated, and their effect on one or more dependent variables is measured.

Example: Determining the efficacy of a new vaccine plan for influenza.

FEATURES OF GOOD RESEARCH DESIGN

A proper design sets your study up for success. Successful research studies provide insights that are accurate and unbiased. You’ll need to create a survey that meets all of the main characteristics of a design. There are four key characteristics:

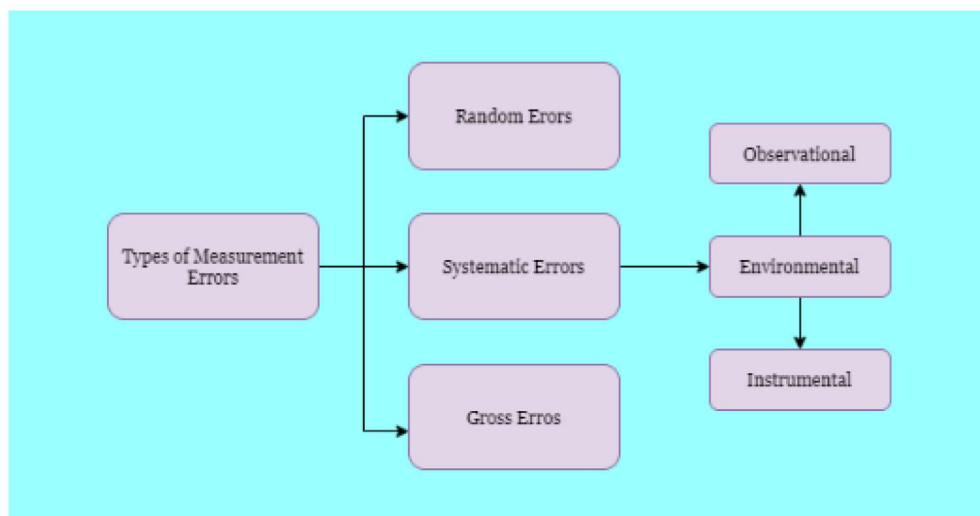


- **Neutrality:** When you set up your study, you may have to make assumptions about the data you expect to collect. The results projected in the research should be free from research bias and neutral. Understand opinions about the final evaluated scores and conclusions from multiple individuals and consider those who agree with the results.
- **Reliability:** With regularly conducted research, the researcher expects similar results every time. You'll only be able to reach the desired results if your design is reliable. Your plan should indicate how to form research questions to ensure the standard of results.
- **Validity:** There are multiple measuring tools available. However, the only correct measuring tools are those which help a researcher in gauging results according to the objective of the research. The questionnaire developed from this design will then be valid.
- **Generalization:** The outcome of your design should apply to a population and not just a restricted sample. A generalized method implies that your survey can be conducted on any part of a population with similar accuracy.

MEASUREMENT ERRORS

Measurement errors, also called observational errors, are defined as the difference between the actual response acquired and the measured response value. In this case, the actual response value is the average of the infinite number of measurements, while the measured response value is the accurate value.

Classification of Measurement Errors



The measurement errors can be classified into three different kinds, and they are

Random Errors

When repeated measurements of value are taken, the inconsistencies in the values account for the so-called random errors. They are always present within the instrument. They occur with fluctuations in the values after each measurement. They can be reduced by taking the average of multiple numbers of measurements.

Systematic Errors

These are not determined by chance but occur due to inaccuracies that are inherent in the system. They are also defined as the error with a non-zero mean, the effect of which is not reduced when the average of the values is taken into consideration. They are sometimes referred to as statistical bias. In general, they are constant and are predictable w.r.t. the true value.

The causes of systematic errors are due to the inappropriate calibration of the instruments or imperfect methods of observation, or the interference of the environment with the measurement process. An example of these errors includes imperfect zeroing of the instrument under study.

Systematic measurement errors are also classified as sampling errors and non-sampling errors.

Sampling Errors: Non-representative sample falls under this category.

Non-Sampling Errors: They include

- Paradigm Error: A scientific approach to the study of the measurable phenomenon.
- Researcher Bias: A researcher is keen to confirm the particular theory which has been devised by them, which then influences the decisions.
- Participant Bias: Participants are influenced by social desirability, supporting or opposing a particular opinion, etc.
- Reliability and validity of measurement tool.

Systematic errors can also be classified into observational errors, instrumental errors and environmental errors.

Gross Error

A gross error arises mainly due to human mistakes, or it can also be said to be a physical error. For instance, a person takes the wrong measurement or reading while using an instrument. Incorrect data is recorded, and this results in a gross error. It can be avoided by

being careful and making sure that the reading that is taken is correct. Additionally, two or more readings can be taken of the measurement quantity.

Error Interval

It is the interval that defines the boundary for the actual value to lie on. For example, consider an experiment that gave you the final reading, say 10 A. If the precision is to be taken into account, say some 2%, then it means the actual value lies between $(10-0.02)$ and $(10+0.02)$ è between 9.98 and 10.02.

The estimation of an error interval gives the absolute error associated with the instrument.

Type A and Type B Evaluation of Uncertainty

Type A measurement takes into account that the knowledge of an input quantity is taken only after considering repeated measured values. This process is called type A evaluation of uncertainty. Here, we consider repeated values for measurement in input, or in other words, we consider the Gaussian distribution.

Thus, there is an expectation which is equal to the average of the measured response values. The standard deviation = the average of all the standard deviations. If we evaluate uncertainty, taking into account only a few measured response values, it can be considered as a t-distribution.

On the other hand, type B measurement takes into account scientific judgement or other information concerning the possible values of the quantity. We can term this as a type B evaluation of uncertainty. Here, the information that is available to us is often the value that lies in a specified interval. Here, the concept of a rectangular probability distribution with limits is used.

Statistical Methods of Assessing Measurement Error

There are certain methods that are adopted to assess measurement error, which include Standard Error of Measurement (SEM): It tells about the deviations or true values of how an instrument, when used multiple times, produces the desired output.

Coefficient of Variation (CV): It is defined as how the values vary on repeated measurements. If the CV is low in value implies the results are closer to the true value.

Limits of Agreement (LOA): It gives the estimate of the interval where a proportion of the differences lie between the measurements.

Ways to Minimise Errors

- Use instruments of higher precision.
- Improve the experimental techniques.

- Adjust the zero of the instruments correctly.
- Take the value of the reading by standing straight to the instrument and not from the sides to avoid parallax errors.
- Repeat the experiment several times and take its algebraic mean for a closer result.
- Take care of the environment, if possible.
- Take the measurements carefully in order to avoid gross errors.

Other Types of Errors

There are various types of errors that can happen in our common day-to-day life. Some of them are

Absolute Error: It is defined as the amount of error in the measurement. For example, we have a sack whose actual weight is said to be 100 kg. Suppose in a weighing machine, the sack weighs out to be 150 kgs, then there is a 50 kg error difference associated with the weighing machine. This associated error of the machine is referred to as an absolute error.

Greatest Possible Error: It is the error which is defined to be one-half of a measuring unit.

Instrument Error: It is the error that is associated with the instrument. This simply tells the inaccuracy of the instrument.

Operator Error: The errors that are caused by the operator is called operating error. For e.g., in an experiment to be conducted in the lab, a man notes the voltmeter to read 5 volts, where it actually was 4 V, then the error caused by the measurement would be due to the man who observed the value. Thus, such type of error is commonly referred to as operator error. They are also called personal errors.

Measurement Location Error: The errors that are caused by the instrument that is kept at a location in which it was not bound to be kept is called a measurement location error. For example, take the case of a thermometer, which is recommended to be kept away from the sun. Suppose the thermometer is placed within the reach of sunlight, and then after the values noted using that particular thermometer, there are chances of the error to occur. Such cumulative errors are broadly classified under this category.

Parallax Error: These errors happen due to wrong sides of measurement being taken. Always take readings by standing straight in front of the instrument and not from its sides.

External Error: The errors that are caused due to external factors like wind, environment, etc., contribute to external errors.

Percentage Error: The error that is defined as the ratio of the difference between the actual value and the measured value to the actual value is called a percentage error.

Relative Error: It is defined as the ratio of the absolute error to the measured value.

MEASUREMENT TECHNIQUES

Measurement is the process observing and recording the observations that are collected as part of a research effort. Measurement is the foundation of all scientific investigation. It may be defined as the assignment of numbers to characteristics of objects or events according to rules. It is important to note that it is the characteristics of objects or events that are measured and not the objects or events themselves. Measurement defines “measurement is the assignment of numbers to objects” it means measuring the personality characteristics by assigning a number” (a score on the test) to an object (a Person). For example, we can measure buyers preference, income and attitude and products’ wetness. A business man engaged in marketing the products is interested in measuring the market potential for a new product, buyer’s attitude, perceptions or preferences towards a new brand. This measurement process gives scope for providing meaningful information for decision making. There are two basic kinds of data Non-Metric (qualitative data) and metric (quantitative).

Non-Metric (Qualitative data)

Non-metric data are attributes, characteristics or categorical properties that identify or describe subject. Non metric data describe differences in type or kind by indicating the presence or absence of a characteristic or property. Many properties are discrete in that by having a particular feature, all other features are excluded’; for example, if one is male, one cannot female. There is no ‘amount’ of gender, just the state of being male or female.

In contrast metric data metrically measured variables reflect relative quantity or degree. Measurements are made so that subjects may be identified as differing in amount or degree.

Metrically measured variables reflect relative quantity or degree. Metric measurements are appropriate for cases involving amount or magnitude, such as height, weight, temperature, rainfall, etc.

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Type	Basic empirical operation	Typical usage	Descriptive	inferential
1. Nominal	Determination of equality (0,1,2,3,.....,9)	Classification: Male-female, purchase, non-purchase, team-A Team b.	Percentage, mode	Chi – Square, Binominal test
2. ordinal	Determination of greater or less (0 < 1 < 2.....<9	Ranking : preference data, market position, attitude measures many psychological measures	Median	Rank – order correlation
3. Interval	Determination of equality of intervals (2-1=7-6)	Index number attitude measures,	Mean, range, standard deviation	Product-moment correlation, T-test, Factor analysis.
4. Ratio	Determination of equality of ratios (2/4=4/8)	Sales, units, produced, number of customer, costs, age	Geometric mean	Co-efficient of variation

Understanding of the different types of measurement scales is important for two reasons: one, to avoid the incorrect use of non metric data as metric data and vice versa and two, to determine which multivariate techniques are the most applicable to the data.

Components of measurement

Ideally speaking, there should be only one component of a measurement and this component should be a direct reflection of the characteristic being measured. Unfortunately, such a situation seldom exists. Very often, a measurement consists of not one but two components- one representing the influence of the characteristic being measured and the other representing the influence of those characteristics which the researcher is not interested in measuring but which still creep in against his wishes. These other characteristics are as follows:

- Additional stable characteristics of the object or event, for example the respondents tendency to give only favorable responses independent of his true feelings.
- Short – term characteristics of the objects, for example fatigue, health, hunger, and emotional state.
- Situational characteristics, for example, the presence or absence of some of the person or of location under which the measurement is taken.

- Characteristics of the measurement process, for example, sex, age, ethnic background, and style of dress for the interviewer or the method of interviewing- telephone, mail, personal interview, etc.
- Characteristics of the measuring instrument, for example, unclear instructions, ambiguous questions, confusing terms, omitted questions, etc.
- Characteristics of the response process, for example, mistakes caused by checking a wrong response.
- Characteristics of the analysis, for example, mistake caused by wrong coding, tabulating, etc.

Six out of these seven characteristics (2 to 7) give rise to variable errors in a measurement, i.e., errors which occur randomly each time something is measured. First characteristic gives rise to a systematic error, i.e., error that occurs in a consistent manner each time something is measured. It is also called bias.

Thus, the general situation is: $M = C + VE + SE$,

where M stands for the measurement, C stands for the characteristic being measured, VE stands for variable errors and SE stands for systematic errors.

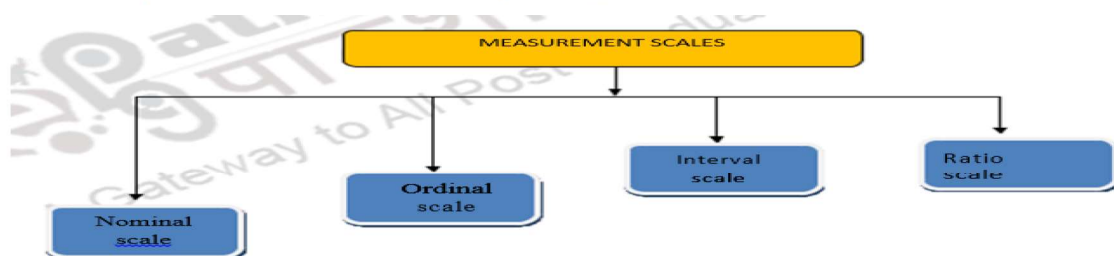
Accuracy of measurement

Accuracy of measurement depends upon the extent to which it is free from systematic and variable errors. Freedom from variable errors is known as the reliability of a measurement and freedom from systematic errors is known as the validity of a measurement. Reliability and validity are thus, two essential criteria of every measure and it is important for every researcher to know how they are measured.

SCALING TECHNIQUES

TYPES OF SCALES

Scale is a yardstick like any instrument for finding length, weight, volume and the like. The scale does not possess the properties associated with most physical measures. The scales can be generally classified into following major categories.



1.Nominal scale

In nominal scale, numbers are used to identify or categories objects or events. For example, the population of a town may be classified according to sex into “male” and Female or according to religion into “Hindus”, “Muslims”, Christians”. In marketing research nominal scales are needed to measure brands, store types, sales territories, geographical locations, heavy versus light users, working versus non-working women and brand awareness versus non-awareness. The following table gives an idea regarding the use of nominal scale.

Reasons for purchase	Male		Female		Total	
	No.	percentage	No.	percentage	No.	percentage
Weight	10	20	28	56	38	38
Speed	31	62	10	20	41	41
Miscellaneous	9	18	12	24	21	21
Total	50	100	50	100	100	100

As it could be seen from the illustration, males are concerned with speed and females are concerned with weight. Nominal scales can be developed from a response to a question Does your car have chakra tyres? The responses may be yes, No and don't know. For nominally scaled data, statistical analysis such as mode, percentages, the binomial test and chi-squared test can be use. A mean or median cannot be calculated.

2.Ordinal Scale

The ordinal scale indicates the relative position of two or more objects or some characteristics. The consumers are asked to rank preference for several brands, flavors or package designs. The measures of such preference are ordinal in nature. One may rank two or more households according to their annual income. Suppose we have five households with annual incomes as shown below:

Annual Income (Rs.)	Households				
	A	B	C	D	E
	6,000	10,000	4,800	12,000	11,000

The house hold with highest income is given o.1 and the next to it No.2 ans so, on then the following order will emerge.

Households	Order of households on the basis of annual income
A	IV
B	111
C	V
D	1
E	II

The marketing data involves ordinal measurement. Most data collected by the process of interrogation have ordinal properties. For example, the measurement of attitude, opinion preference and perception involves 'greater than' or 'less than' Judgments.

3.Interval scale:

The interval scale has all characteristics of the ordinal scale and in addition, the units of measure or intervals between successive positions are equal. For example, a research scaled brands A, B and C on an interval scale regarding the buyers' degree of liking of the brands. Brand a receives the highest liking score of 6 B receives 3 and C receives 2. First the linking for brand A is more favorable than that for brand B. Second the degree of liking between A and B is three times greater than the liking between B and C.

The statistical tools used to analyzer the interval data are range, mean, standard deviation and the like. Interval scales are frequently used in commercial marketing research studies, especially in collecting attitudinal and overall brand rating information.

4. Ratio Scales

The ratio scales have all the properties of nominal scale, ordinal scale and intervals scale. They have order, distance and unique origin. Once a ration scale has been established, its

value can be transformed only by multiplying each value by a constant. Thus, on a ratio scale, a score of 90 is twice that of 45.

Errors

Re-zero the instrument if possible, or measure the displacement of the zero reading from the true zero and correct any measurements accordingly. ... Parallax (systematic or random) – This error can occur whenever there is some distance between the measuring scale and the indicator used to obtain a measurement. The measurement of an amount is based on some international standards which are completely accurate compared with others. Generally, measurement of any quantity is done by comparing it with derived standards with which they are not completely accurate. Thus, the errors in measurement are not only due to error in methods, but are also due to derivation being not done perfectly well. So, 100% measurement error is not possible with any methods. It is very important for the operator to take proper care of the experiment while performing on industrial instruments so that the error in measurement can be reduced. Some of the errors are constant in nature due to the unknown reasons, some will be random in nature, and the other will be due to gross blunder on the part of the experimenter.

Sources of Error in Measurement scale

Measurement scale should be precise the source of error and unequivocal in an ideal researches. This objective, however, is often not met with in entirety. As such the researcher must be aware about the sources of error in measurement. The following content may expresses the possible sources of error on measurement.

Instrument error refers to the combined accuracy and precision of a measuring **instrument**, or the difference between the actual value and the value indicated by the **instrument (error)**. Measuring instruments are usually calibrated on some regular frequency against a standard.

Respondent error

- The respondent error consist the respondent providing information for research it may be inaccurate or wrong information.
- They occur because of memory biases or respondents giving inaccurate or false information when they believe that they are protecting their personal interests or integrity.
- They can also arise from the way the respondent interprets the questionnaire and the wording of the answer that the respondent gives.

- Careful questionnaire design and effective questionnaire testing can overcome these problems to some extent

Respondent: The respondent it may be reluctant to explain the thought about the negative(-) feelings, instead that just favorable that they have very little knowledge but may not admit his ignorance. All this lack of enthusiasm is likely to result in an interview of 'guesses.' Transient factors like fatigue, boredom, anxiety, etc. may limit the ability of the respondent to respond accurately and fully.

Situation: Situational factors may also come in the way of correct measurement. The condition which is places a strain on interview can have serious effects on the interviewer-responder rapport. For instance, if someone else is present, he can distort responses by joining in or merely by being present. If the respondent feels that anonymity is not assured, he may be reluctant to express certain feelings.

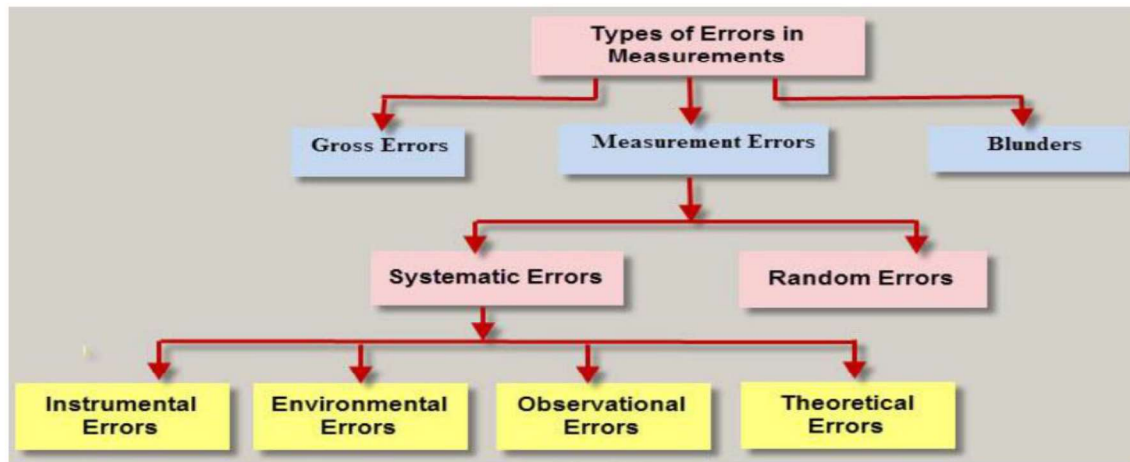
Measurer: The interviewer can distort responses by rewording or reordering questions. His behaviour, style and looks may encourage or discourage certain replies from respondents. Careless mechanical processing may distort the findings. Errors may also creep in because of incorrect coding, faulty tabulation and/or statistical calculations, particularly in the data-analysis stage.

Instrument: Error may arise because of the defective measuring instrument. The study beyond the understanding of the respondent would make the ambiguous meanings of poor printing, and insufficient space for response and choice omissions, there are a few things that make the measuring instrument defective and it result in measurement errors. Another type of instrument deficiency is the poor sampling of the universe of items of concern. The study it may extent possible and try to eliminate, neutralize or otherwise deal with all the possible sources of error so that the final results may not be contaminated.

Errors in Measurement System

An error may be defined as the difference between the measured value and the actual value. For example, if the two operators use the same device or instrument for finding the errors in measurement, it is not necessary that they may get the similar results. There may be a difference between both measurements. The difference that occurs between both the measurements is referred to as an ERROR. Sequentially, to understand the concept of errors in measurement, you should know the two terms that define the error. They are true value and measured value. The true value is impossible to find out the truth of quantity by experimental means. It may be defined as the average value of an infinite number of measured values.

Measured value can be defined as the estimated value of true value that can be found by taking several measured values during an experiment.



SCALE CONSTRUCTION TECHNIQUES

The technique of developing measurement tools consists four-stages of process. It consisting of the following:

Understanding of major concepts.

The first and foremost step is that of concept development which means that the analysis should clarified and an understanding of the major concepts needed to his study. This study concentrate to develop the concept and making the analysis should be most suitable apparent in theoretical studies than in the more pragmatic research, where the fundamental concepts are often already established.

Concepts of Dimensions

The II steps requires the researcher to specify the dimensions of the concepts that he developed in the first stage. This task should be accomplished based on deduction. It means that the adopting a new intuitive approach to taken by empirical correlation which is based on the individual dimensions with the total concept. It think more or less several dimensions like that reputation of products, treatment of customers, leading the corporate leadership, concern for individuals, sense of social responsibility and so forth when one is thinking about the image of a certain industry. The dimensions concepts specified, the analysis should in-depth develop indicators for measuring step by step development of research elements.

Perfect measure of a concept

Indicators are specific questions, scales, or other devices by which respondent's knowledge, opinion, expectation, etc., are measured. As there is seldom a perfect measure of a concept,

the researcher should consider several alternatives for the purpose. The use of more than one indicator gives stability to the scores and it also improves their validity.

Combining the various indicators

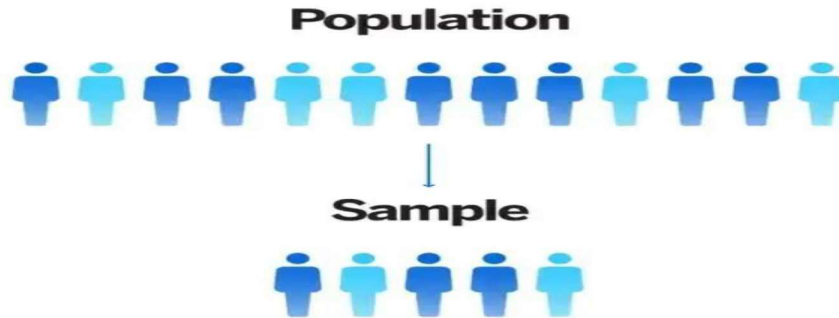
The fourth step is that of combining the different indicators into an index that is formation of an index. The study have different kind of dimensions concept, and it may be consists the various dimension of measurements, it need to combine them into a single index. The one or more simplest methods are there to getting an overall index is to proved the scale values to the customers and calculate the total score with the help of scale value to find out the corresponding scores. The overall index given a satisfactory measurement tool to indicate a single indicator i.e that the fact which indicate the individual indicator had the probability chance identification and what we really want to know. This way it may obtain an overall index for the different dimension concepts of the research study.

Conclusion

The measurement research in science or social science which is help to provide meaningful information for decision making process. The scale of measurement which is help to marketing research to measure brands, store types, sales territories, geographical locations, heavy versus light users, working versus non-working women and brand awareness versus non-awareness. Measurement is the assignment of numbers or other symbols to characteristics of objects according to set rules. Scaling involves the generation of a continuum upon which measure objects or located. Scaling techniques can be classified as comparative or non comparative. It involves a direct comparison of stimulus objects. The statistics measurement techniques which are highly need to research.

UNIT-III
SAMPLING DESIGN

MEANING OF SAMPLING DESIGN



Meaning:

Sample design refers to the plan, strategy, or blueprint for selecting a sample from a larger population. It outlines how many units will be selected, which units will be selected, and how they will be selected so that the sample accurately represents the population.

Definition:

Sample design is the structured plan that specifies the method and procedures used to choose a subset (sample) from a given population for the purpose of making statistical inferences. It includes decisions about the sampling method (e.g., random, stratified, cluster), sample size, sampling frame, and selection procedures.

Sample design is the plan for picking the people or items you will study so that your findings reflect the larger group accurately.

STEPS IN SAMPLING

Sampling design is a crucial aspect of conducting high-quality research. It involves selecting a representative subset of the population to participate in the research project. In this article, we will discuss the steps involved in sampling design to ensure that the sample is representative of the population.

1. Define the Population

The first step in sampling design is to define the population. The population should be well-defined and identifiable. The population should also be relevant to the research question and objectives.

2. Determine the Sample Size

The second step in sampling design is to determine the sample size. The sample size should be large enough to ensure that the research findings are generalizable to the population. The sample size should also be cost-effective and time-saving.



3. Select the Sampling Technique

The third step in sampling design is to select the appropriate sampling technique. The sampling technique should be relevant to the research question and objectives. The sampling technique should also be appropriate for the population and the sample size.

4. Create the Sampling Frame

The fourth step in sampling design is to create the sampling frame. The sampling frame is a list of individuals or units in the population. The sampling frame should be complete, up-to-date, and accurate.

5. Randomize the Sample

The fifth step in sampling design is to randomize the sample. Randomization ensures that every member of the population has an equal chance of being selected for the sample. Randomization also reduces the risk of sampling bias.

6. Collect Data from the Sample

The sixth step in sampling design is to collect data from the sample. The data should be collected using appropriate data collection techniques, such as surveys, interviews, or observations. The data should also be collected in a standardized and consistent manner.

7. Analyze the Data

The seventh step in sampling design is to analyze the data. The data should be analyzed using appropriate statistical techniques. The analysis should also be relevant to the research question and objectives.

Conclusion

In conclusion, sampling design is a crucial aspect of conducting high-quality research. The steps involved in sampling design include defining the population, determining the sample size, selecting the sampling technique, creating the sampling frame, randomizing the sample, collecting data from the sample, and analyzing the data. By following these steps, researchers can ensure that the sample is representative of the population and that the research findings are generalizable to the population. By conducting high-quality sampling in research, researchers can produce high-quality research findings.

CRITERIA FOR GOOD SAMPLE DESIGN

1. Representativeness:

- ❖ A researcher selects a relatively small number for a sample from an entire population.
- ❖ This sample needs to closely match all the characteristics of the entire population.
- ❖ It will help generalize the results from a small group to large population being studied.

2. Absence of sampling error:

- ❖ Sampling error is the error caused by taking a small sample instead of the whole population for study.
- ❖ Sampling error is reduced by selecting a large sample and by using efficient sample design and estimation strategies.

3. Economically viable:

- ❖ Studies have a limited budget called the research budget.
- ❖ The sample should be done in such a way that it is within the research budget and not too expensive to be replicated

4. Generalized and applicable:

- ❖ Samples that it covers the whole population of the study and is not limited to a part.

5. Goal oriented:

- ❖ A sample should be goal oriented.
- ❖ It is means and should be oriented to the research objectives and fitted to the research conditions.

6. Proportional:

- ❖ A sample should be proportional.
- ❖ It should be large enough to represent the population properly.
- ❖ The sample size should be sufficiently large to provide statistical stability or reliability.
- ❖ The sample size should give accuracy required for the purpose of particular study.

7. Randomly Selected:

- ❖ A sample should be selected at random.
- ❖ This means that any item in the group has a full and equal chance of being selected and included in the sample.
- ❖ This makes the selected sample truly representative in character.

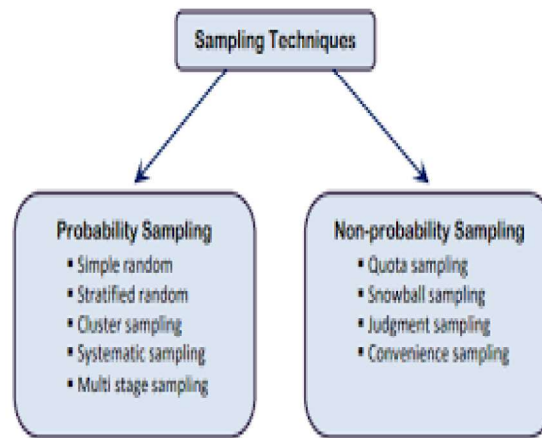
8. Actual information provider:

- ❖ A sample should be as to provide actual information required for the study or research.

9. Practical:

- ❖ A sample should be practical.
- ❖ The sample should be simple i.e. it should be capable of being understood and followed in the fieldwork.

TYPES OF SAMPLING DESIGNS



Sampling helps a lot in research. It is one of the most important factors which determines the accuracy of your research/survey result. If anything goes wrong with your sample then it will be directly reflected in the final result. There are lot of techniques which help us to gather sample depending upon the need and situation. This blog post tries to explain some of those techniques.

Population is the collection of the elements which has some or the other characteristic in common. Number of elements in the population is the size of the population.

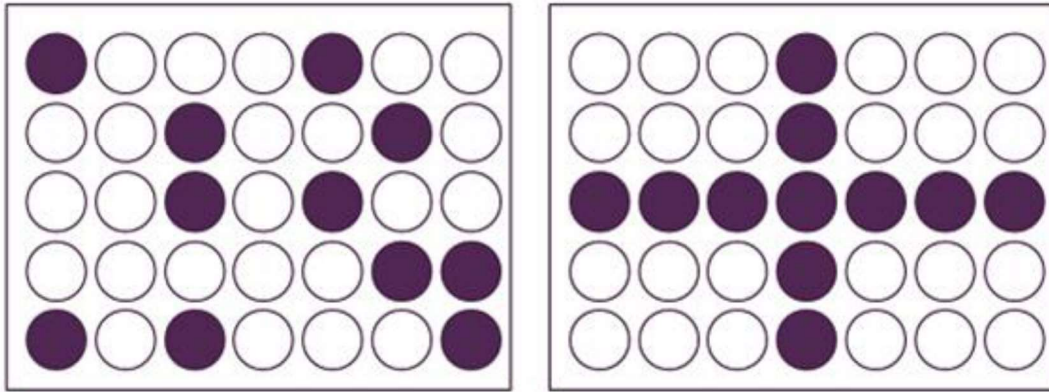
Sample is the subset of the population. The process of selecting a sample is known as sampling. Number of elements in the sample is the sample size.



Sampling There are lot of sampling techniques which are grouped into two categories as

1. PROBABILITY SAMPLING

2. NON- PROBABILITY SAMPLING



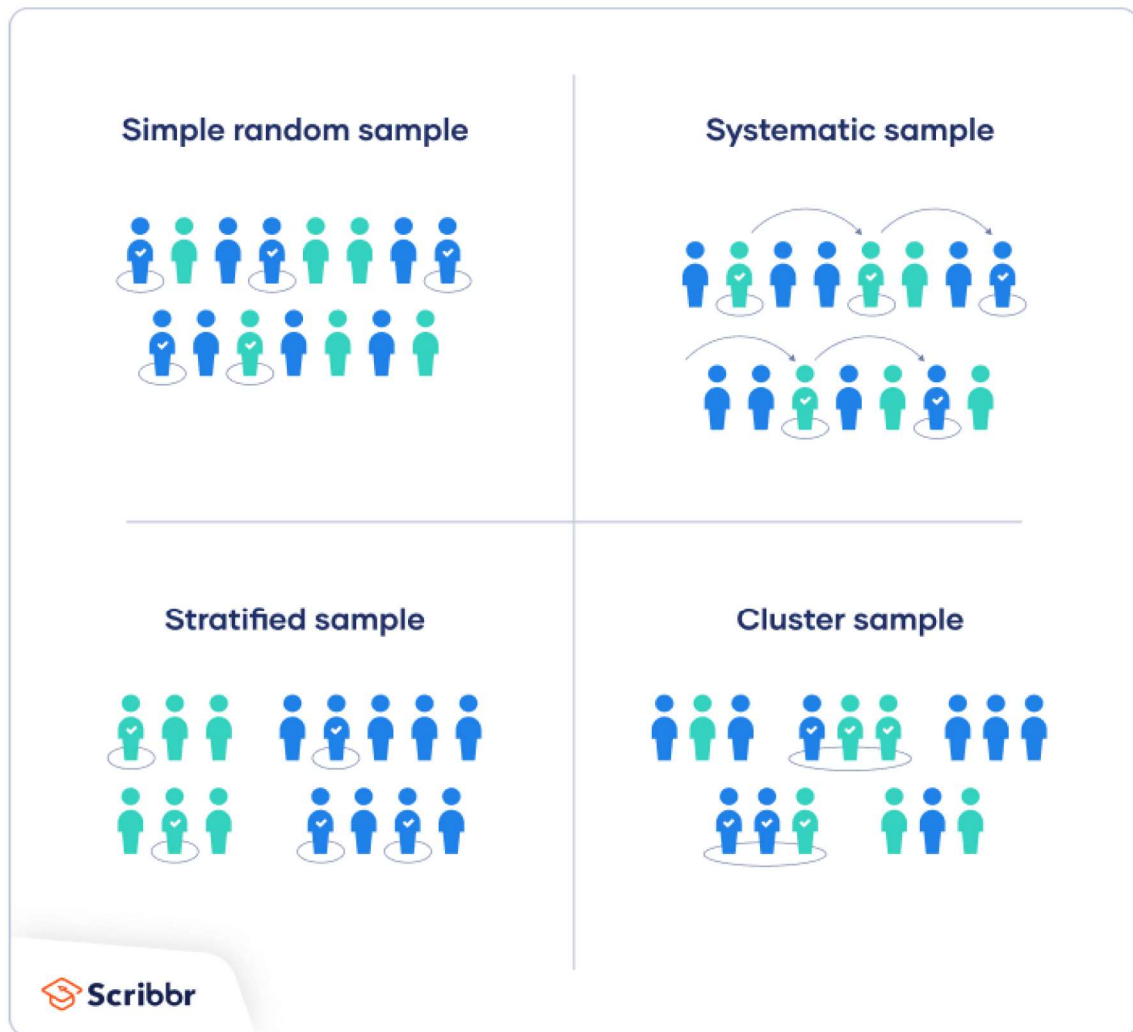
Probability Sampling Vs Non-Probability Sampling

The difference lies between the above two is whether the sample selection is based on randomization or not. With randomization, every element gets equal chance to be picked up and to be part of sample for study.

PROBABILITY SAMPLING

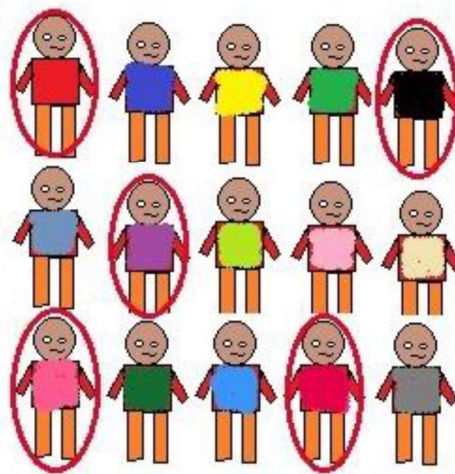
This Sampling technique uses randomization to make sure that every element of the population gets an equal chance to be part of the selected sample. It's alternatively known as random sampling.

- Simple Random Sampling
- Stratified sampling
- Systematic sampling
- Cluster Sampling
- Multi stage Sampling

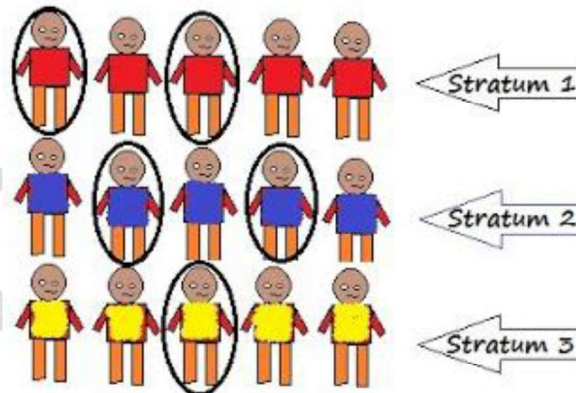


1. Simple Random Sampling: Every element has an equal chance of getting selected to be the part sample. It is used when we don't have any kind of prior information about the target population.

For example: Random selection of 20 students from class of 50 student. Each student has equal chance of getting selected. Here probability of selection is $1/50$



2. Stratified Sampling: This technique divides the elements of the population into small subgroups (strata) based on the similarity in such a way that the elements within the group are homogeneous and heterogeneous among the other subgroups formed. And then the elements are randomly selected from each of these strata. We need to have prior information about the population to create subgroups.

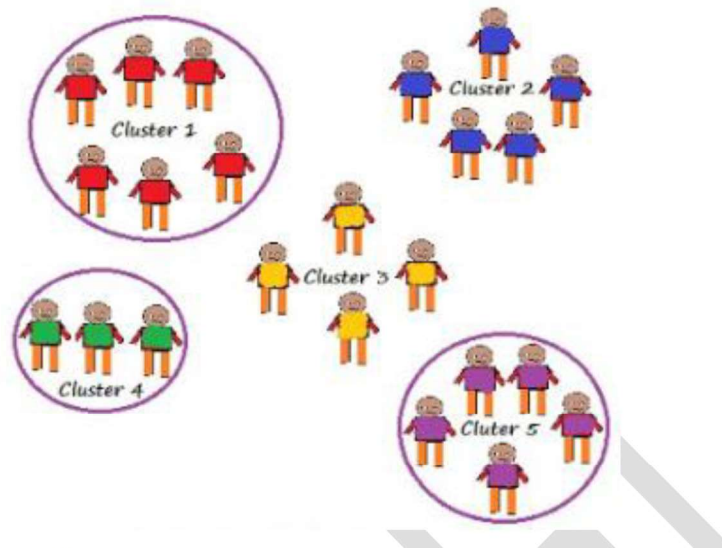


3. Cluster Sampling: Our entire population is divided into clusters or sections and then the clusters are randomly selected. All the elements of the cluster are used for sampling. Clusters are identified using details such as age, sex, location etc.

Cluster sampling can be done in following ways:

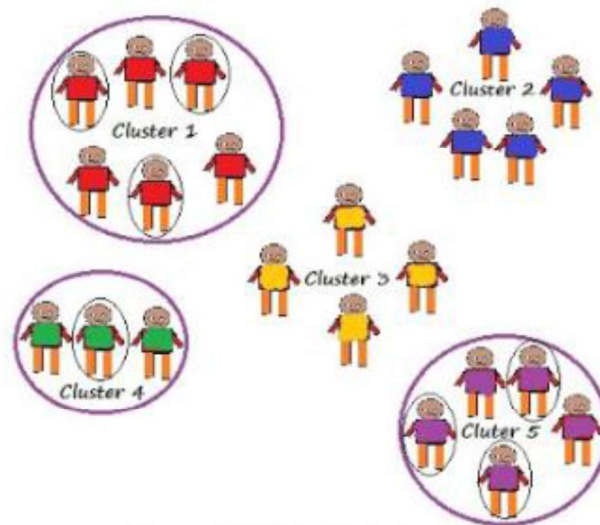
Single Stage Cluster Sampling

Entire cluster is selected randomly for sampling.



Two Stage Cluster Sampling

Here first we randomly select clusters and then from those selected clusters we randomly select elements for sampling



4. Systematic Clustering

Here the selection of elements is systematic and not random except the first element. Elements of a sample are chosen at regular intervals of population. All the elements are put together in a sequence first where each element has the equal chance of being selected.

For a sample of size n , we divide our population of size N into subgroups of k elements.

We select our first element randomly from the first subgroup of k elements.

To select other elements of sample, perform following:

We know number of elements in each group is k i.e N/n

So if our first element is n_1 then

Second element is n_1+k i.e n_2

Third element n_2+k i.e n_3 and so on..

Taking an example of $N=20$, $n=5$

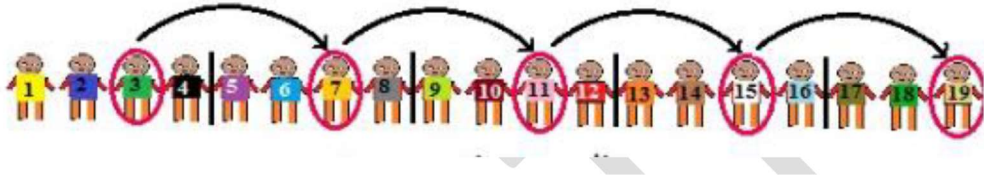
No of elements in each of the subgroups is N/n i.e $20/5 = 4 = k$

Now, randomly select first element from the first subgroup.

If we select $n_1 = 3$

$n_2 = n_1 + k = 3 + 4 = 7$

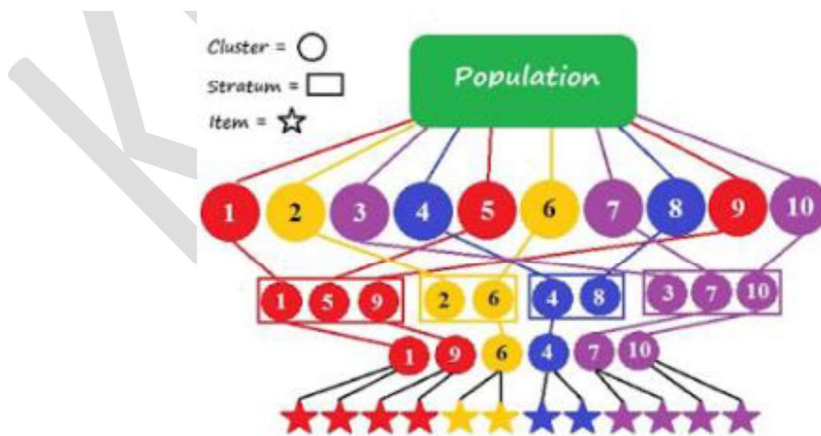
$n_3 = n_2 + k = 7 + 4 = 11$



5. Multi-Stage Sampling

It is the combination of one or more methods described above.

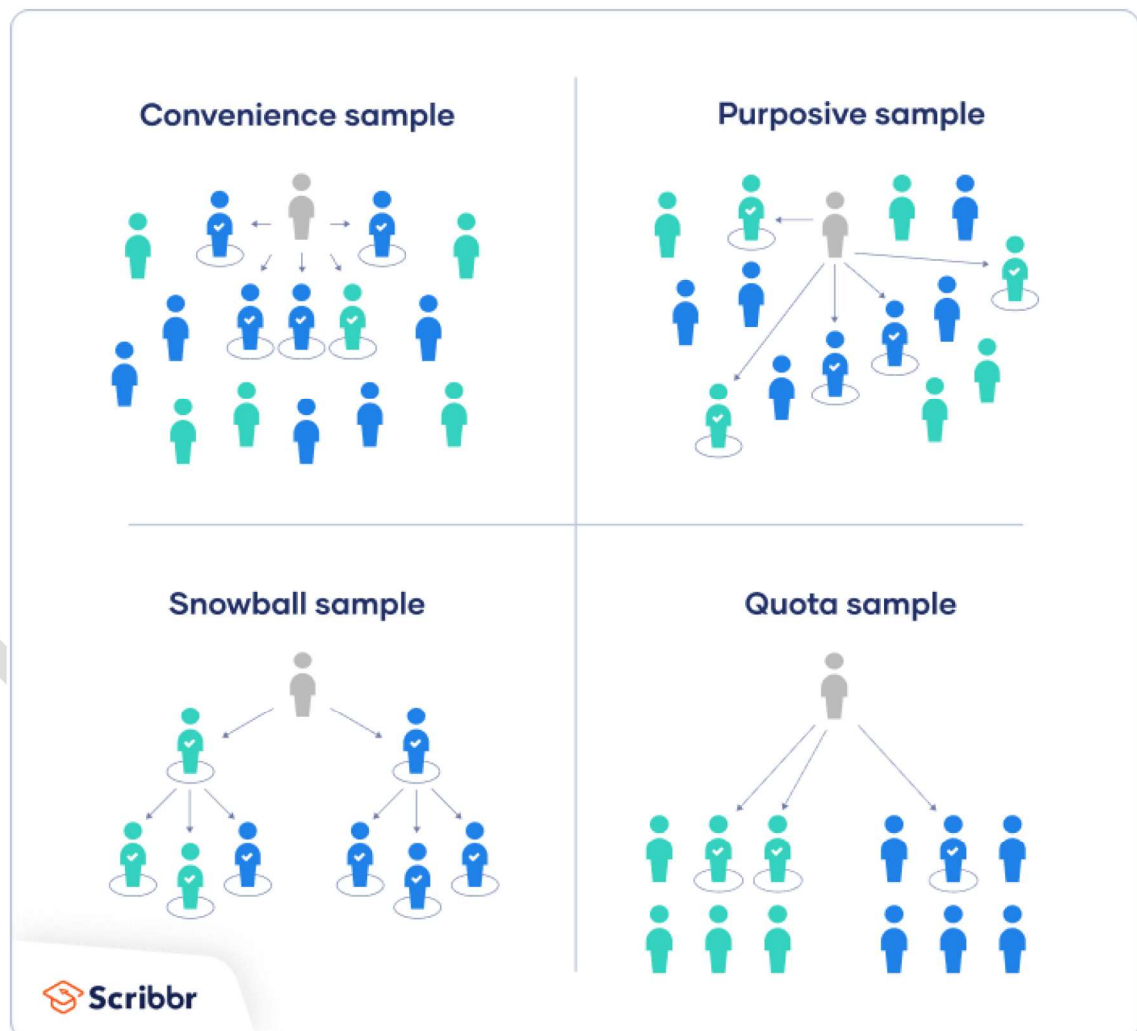
Population is divided into multiple clusters and then these clusters are further divided and grouped into various sub groups (strata) based on similarity. One or more clusters can be randomly selected from each stratum. This process continues until the cluster can't be divided anymore. For example country can be divided into states, cities, urban and rural and all the areas with similar characteristics can be merged together to form a strata.



NON-PROBABILITY SAMPLING

It does not rely on randomization. This technique is more reliant on the researcher's ability to select elements for a sample. Outcome of sampling might be biased and makes difficult for all the elements of population to be part of the sample equally. This type of sampling is also known as non-random sampling.

- Convenience Sampling
- Purposive Sampling
- Quota Sampling
- Referral /Snowball Sampling



1. Convenience Sampling: Here the samples are selected based on the availability. This method is used when the availability of sample is rare and also costly. So based on the convenience samples are selected.

For example: Researchers prefer this during the initial stages of survey research, as it's quick and easy to deliver results.

2. Purposive Sampling: This is based on the intention or the purpose of study. Only those elements will be selected from the population which suits the best for the purpose of our study.

For Example: If we want to understand the thought process of the people who are interested in pursuing master's degree then the selection criteria would be "Are you interested for Masters in..?"

All the people who respond with a "No" will be excluded from our sample.

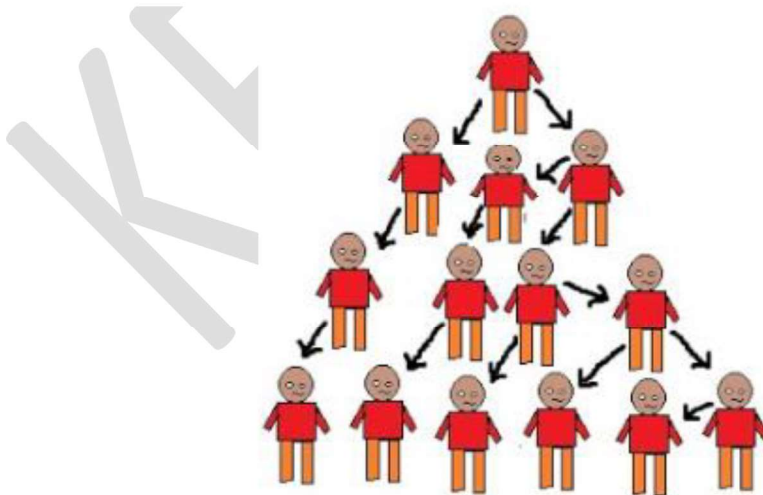
3. Quota Sampling: This type of sampling depends of some pre-set standard. It selects the representative sample from the population. Proportion of characteristics/ trait in sample should be same as population. Elements are selected until exact proportions of certain types of data is obtained or sufficient data in different categories is collected.

For example: If our population has 45% females and 55% males then our sample should reflect the same percentage of males and females.

4. Referral /Snowball Sampling: This technique is used in the situations where the population is completely unknown and rare.

Therefore we will take the help from the first element which we select for the population and ask him to recommend other elements who will fit the description of the sample needed.

So this referral technique goes on, increasing the size of population like a snowball.



Referral /Snowball Sampling

For example: It's used in situations of highly sensitive topics like HIV Aids where people will not openly discuss and participate in surveys to share information about HIV Aids.

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Not all the victims will respond to the questions asked so researchers can contact people they know or volunteers to get in touch with the victims and collect information

Helps in situations where we do not have the access to sufficient people with the characteristics we are seeking. It starts with finding people to study.

DATA COLLECTION

Data Collection refers to the systematic process of gathering, measuring, and analyzing information from various sources to get a complete and accurate picture of an area of interest. Data collection is a critical step in any research or data-driven decision-making process, ensuring the accuracy and reliability of the results obtained. By employing various methods of data collection researchers and organizations can gather the necessary data to support their objectives effectively.

Data Collection is the process of collecting information from relevant sources to find a solution to the given statistical inquiry. Collection of Data is the first and foremost step in a statistical investigation. It's an essential step because it helps us make informed decisions, spot trends, and measure progress.

Different methods of collecting data include

- ❖ Interviews
- ❖ Questionnaires
- ❖ Observations
- ❖ Experiments
- ❖ Published Sources and Unpublished Sources

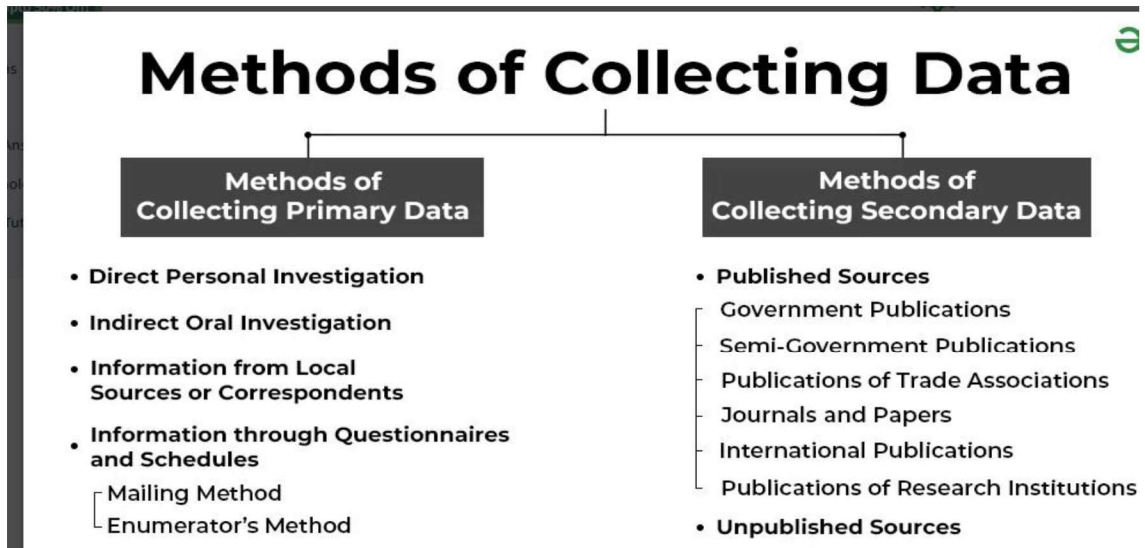
Here, statistical inquiry means an investigation by any agency on a topic in which the investigator collects the relevant quantitative information. In simple terms, a statistical inquiry is a search for truth by using statistical methods of collection, compiling, analysis, interpretation, etc. The basic problem for any statistical inquiry is the collection of facts and figures related to this specific phenomenon that is being studied. Therefore, the basic purpose of data collection is collecting evidence to reach a sound and clear solution to a problem.

Terms Related to Data Collection

- ❖ **Data:** Data is a tool that helps an investigator in understanding the problem by providing him with the information required. Data can be classified into two types; viz., **Primary Data** and **Secondary Data**.

- ❖ **Investigator:** An investigator is a person who conducts the statistical enquiry.
- ❖ **Enumerators:** In order to collect information for statistical enquiry, an investigator needs the help of some people. These people are known as enumerators.
- ❖ **Respondents:** A respondent is a person from whom the statistical information required for the enquiry is collected.
- ❖ **Survey:** It is a method of collecting information from individuals. The basic purpose of a survey is to collect data to describe different characteristics such as usefulness, quality, price, kindness, etc. It involves asking questions about a product or service from a large number of people.

DATA COLLECTION METHODS



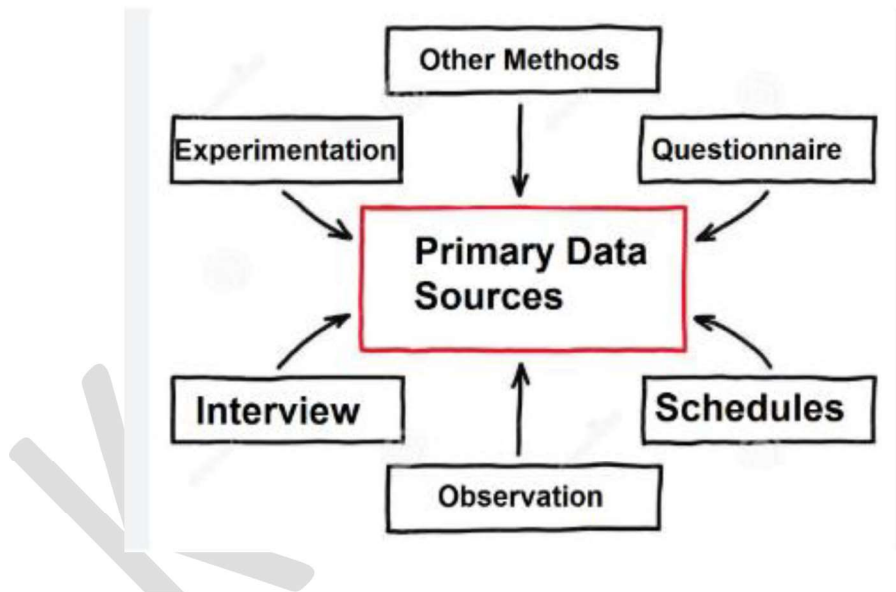
There are two different methods of collecting data: Primary Data Collection and Secondary Data Collection.

TYPES OF DATA

1. Primary data
2. Secondary data

unstop	
Difference Between Primary & Secondary Data	
Primary Data	Secondary Data
• First-hand data • More reliable • Control over collection method • Time & money intensive	• Second-hand Data • Comparatively less reliable • No control over collection method • Time & cost effective

PRIMARY DATA



Primary data refers to information collected directly from first-hand **sources** specifically for a particular research purpose. This type of data is gathered through various methods, including surveys, interviews, experiments, observations, and focus groups. One of the main advantages of primary data is that it provides current, relevant, and specific information tailored to the researcher's needs, offering a high level of accuracy and control over data quality.

Methods of Collecting Primary Data

There are a number of methods of collecting primary data, Some of the common methods are as follows:

1. Interviews: Collect data through direct, one-on-one conversations with individuals. The investigator asks questions either directly from the source or from its indirect links.

Direct Personal Investigation: The method of direct personal investigation involves collecting data personally from the source of origin. In simple words, the investigator makes direct contact with the person from whom he/she wants to obtain information.

For example, direct contact with the household women to obtain information about their daily routine and schedule.

Indirect Oral Investigation: In the indirect oral investigation method of collecting primary data, the investigator does not make direct contact with the person from whom he/she needs information, instead they collect the data orally from some other person who has the necessary required information.

For example, collecting data of employees from their superiors or managers.

- **Advantage:** Provides real-time, natural data; no reliance on self-reported information.
- **Disadvantage:** Observer bias; limited to what can be seen; may influence subjects' behavior.
- **Suitable Use Case:** Behavioral studies, user experience research.

2. Questionnaires: Collect data by asking people a set of questions, either online, on paper, or face-to-face. In this method the investigator prepares a questionnaire to collect Information through Questionnaires and Schedules , while keeping in mind the motive of the study, . The investigator can collect data through the questionnaire in two ways:

Mailing Method: This method involves mailing the questionnaires to the informants for the collection of data. The investigator attaches a letter with the questionnaire in the mail to define the purpose of the study or research.

Enumerator's Method: This method involves the preparation of a questionnaire according to the purpose of the study or research. However, in this case, the enumerator reaches out to the informants himself with the prepared questionnaire.

- **Advantage:** Can reach a large audience quickly and cost-effectively.

- **Disadvantage:** Responses may be biased or inaccurate; low response rates.
- **Suitable Use Case:** Customer satisfaction surveys, market research.

3. Observations: The observation method involves collecting data by watching and recording behaviors, events, or conditions as they naturally occur. The observer systematically watches and notes specific aspects of a subject's behavior or the environment, either covertly or overtly.

- **Advantage:** Provides real-time, authentic data without reliance on self-reported information.
- **Disadvantage:** Observer bias can influence the results, and the presence of an observer might alter subjects' behavior.
- **Suitable Use Case:** Studying user interactions with a product in a natural setting, monitoring wildlife behavior, or assessing classroom dynamics.

4. Experiments: The experiment method involves manipulating one or more variables to determine their effect on another variable, within a controlled environment. Researchers create two groups (control and experimental), apply the treatment or variable to the experimental group, and compare the outcomes between the groups.

- **Advantage:** Allows for the establishment of cause-and-effect relationships with high precision.
- **Disadvantage:** Experiments can be artificial, limiting the ability to generalize findings to real-world settings, and they can be resource-intensive.
- **Suitable Use Case:** Testing the efficacy of a new drug, assessing the impact of a new teaching method, or evaluating the effect of a marketing campaign.

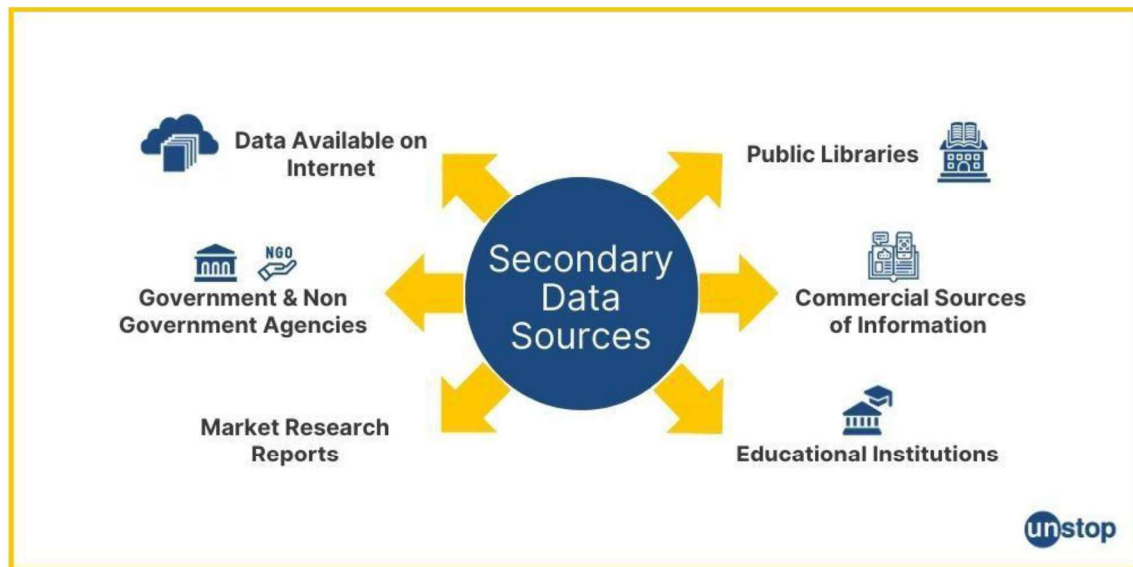
5. Focus Group: The focus group method involves gathering a small group of people to discuss a specific topic or product, facilitated by a moderator. A group of 6-12 participants engages in a guided discussion led by a moderator who asks open-ended questions to elicit opinions, attitudes, and perceptions.

- **Advantage:** Provides in-depth insights and diverse perspectives through interactive discussions, revealing the reasoning behind participants' thoughts and feelings.
- **Disadvantage:** Results can be influenced by dominant participants or groupthink, and the findings are not easily generalizable due to the small, non-representative sample size.

- **Suitable Use Case:** Exploring customer attitudes towards a new product, gathering feedback on a marketing campaign, or understanding public opinion on social issues.

6. Information from Local Sources or Correspondents: In this method, for the collection of data, the investigator appoints correspondents or local persons at various places, which are then furnished by them to the investigator. With the help of correspondents and local persons, the investigators can cover a wide area.

SECONDARY DATA



Secondary data refers to information that has already been collected, processed, and published by others. This type of data can be sourced from existing research papers, government reports, books, statistical databases, and company records. The advantage of secondary data is that it is readily available and often free or less expensive to obtain compared to primary data. It saves time and resources since the data collection phase has already been completed.

Methods of Collecting Secondary Data

Secondary data can be collected through different published and unpublished sources. Some of them are as follows:

1. Published Sources

- **Government Publications:** Government publishes different documents which consists of different varieties of information or data published by the Ministries, Central and State Governments in India as their routine activity. As the government publishes these

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Statistics, they are fairly reliable to the investigator. **Examples** of Government publications on Statistics are the Annual Survey of Industries, Statistical Abstract of India, etc.

- **Semi-Government Publications:** Different Semi-Government bodies also publish data related to health, education, deaths and births. These kinds of data are also reliable and used by different informants. Some **examples** of semi-government bodies are Metropolitan Councils, Municipalities, etc.
- **Publications of Trade Associations:** Various big trade associations collect and publish data from their research and statistical divisions of different trading activities and their aspects.

For example, data published by Sugar Mills Association regarding different sugar mills in India.

- **Journals and Papers:** Different newspapers and magazines provide a variety of statistical data in their writings, which are used by different investigators for their studies.
- **International Publications:** Different international organizations like IMF, UNO, ILO, World Bank, etc., publish a variety of statistical information which are used as secondary data.
- **Publications of Research Institutions:** Research institutions and universities also publish their research activities and their findings, which are used by different investigators as secondary data. For example National Council of Applied Economics, the Indian Statistical Institute, etc.

2. Unpublished Sources

Unpublished sources are another source of collecting secondary data. The data in unpublished sources is collected by different government organizations and other organizations. These organizations usually collect data for their self-use and are not published anywhere. For example, research work done by professors, professionals, teachers and records maintained by business and private enterprises.

Example:

The table below shows the production of rice in India.

DATA COLLECTION TOOLS

Now that we've explained the various techniques let's narrow our focus even further by looking at some specific tools. For example, we mentioned interviews as a technique, but we can further break that down into different interview types (or "tools").

- **Word Association**

The researcher gives the respondent a set of words and asks them what comes to mind when they hear each word.

- **Sentence Completion**

Researchers use sentence completion to understand the respondent's ideas. This tool involves giving an incomplete sentence and seeing how the interviewee finishes it.

- **Role-Playing**

Respondents are presented with an imaginary situation and asked how they would act or react if it were real.

- **In-Person Surveys**

The researcher asks questions in person.

- **Online/Web Surveys**

These surveys are easy to accomplish, but some users may be unwilling to answer truthfully, if at all.

- **Mobile Surveys**

These surveys take advantage of the increasing proliferation of mobile technology. Mobile collection surveys rely on mobile devices like tablets or smartphones to conduct surveys via SMS or mobile apps.

- **Phone Surveys**

No researcher can call thousands of people at once, so they need a third party to handle the chore. However, many people have call screening and won't answer.

- **Observation**

Sometimes, the simplest method is the best. Researchers who make direct observations collect data quickly and easily, with little intrusion or third-party bias. Naturally, this method is only effective in small-scale situations.

Production of Rice in India (in Million Tonnes)

Year (X)	Production of Rice (Y)
1950-1951	20.58
1966-1967	30.34
1975-1976	48.74
1998-1999	86.03
2002-2003	77.7
2020-2021	120
2022-2023 (Estimated)	135.52

The above table contains the production of rice in India in different years. It can be seen that these values vary from one year to another. Therefore, they are known as **variable**. A variable is a quantity or attribute, the value of which varies from one investigation to another. In general, the variables are represented by letters such as X, Y, or Z. In the above example, years are represented by variable X, and the production of rice is represented by variable Y. The values of variable X and variable Y are **data** from which an investigator and enumerator collect information regarding the trends of rice production in India.

Conclusion

Data collection is the backbone of any research or statistical investigation, providing the necessary information **to** make informed decisions, identify trends, and measure progress. By understanding the various methods of data collection—such as direct personal investigation, indirect oral investigation, questionnaires, observations, experiments, and focus groups—researchers can choose the most suitable approach to gather primary data that is current, relevant, and accurate. Similarly, using secondary data from published and unpublished sources like government reports, trade associations, and research institutions can save time and resources while offering valuable insights. Mastering these data collection techniques ensures the reliability and validity of the research, ultimately leading to sound and actionable conclusions.

CONSTRUCTING QUESTIONNAIRE

Meaning of questionnaire

A **questionnaire** is a list of questions or items used to gather data from respondents about their attitudes, experiences, or opinions. Questionnaires can be used to collect quantitative and/or qualitative information.

Questionnaires are commonly used in market research as well as in the social and health sciences. For example, a company may ask for feedback about a recent customer service experience, or psychology researchers may investigate health risk perceptions using questionnaires.

Step-by-step guide to constructing questionnaire

Step 1: Define your goals and objectives

The first step of designing a questionnaire is determining your aims.

- ❖ What topics or experiences are you studying?
- ❖ What specifically do you want to find out?
- ❖ Is a self-report questionnaire an appropriate tool for investigating this topic?

Once you've specified your research aims, you can operationalize your variables of interest into questionnaire items. Operationalizing concepts means turning them from abstract ideas into concrete measurements. Every question needs to address a defined need and have a clear purpose.

Step 2: Use questions that are suitable for your sample

Create appropriate questions by taking the perspective of your respondents. Consider their language proficiency and available time and energy when designing your questionnaire.

- ❖ Are the respondents familiar with the language and terms used in your questions?
- ❖ Would any of the questions insult, confuse, or embarrass them?
- ❖ Do the response items for any closed-ended questions capture all possible answers?
- ❖ Are the response items mutually exclusive?
- ❖ Do the respondents have time to respond to open-ended questions?

Consider all possible options for responses to closed-ended questions. From a respondent's perspective, a lack of response options reflecting their point of view or true answer may make them feel alienated or excluded. In turn, they'll become disengaged or inattentive to the rest of the questionnaire.

Step 3: Decide on your questionnaire length and question order

Once you have your questions, make sure that the length and order of your questions are appropriate for your sample.

If respondents are not being incentivized or compensated, keep your questionnaire short and easy to answer. Otherwise, your sample may be biased with only highly motivated respondents completing the questionnaire.

Decide on your question order based on your aims and resources. Use a logical flow if your respondents have limited time or if you cannot randomize questions. Randomizing questions helps you avoid bias, but it can take more complex statistical analysis to interpret your data.

Step 4: Pretest your questionnaire

When you have a complete list of questions, you'll need to pretest it to make sure what you're asking is always clear and unambiguous. Pretesting helps you catch any errors or points of confusion before performing your study.

Ask friends, classmates, or members of your target audience to complete your questionnaire using the same method you'll use for your research. Find out if any questions were particularly difficult to answer or if the directions were unclear or inconsistent, and make changes as necessary.

If you have the resources, running a pilot study will help you test the validity and reliability of your questionnaire. A pilot study is a practice run of the full study, and it includes sampling, data collection, and analysis. You can find out whether your procedures are unfeasible or susceptible to bias and make changes in time, but you can't test a hypothesis with this type of study because it's usually statistically underpowered.

Questionnaire methods:

Questionnaires can be **self-administered** or **researcher-administered**. Self-administered questionnaires are more common because they are easy to implement and inexpensive, but researcher-administered questionnaires allow deeper insights.

1. Self-administered questionnaires

Self-administered questionnaires can be delivered online or in paper-and-pen formats, in person or through mail. All questions are standardized so that all respondents receive the same questions with identical wording.

Self-administered questionnaires can be:

- ❖ cost-effective
- ❖ easy to administer for small and large groups

- ❖ anonymous and suitable for sensitive topics
- ❖ self-paced

But they may also be:

- ❖ unsuitable for people with limited literacy or verbal skills
- ❖ susceptible to a nonresponse bias (most people invited may not complete the questionnaire)
- ❖ biased towards people who volunteer because impersonal survey requests often go ignored.

2. Researcher-administered questionnaires

Researcher-administered questionnaires are interviews that take place by phone, in-person, or online between researchers and respondents.

Researcher-administered questionnaires can:

- ❖ help you ensure the respondents are representative of your target audience
- ❖ allow clarifications of ambiguous or unclear questions and answers
- ❖ have high response rates because it's harder to refuse an interview when personal attention is given to respondents

But researcher-administered questionnaires can be limiting in terms of resources. They are:

- ❖ costly and time-consuming to perform
- ❖ more difficult to analyze if you have qualitative responses
- ❖ likely to contain experimenter bias or demand characteristics
- ❖ likely to encourage social desirability bias in responses because of a lack of anonymity

3. Open-ended vs. closed-ended questions

Your questionnaire can include open-ended or closed-ended questions or a combination of both.

Using closed-ended questions limits your responses, while open-ended questions enable a broad range of answers. You'll need to balance these considerations with your available time and resources.

4. Closed-ended questions

Closed-ended, or restricted-choice, questions offer respondents a fixed set of choices to select from. Closed-ended questions are best for collecting data on categorical or quantitative variables.

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Categorical variables can be nominal or ordinal. Quantitative variables can be interval or ratio. Understanding the type of variable and level of measurement means you can perform appropriate statistical analyses for generalizable results.

Examples of closed-ended questions for different variables

Nominal variables include categories that can't be ranked, such as race or ethnicity. This includes binary or dichotomous categories.

It's best to include categories that cover all possible answers and are mutually exclusive. There should be no overlap between response items.

In binary or dichotomous questions, you'll give respondents only two options to choose from.

Example: Nominal variables

What is your race?

- ☐ White
- ☐ Black or African American
- ☐ American Indian or Alaska Native
- ☐ Asian
- ☐ Native Hawaiian or Other Pacific Islander

Are you satisfied with the current work-from-home policies?

- ☐ Yes
- ☐ No

Ordinal variables include categories that can be ranked. Consider how wide or narrow a range you'll include in your response items, and their relevance to your respondents.

Example: Ordinal variables

What is your age?

- ☐ 15 or younger
- ☐ 16–35
- ☐ 36–60
- ☐ 61–75
- ☐ 76 or older

Likert scale questions collect ordinal data using rating scales with 5 or 7 points.

Example: Likert-type questions

How satisfied or dissatisfied are you with your online shopping experience today?

- ☐ Very dissatisfied
- ☐ Somewhat dissatisfied

- Neither satisfied nor dissatisfied
- Somewhat satisfied
- Very satisfied

When you have four or more Likert-type questions, you can treat the composite data as quantitative data on an interval scale. Intelligence tests, psychological scales, and personality inventories use multiple Likert-type questions to collect interval data.

With interval or ratio scales, you can apply strong statistical hypothesis tests to address your research aims.

Pros and cons of closed-ended questions

Well-designed closed-ended questions are easy to understand and can be answered quickly. However, you might still miss important answers that are relevant to respondents. An incomplete set of response items may force some respondents to pick the closest alternative to their true answer. These types of questions may also miss out on valuable detail.

To solve these problems, you can make questions partially closed-ended, and include an open-ended option where respondents can fill in their own answer.

Open-ended questions

Open-ended, or long-form, questions allow respondents to give answers in their own words. Because there are no restrictions on their choices, respondents can answer in ways that researchers may not have otherwise considered. For example, respondents may want to answer “multiracial” for the question on race rather than selecting from a restricted list.

Example: Open-ended questions

1. How do you feel about open science?
2. How would you describe your personality?
3. In your opinion, what is the biggest obstacle for productivity in remote work?

Open-ended questions have a few downsides.

They require more time and effort from respondents, which may deter them from completing the questionnaire.

For researchers, understanding and summarizing responses to these questions can take a lot of time and resources. You'll need to develop a systematic coding scheme to categorize answers, and you may also need to involve other researchers in data analysis for high reliability.

PILOT STUDY

Definition

pilot study: A small scale preliminary observational study or experiment used to evaluate feasibility, cost, and potential complications of a large scale research project.

Definition

A pilot study (also known as a "feasibility study", "feasibility test", "pilot experiment" or "pilot trial",) is a preliminary investigation designed to test the feasibility and design of a larger, more comprehensive study. It's essentially a 'trial run' conducted on a smaller scale, aiming to identify any potential issues and make necessary adjustments before the main research is undertaken. This type of study can be used in both qualitative research and quantitative research methodologies, serving to refine research questions, methods, and even the practicalities of data collection.

Pilot studies are not intended to provide definitive data or outcomes but rather to illuminate the path for smoother execution of the subsequent full-scale investigation.

Where Do We Use Pilot Studies In Research?

Pilot studies are used across many different research areas, particularly where new theories or methodologies are being tested. In healthcare research, they are used to evaluate the feasibility of large-scale clinical trials, assessing new interventions, or exploring patient responses.

Pilot studies are widely used in various disciplines such as psychology, public health, as well as in business, social sciences, and the humanities, to test and improve experimental designs, sampling techniques, data collection methods, and interpretation of results. Their versatility extends to the testing of new technology or software designed to enhance patient care or data analysis, where initial feedback can guide further development or refinement.

Advantages of A Pilot Study

Engaging in a pilot study presents benefits that streamline the research process and enhance its overall effectiveness such as:

- ❖ This study type provides the opportunity to identify and address the feasibility and potential pitfalls in the study design, methodology, and data collection techniques. This significantly reduces the risk of errors in the larger study, saving both time and resources.
- ❖ Pilot studies give researchers the chance to refine their research questions and objectives, ensuring they are sharply focused and aligned with the study's aims.
- ❖ This type of study provides insights into the practical aspects of conducting research, such as participant recruitment and retention strategies, which are crucial for the successful execution of the main study.
- ❖ Conducting a pilot study helps to ensure that the research is ethically sound. By testing the study procedures on a smaller scale, researchers can identify and mitigate any potential ethical issues.

Disadvantages of A Pilot Study

Whilst pilot studies are invaluable in many aspects of research, they are not without their drawbacks. Some of these include:

- ❖ Pilot studies have the potential for resource allocation issues; conducting a pilot study requires funding, time, and human resources, which could be limited or at the expense of the main study.
- ❖ Sometimes, the results from a pilot study may not be generalizable due to the smaller scale and sample size, leading to questions over the applicability of the findings to the broader research.
- ❖ There is also the risk that the pilot study's outcomes could inadvertently steer the direction of the main study, potentially biasing the research process and its conclusions.
- ❖ Despite their invaluable insights, pilot studies do not necessarily guarantee the success of the main research, which combined with their time-consuming and costly nature, may pose additional challenges for researchers.

What Types Of Studies Are Similar?

Several types of studies share similarities with pilot studies, albeit with distinct purposes and scopes:

- ❖ **Feasibility studies** are closely aligned with pilot studies in that they assess the practical aspects of conducting larger-scale research, focusing on whether it can be done, whether it should proceed, and if so, how.
- ❖ **Exploratory studies**, which, like a pilot study, aims to investigate a phenomenon or issue without aiming for conclusive results, often used to gather preliminary data or inform a more comprehensive study design.
- ❖ **Proof of concept studies** also bear a resemblance, designed to demonstrate the feasibility of a particular concept or theory, although often with a more focused aim related to product or process development rather than broader research questions

Pilot Study Methodology:

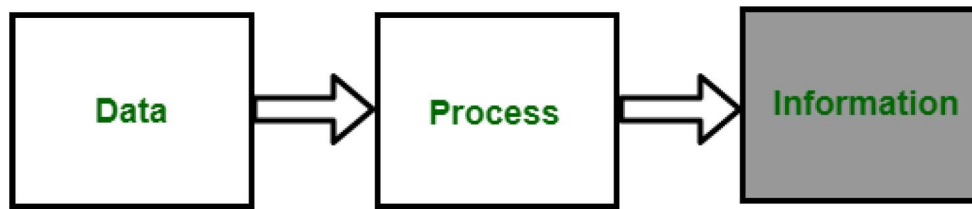
The methodology for a pilot study is a structured approach that mirrors the steps of a full-scale study but on a smaller, exploratory scale. Here's a detailed breakdown of how to conduct a pilot study:

- ❖ **Define Objectives:** Identify the specific goals of the pilot study, such as testing research tools or procedures.
- ❖ **Design the Study:** Choose a small, representative sample, and develop research instruments (e.g., surveys, interview guides).
- ❖ **Obtain Ethical Approval:** Secure approval from relevant ethics boards and obtain informed consent from participants.
- ❖ **Recruit Participants:** Use a strategy that mirrors the main study to recruit a small group of participants.
- ❖ **Collect Data:** Conduct the study using the planned methods (e.g., surveys, experiments, interviews).
- ❖ **Analyse Data:** Perform an initial analysis to identify any issues or patterns.
- ❖ **Review and Adjust:** Identify problems and refine the research design, tools, and procedures as needed.
- ❖ **Prepare for the Main Study:** Finalise the study design and ensure all necessary resources are in place.

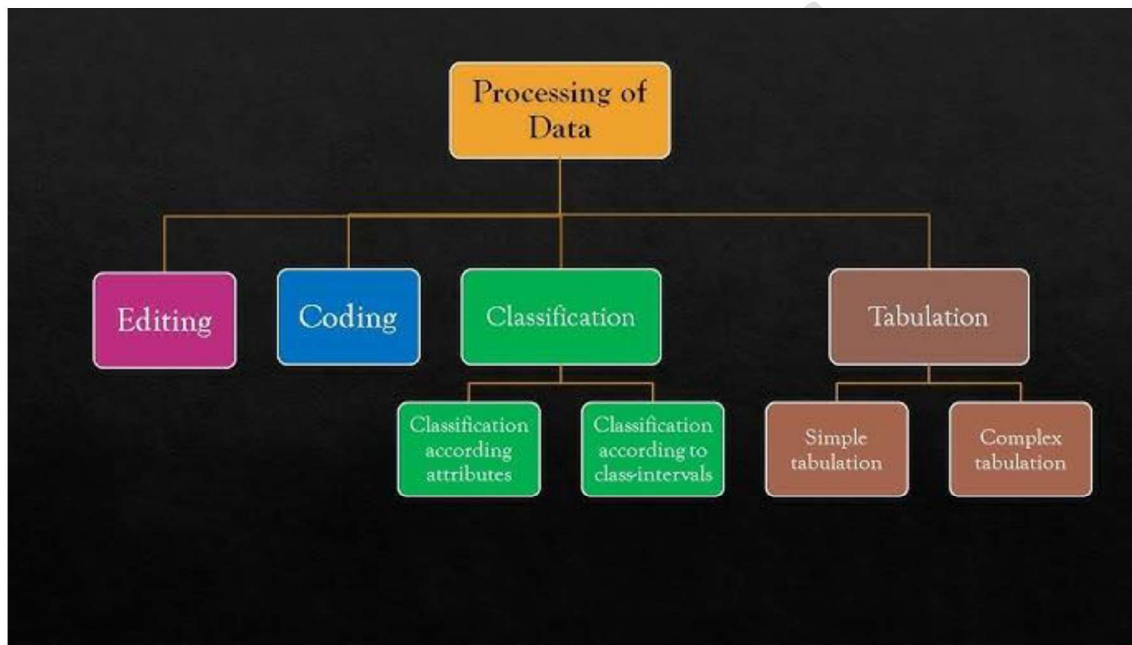
DATA PROCESSING

Data processing means to process of data i.e. to convert its format. As we all know data is the very useful and when it is well presented, and it becomes informative and useful. Data processing process system is also referred as information system. It is also right to say that

data processing becomes the process of converting information into data and also vice-versa.



EDITING



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 thorough check up is made. **Mildred B. Parten** in his book points out that the editor is responsible for seeing that the data are;

- Accurate as possible,
- Consistent with other facts secured,
- Uniformly entered,
- As complete as possible,
- Acceptable for tabulation and arranged to facilitate coding tabulation.

There are different **types of editing**. They are:

1. **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 and are there any wanton manipulation of data.

2. **Editing for tabulation** does certain accepted modification to data or even rejecting certain pieces of data in order to facilitate tabulation. or instance, extremely high or low value data item may be ignored or bracketed with suitable class interval.
3. **Field Editing** is done by the enumerator. The schedule filled up by the enumerator or the respondent might have some abbreviated writings, illegible writings and the like. These are rectified by the enumerator. This should be done soon after the enumeration or interview before the loss of memory. The field editing should not extend to giving some guess data to fill up omissions.
4. **Central Editing** is done by the researcher after getting all schedules or questionnaires or forms from the enumerators or 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.

Important points while performing editing:

- They should be familiar with instructions given to the interviewers and coders as well as with the editing instructions supplied to them for the purpose,
- While crossing out an original entry for one reason or another, they should just draw a single line on it so that the same may remain legible,
- They must make entries (if any) on the form in some distinctive color and that too in a standardized form,
- They should initial all answers which they change or supply,
- Editor's initials and the data of editing should be placed on each completed form or schedule.

Significance of editing

- ❖ It is pre-requisite for accuracy.
- ❖ It is useful in the elimination of the incorrect reply.
- ❖ It ensures the consistency of the data collected and avoids contradictions.
- ❖ It is useful to convert the answers into uniform units of measurement before coding

CODING

Coding is the process of organizing the data or response into classes or categories and assigning numerical or other symbols to responses according to the class or category in which they fall. Hence coding is considered as the classification process. Coding is necessary for efficient analysis. Coding is used to compartmentalize several replies effective into a small number of classes which contain the critical information required for analysis. In the process of coding, the study of answer is the first step and the last step is transfer of information from the schedule to the separate sheet called transcription sheet. Transcription sheet is a large summary sheet which contains the answers or codes of all the respondents. Transcription may not be necessary when only simple tables are required and the number of respondents is few. Coding is done with the help of set rules. The classes or categories should be reasonable and should be appropriate to the research problem, under study. The coding must be exhaustive; it means there should be class for each item of the data. For each answer it should be assigned with separate number. The coding should be based on the fact of mutual exclusivity it means specific answer can be place only in one category. The coding must observe the rule of single dimension; it means every class in the category set is defined in terms of only one concept. Coding provides base for analysis. It can be simplified if use of pre-coded questionnaire is made. The decision of coding should be taken well in advance at the stage of designing of questionnaire. By the investigators time is saved. Standard method should be used in case of hand coding. The process of coding is hard task but can be simplified if coding preparation is made prior to designing of schedule or questionnaire. Whatever method is adopted for coding the main important fact is that coding errors should be reduced.

Significance of coding

- ❖ It is useful in the classification of the responses into meaningful categories.
- ❖ It simplifies the difficult task of processing the qualitative information.
- ❖ One code is specific to only one kind of information so that a given response falls in only one category.

Rules for coding

- ❖ Give code numbers to each respondent and to each response.
- ❖ Give code numbers to qualitative response also.
- ❖ Prepare the coding frame.

While the editing is being performed, special coding actions are carried out. Geographic coding is in essence for the completion of automatic coding.

Example

S.N.	Category	Code
1	Male	1
2	Female	2

S.N.	Category	Code
1	Male	M
2	Female	F

S.N.	Category	Code
1	Male	
2	Female	

TABULATION OF DATA

What is Tabulation?

- A table is a systematic arrangement of statistical data in column and rows”.
- “Tabulation is a process where by raw data on completed questionnaire are transformed in to the “list of needed information”.
- The purpose of table is to simplify the presentation and facilitate comparison
- Now, to analyze the collected data, it is essential to present it in an easy-to-understand and interpretable way. The different ways the classified data can be presented are textual, tabular, diagrammatic, and graphical.

Tabular Presentation or **Tabulation** is a systematic way of presenting numerical data in rows and columns.

- The tabular presentation helps the investigator in simplifying the presentation and facilitating analysis.
- It can bring the related information close to each other such that the investigator can easily make comparisons between them, and also helps in further statistical analysis and interpretation of the data.

According to L.R. Connor, "Tabulation involves the orderly and systematic presentation of numerical data in a form designed to elucidate the problem under consideration."

Components of table:

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

To tabulate data correctly, one must learn about the eight essential parts of a table. These are as follows –

- **Table Number** – This is the first part of a table and is given on top of any table to facilitate easy identification and for further reference.
- **Title of the Table** – One of the most important parts of any table is its title. The title is either placed just below the table number or at its right. It is imperative for the title to be brief, crisp and carefully-worded to describe the tables' contents effectively.
- **Headnote** – The headnote of a table is presented in the portion just below the title. It provides information about the unit of data in the table, like “amount in Rupees” or “quantity in kilograms”, etc.
- **Column Headings or Captions** – The headings of the columns are referred to as the caption. It consists of one or more column heads. A caption should be brief, short, and self-explanatory. Column heading is written in the middle of a column in small letters.
- **Row Headings or Stubs** – The title of each horizontal row is called a stub.
- **Body of a Table** – This is the portion that contains the numeric information collected from investigated facts. The data in the body is presented in rows which are read horizontally from left to right and in columns, read vertically from top to bottom.
- **Footnote** – Given at the bottom of a table above the source note, a footnote is used to state any fact that is not clear from the table's title, headings, caption or stub. For instance, if a table represents the profit earned by a company, a footnote can be used to state if said profit is earned before, or after tax calculations.
- **Source Note** – As its name suggests, a source note refers to the source from where the table's information has been collected.

Structure of Tabulation of Data

Tabulation is the process of systematically arranging data in rows and columns to make it easy to read, analyze, and interpret. A well-structured table has several **essential components**, each serving a specific purpose.

Stub(Heading)	Caption-1	Caption-2
Category-1	Data	Data

Stub(Heading)	Caption-1	Caption-2
Category-1	Data	Data
Category-1	Data	Data

Types of Tabulation

Tabulation is divided into three main types:

1. Simple Tabulation
2. Double Tabulation
3. Complex Tabulation

1. Simple Tabulation

This type of tabulation presents data based on a single characteristic or variable. It shows the number or frequency of observations in different categories of just one factor.

Gender	Number of People
Male	120
Female	130

This table focuses only on **gender**, which is **one characteristic**.

2. Double Tabulation

Double tabulation presents data based on two related characteristics. It helps in comparing the interaction or relationship between two variables at the same time.

Stream	Male	Female	Total
Science	40	30	70
Commerce	50	40	90
Arts	30	60	90

Stream	Male	Female	Total
Total	120	130	250

Here, data is classified by two characteristics: gender and stream.

3. Complex Tabulation

This type of tabulation involves **more than two characteristics** at the same time. It gives a more detailed and in-depth analysis of the data.

Location	Stream	Male	Female	Total
Urban	Science	25	20	45
	Commerce	30	25	55
	Arts	15	30	45
Rural	Science	15	10	25
	Commerce	20	15	35
	Arts	15	30	45

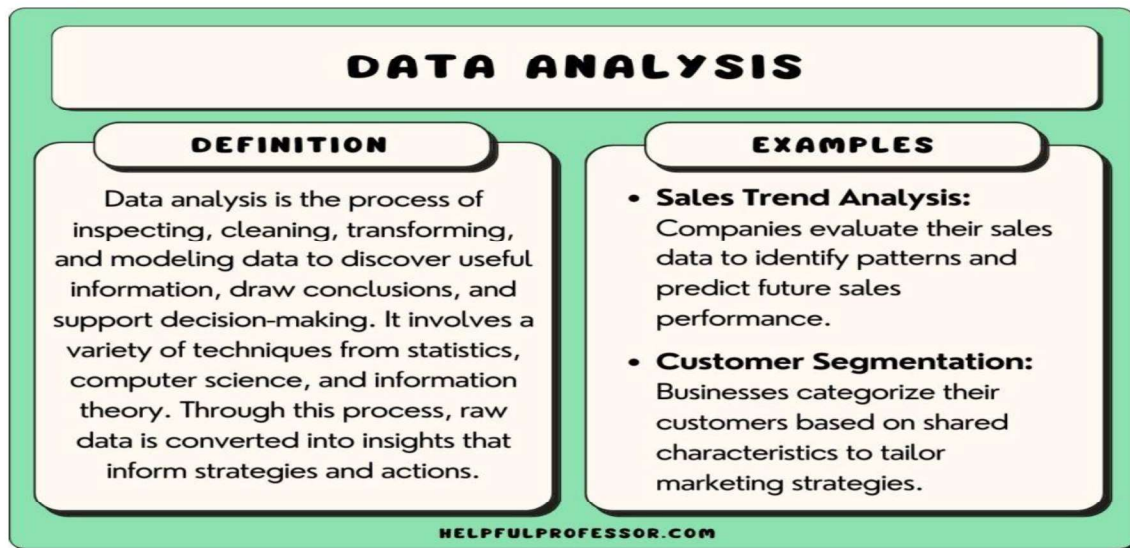
This table involves **three characteristics: Location, Stream, and Gender.**

Advantages of Tabulation of Data

- Tabulation arranges data systematically in rows and columns, making it easier to read and understand.
- Data presented in tables allows for straightforward comparison across different categories or groups.
- Tabulated data is easier to analyze, as it provides a clear and concise format that highlights key information.
- Tables condense large amounts of data into a compact format, saving time for those who need to interpret or analyze the data.

- By presenting data in a structured format, tabulation eliminates confusion and makes it easier to identify trends and patterns.
- Well-tabulated data provides a clear overview, aiding in quicker and more informed decision-making.
- Tables are essential for performing various statistical analyses, providing a foundation for calculations and evaluations.

DATA ANALYSIS



According to (Upton & Brawn, 2023).Data analysis is the process of inspecting, cleaning, transforming, and modeling data to discover useful information, derive conclusions, and support decision-making

Data analysis in research is the systematic use of statistical and analytical tools to describe, summarize, and draw conclusions from datasets. This process involves organizing, analyzing, modeling, and transforming data to identify trends, establish connections, and inform decision-making. The main goals include describing data through visualization and statistics, making inferences about a broader population, predicting future events using historical data, and providing data-driven recommendations. The stages of data analysis involve collecting relevant data, preprocessing to clean and format it, conducting exploratory data analysis to identify patterns, building and testing models, interpreting results, and effectively reporting findings.

- ❖ **Main Goals:** Describe data, make inferences, predict future events, and provide data-driven recommendations.

- ❖ **Stages of Data Analysis:** Data collection, preprocessing, exploratory data analysis, model building and testing, interpretation, and reporting.

Types of Data Analysis

1. Descriptive Analysis

Descriptive analysis focuses on summarizing and describing the features of a dataset. It provides a snapshot of the data, highlighting central tendencies, dispersion, and overall patterns.

- ❖ **Central Tendency Measures:** Mean, median, and mode are used to identify the central point of the dataset.
- ❖ **Dispersion Measures:** Range, variance, and standard deviation help in understanding the spread of the data.
- ❖ **Frequency Distribution:** This shows how often each value in a dataset occurs.

2. Inferential Analysis

Inferential analysis allows researchers to make predictions or inferences about a population based on a sample of data. It is used to test hypotheses and determine the relationships between variables.

- ❖ **Hypothesis Testing:** Techniques like t-tests, chi-square tests, and ANOVA are used to test assumptions about a population.
- ❖ **Regression Analysis:** This method examines the relationship between dependent and independent variables.
- ❖ **Confidence Intervals:** These provide a range of values within which the true population parameter is expected to lie.

3. Exploratory Data Analysis (EDA)

EDA is an approach to analyzing data sets to summarize their main characteristics, often with visual methods. It helps in discovering patterns, spotting anomalies, and checking assumptions with the help of graphical representations.

- ❖ **Visual Techniques:** Histograms, box plots, scatter plots, and bar charts are commonly used in EDA.
- ❖ **Summary Statistics:** Basic statistical measures are used to describe the dataset.

4. Predictive Analysis

Predictive analysis uses statistical techniques and machine learning algorithms to predict future outcomes based on historical data.

- ❖ **Machine Learning Models:** Algorithms like linear regression, decision trees, and neural networks are employed to make predictions.
- ❖ **Time Series Analysis:** This method analyzes data points collected or recorded at specific time intervals to forecast future trends.

5. Causal Analysis

Causal analysis aims to identify cause-and-effect relationships between variables. It helps in understanding the impact of one variable on another.

- ❖ **Experiments:** Controlled experiments are designed to test the causality.
- ❖ **Quasi-Experimental Designs:** These are used when controlled experiments are not feasible.

6. Mechanistic Analysis

Mechanistic analysis seeks to understand the underlying mechanisms or processes that drive observed phenomena. It is common in fields like biology and engineering.

Methods of Data Analysis

1. Quantitative Methods

Quantitative methods involve numerical data and statistical analysis to uncover patterns, relationships, and trends.

- ❖ **Statistical Analysis:** Includes various statistical tests and measures.
- ❖ **Mathematical Modeling:** Uses mathematical equations to represent relationships among variables.
- ❖ **Simulation:** Computer-based models simulate real-world processes to predict outcomes.

2. Qualitative Methods

Qualitative methods focus on non-numerical data, such as text, images, and audio, to understand concepts, opinions, or experiences.

- ❖ **Content Analysis:** Systematic coding and categorizing of textual information.
- ❖ **Thematic Analysis:** Identifying themes and patterns within qualitative data.
- ❖ **Narrative Analysis:** Examining the stories or accounts shared by participants.

3. Mixed Methods

Mixed methods combine both quantitative and qualitative approaches to provide a more comprehensive analysis.

- ❖ **Sequential Explanatory Design:** Quantitative data is collected and analyzed first, followed by qualitative data to explain the quantitative results.

- ❖ **Concurrent Triangulation Design:** Both qualitative and quantitative data are collected simultaneously but analyzed separately to compare results.

4. Data Mining

Data mining involves exploring large datasets to discover patterns and relationships.

- ❖ **Clustering:** Grouping data points with similar characteristics.
- ❖ **Association Rule Learning:** Identifying interesting relations between variables in large databases.
- ❖ **Classification:** Assigning items to predefined categories based on their attributes.

5. Big Data Analytics

Big data analytics involves analyzing vast amounts of data to uncover hidden patterns, correlations, and other insights.

- ❖ **Hadoop and Spark:** Frameworks for processing and analyzing large datasets.
- ❖ **NoSQL Databases:** Designed to handle unstructured data.
- ❖ **Machine Learning Algorithms:** Used to analyze and predict complex patterns in big data.

Applications and Case Studies

Numerous fields and industries use data analysis methods, which provide insightful information and facilitate data-driven decision-making. The following case studies demonstrate the effectiveness of data analysis in research:

Medical Care:

- ❖ **Predicting Patient Readmissions:** By using data analysis to create predictive models, healthcare facilities may better identify patients who are at high risk of readmission and implement focused interventions to enhance patient care.
- ❖ **Disease Outbreak Analysis:** Researchers can monitor and forecast disease outbreaks by examining both historical and current data. This information aids public health authorities in putting preventative and control measures in place.

Finance:

- ❖ **Fraud Detection:** To safeguard clients and lessen financial losses, financial institutions use data analysis tools to identify fraudulent transactions and activities.
- ❖ **Investing Strategies:** By using data analysis, quantitative investing models that detect trends in stock prices may be created, assisting investors in optimizing their portfolios and making well-informed choices.

Marketing:

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

- ❖ Customer Segmentation: Businesses may divide up their client base into discrete groups using data analysis, which makes it possible to launch focused marketing efforts and provide individualized services.
- ❖ Social Media Analytics: By tracking brand sentiment, identifying influencers, and understanding consumer preferences, marketers may develop more successful marketing strategies by analyzing social media data.

Education:

- ❖ Predicting Student Performance: By using data analysis tools, educators may identify at-risk children and forecast their performance. This allows them to give individualized learning plans and timely interventions.
- ❖ Education Policy Analysis: Data may be used by researchers to assess the efficacy of policies, initiatives, and programs in education, offering insights for evidence-based decision-making.

Social Science Fields:

- ❖ Opinion mining in politics: By examining public opinion data from news stories and social media platforms, academics and policymakers may get insight into prevailing political opinions and better understand how the public feels about certain topics or candidates.
- ❖ Crime Analysis: Researchers may spot trends, anticipate high-risk locations, and help law enforcement use resources wisely in order to deter and lessen crime by studying crime data.

Conclusion

Data analysis is a crucial step in the research process because it enables companies and researchers to glean insightful information from data. By using diverse analytical methodologies and approaches, scholars may reveal latent patterns, arrive at well-informed conclusions, and tackle intricate research inquiries. Numerous statistical, machine learning, and visualization approaches are among the many data analysis tools available, offering a comprehensive toolbox for addressing a broad variety of research problems.

UNIT-IV

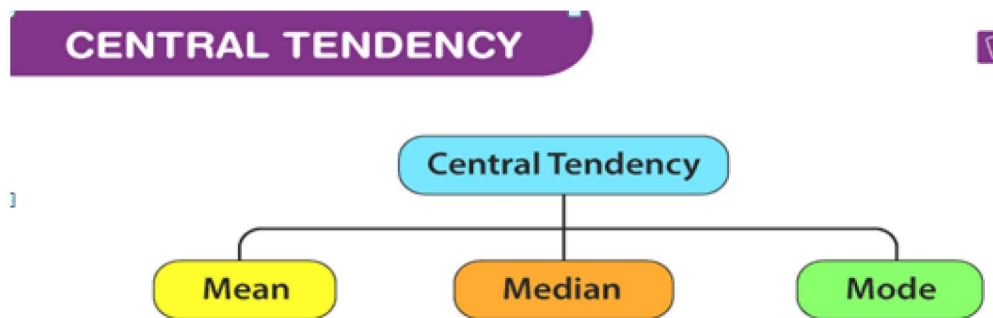
STATISTICAL TOOLS FOR ANALYSIS

MEASURES OF CENTRAL TENDENCY IN STATISTICS

Central Tendencies are the numerical values that are used to represent a large collection of numerical data. These numerical values are called measures of central tendency or average values. A central or average value of any statistical data or series is the variable's value representative of the entire data or its associated frequency distribution.

Some commonly used measures of central tendency are:

- **Mean**
- **Median**
- **Mode**



MEAN (OR) ARITHMETIC MEAN($\bar{x}$) (OR) AVERAGE

The mean is a statistical measure used to determine the average value in a dataset. It represents the central or typical value of the data. While the arithmetic mean is the most common, there are also other types like the geometric mean and harmonic mean. These are calculated differently, depending on the situation. It is sensitive to outliers, i.e, if there are extreme values in our dataset then it could cause a different mean rather than the typical mean.

Mean (in statistics) is the *average* of a set of numbers. It is one of the most important measures of central tendency in distributed data. To calculate the mean, add all the values in the data set and then divide the total by the number of values.

The mean is usually denoted by $\bar{x}$ (read as “x bar”).

Mean Formula

The mean formula in statistics is defined as the sum of all observations in the given dataset divided by the total number of observations.

Mean Formula

$$\text{Mean} = \frac{\text{Sum of All Data Points}}{\text{Number of Data points}}$$

1. INDIVIDUAL SERIES

Example: 2, 3, 8, 4

2. DISCRETE SERIES

Example:

x : 20	30	40	50	60	70
f : 8	12	20	10	6	4

3. CONTINUOUS SERIES

Example:

CI	f
10-20	4
20-30	7
30-40	16
40-50	20
50-60	15
60-70	8

ARITHMETIC MEAN PROBLEMS

1. ARITHMETIC MEAN FOR INDIVIDUAL SERIES

1. Mean for Individual Series

Formula:

$$\bar{x} = (\text{Sum of observations}) / (\text{Number of observations})$$

Example:

14, 24, 29, 10, 16, 12, 4, 9, 18

Sum = 136

Number of items = 9

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

$$\text{Mean} = 136 / 9 = 15.11$$

2. ARITHMETIC MEAN FOR DISCRETE SERIES

Formula:

$$\bar{x} = \Sigma fx / \Sigma f$$

Where:

f = frequency

x = variable

Σfx = total of f multiplied by x

Σf = total frequency

Problem:

x	20	30	40	50	60	70
f	8	12	20	10	6	4

Solution:

x	f	fx
20	8	160
30	12	360
40	20	800
50	10	500
60	6	360
70	4	280

$$\Sigma f = 60$$

$$\Sigma fx = 2460$$

$$\text{Mean} = 2460 / 60 = 41$$

3. ARITHMETIC MEAN FOR CONTINUOUS SERIES

Formula:

$$\bar{x} = \Sigma fx / \Sigma f$$

Problem:

class interval	f
10-20	4
20-30	7
30-40	16
40-50	20

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

50-60	15
60-70	8

Solution:

Mid-values (x): 15, 25, 35, 45, 55, 65

x mid values	f	fx
15	4	60
25	7	175
35	16	560
45	20	900
55	15	825
65	8	520

$$\Sigma f = 70$$

$$\Sigma fx = 3040$$

$$\text{Mean} = 3040 / 70 = 43.43$$

MEAN FOR CONTINUOUS SERIES – STEP DEVIATION METHOD

Formula:

$$\bar{x} = A + (\Sigma f \cdot d) / N \times C$$

Where:

A = assumed mean

$$d = (x - A) / C$$

C = class interval

x = mid value

Problem and solution

CI	f	x mid values	dx=x-a%c	fdx
10-20	15	15	-3	-45
20-30	8	25	-2	-16
30-40	6	35	-1	-6
40-50	9	45 Assumed mean	0	0
50-60	2	55	1	2
60-70	18	65	2	36

$$N = 58$$

$$\Sigma f \cdot d = -29$$

$$A = 45$$

$$C = 10$$

$$\text{Mean} = 45 + (-29 / 58) \times 10$$

$$\text{Mean} = 45 - 5$$

$$\text{Mean} = 40 \text{ (approx)}$$

MEDIAN

The median is the middle value in a set of numbers when they're arranged in order, either from smallest to largest or vice versa. If there's an odd number of numbers, the median is the one right in the middle. If there's an even number, the median is the average of the two middle numbers.

MEDIAN PROBLEMS

1. CALCULATION OF MEDIAN FOR INDIVIDUAL SERIES

(i) When number of terms is odd

Example Data:

16, 19, 21, 16, 18, 24, 22, 20

Arrange:

16, 17, 18, 19, 20, 21, 22, 24

Median = Middle value

Median = 19

(ii) When number of terms is even

Example Data:

13, 17, 29, 18, 24, 21, 31, 27, 34, 28

Arrange:

13, 17, 18, 21, 24, 25, 27, 28, 29, 31, 34

Middle values = 24 and 25

Median = $(24 + 25) / 2$

Median = 24.5

2. MEDIAN FOR DISCRETE SERIES

Formula:

Median = Value of $(N + 1) / 2$ th term

Example problem:

$x = 10, 20, 30, 40, 50, 60, 70$

$f = 4, 7, 21, 32, 25, 12, 3$

+-----+-----+-----+

Solution:

| x | f | cf |

+-----+-----+-----+

| 10 | 4 | 4 |

| 20 | 7 | 11 |

| 30 | 21 | 32 |

| 40 | 32 | 64 |

| 50 | 25 | 89 |

| 60 | 12 | 101 |

| 70 | 3 | 104 |

+-----+-----+-----+

$N = 104$

Median Position = $(104 + 1) / 2 = 52.5$ th term

The 52.5th term lies in $x = 40$

Therefore:

Median = 40

3. MEDIAN FOR CONTINUOUS SERIES

Formula:

Median = $L + [(N/2 - cf) / f] \times C$

Where:

L = Lower limit of median class

cf = cumulative frequency before median class

f = frequency of median class

C = class interval

Problem

+-----+-----+

| C.I | f |

+-----+-----+

45–50 10
40–45 15
35–40 26
30–35 30
25–30 48
20–25 31
15–20 24
10–15 7
5–10 15
+-----+-----+

SOLUTION:

Now add cumulative frequency:

+-----+-----+
C.I f cf
+-----+-----+
5–10 15 15
10–15 7 22
15–20 24 46
20–25 31 77
25–30 48 125
30–35 30 155
35–40 26 181
40–45 15 196
45–50 10 206
+-----+-----+

$$N = 206$$

$$N/2 = 103$$

The 103rd term falls in **25–30** → Median Class

$$L = 25$$

$$cf = 77$$

$$f = 48$$

$$C = 5$$

$$\text{Median} = 25 + [(103 - 77) / 48] \times 5$$

$$\text{Median} = 25 + (26 / 48) \times 5$$

$$\text{Median} = 25 + 2.708$$

$$\text{Median} \approx 27.71$$

MODE

Mode is the value of the observation that occurs with the highest frequency. In other, that observation of the data occurs the maximum number of times in a dataset.

MODE PROBLEMS

1. MODE FOR INDIVIDUAL SERIES

Definition:

Mode = the value that occurs the maximum number of times.

Example:

Find the mode for the following data:

18, 14, 16, 18, 26, 16, 20, 16, 11, 12, 16, 15, 20, 20

Most repeating value = **16**

Mode = 16

2. MODE FOR DISCRETE SERIES

Example:

x = 28, 29, 30, 31, 28, 33

f = 10, 20, 40, 65, 50, 15

Highest frequency = 65

Corresponding value = 31

Mode = 31

3. MODE FOR CONTINUOUS SERIES

Formula:

$$\text{Mode} = L + (f_1 - f_0) / (2f_1 - f_0 - f_2) \times C$$

Where:

L = Lower limit of modal class

f₀ = frequency of class before modal class

f₁ = frequency of modal class

f₂ = frequency of class after modal class

C = class interval

MEASURES OF DISPERSION

Dispersion refers to the degree to which data values in a dataset are spread out or scattered around an average (such as the mean or median). It helps us understand the variability or consistency within the data — whether the values are closely grouped or widely spread.

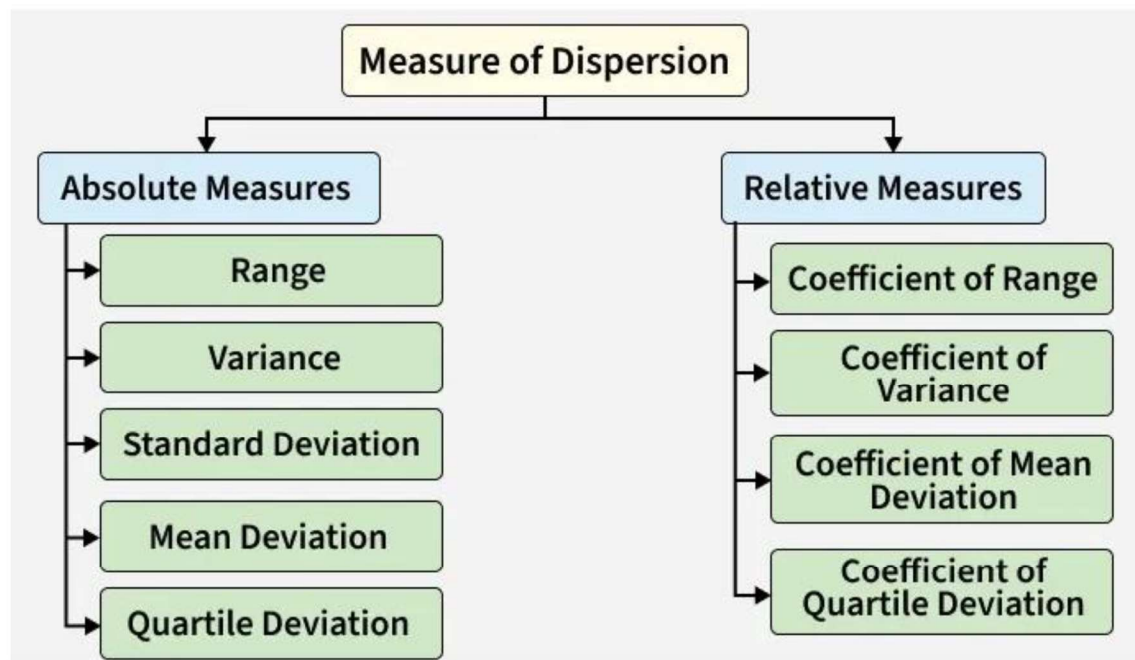
In statistics, measures of dispersion are numerical values that describe the extent of variation among data points. They provide insight into how the data is distributed and help compare the reliability or uniformity of different datasets.

The most common measures of dispersion include:

- **Range** – the difference between the maximum and minimum values.
- **Variance** – the average of the squared deviations from the mean.
- **Standard Deviation** – the square root of the variance, showing dispersion in the same units as the data.

Types of Measures of Dispersion

Measures of dispersion can be classified into the following two types:



These measures of dispersion can be further divided into various categories. They have various parameters, and these parameters have the same unit.

Absolute Measure of Dispersion

The measures of dispersion that are measured and expressed in the units of data themselves are called the Absolute Measure of Dispersion.

For example - Meters, Dollars, Kg, etc.

Some absolute measures of dispersion are:

- **Range:** It is defined as the difference between the largest and the smallest value in the distribution. The measure of dispersion is used to variability of the data.
- **Mean deviation:** It is the arithmetic mean of the differences between the values and their mean.
- **Standard Deviation:** It is the square root of the arithmetic average of the squares of the deviations measured from the mean.
- **Variance:** It is defined as the average of the square deviation from the mean of the given data set.
- **Quartile Deviation:** It is defined as half of the difference between the third quartile and the first quartile in a given data set.
- **Interquartile Range:** The difference between the upper(Q3) and lower(Q1) quartiles is called the Interquartile Range. Its formula is given as $Q3 - Q1$.

Relative Measure of Dispersion

We use relative measures of dispersion to measure the two quantities that have different units to get a better idea of the scattering of the data.

Here are some of the relative measures of dispersion:

- **Coefficient of Range:** It is defined as the ratio of the difference between the highest and lowest value in a data set to the sum of the highest and lowest value.
- **Coefficient of Variation:** It is defined as the ratio of the standard deviation to the mean of the data set. We use percentages to express the coefficient of variation.
- **Coefficient of Mean Deviation:** It is defined as the ratio of the mean deviation to the value of the central point of the data set.
- **Coefficient of Quartile Deviation:** It is defined as the ratio of the difference between the third quartile and the first quartile to the sum of the third and first quartiles.

PROBLEMS ON STANDARD DEVIATION (σ)

1. STANDARD DEVIATION FOR INDIVIDUAL SERIES

Formula:

$$\sigma = \sqrt{(\Sigma x^2 / n)}$$

Where:

- $x = X - \bar{X}$
- n = number of observations

Example:

Calculate standard deviation for the following data:

240, 260, 290, 245, 255, 288, 272, 263, 277, 251

Solution:

240, 260, 290, 245, 255, 288, 272, 263, 277, 251

Mean ($\bar{X}$):

$$\Sigma X = 2641$$

$$n = 10$$

$$\bar{X} = \Sigma X / n = 2641 / 10 = 264.1$$

Table

X	$X - \bar{X}$	$(X - \bar{X})^2$
240	-24.1	580.81
260	-4.1	16.81
290	25.9	670.81
245	-19.1	364.81
255	-9.1	82.81
288	23.9	571.21
272	7.9	62.41
263	-1.1	1.21
277	12.9	166.41
251	-13.1	171.61

$$\Sigma x^2 = 2688.9$$

Formula:

$$\sigma = \sqrt{(\Sigma x^2 / n)}$$

$$\Sigma x^2 = 2688.9$$

$$n = 10$$

So,

$$\sigma = \sqrt{(2688.9 / 10)}$$

$$\sigma = \sqrt{268.89}$$

$$\sigma = 16.3979$$

Standard Deviation (σ) = **16.3979**

2. STANDARD DEVIATION FOR DISCRETE SERIES

Direct Method Formula:

$$\sigma = \sqrt{(\Sigma fX^2 / N - (\Sigma fX / N)^2)}$$

Assumed Mean Method Formula:

$$\sigma = \sqrt{(\Sigma fd^2 / N - (\Sigma fd / N)^2) \times C}$$

Where:

- $d = X - A$
- A = assumed mean
- C = class width

Example:

Calculate standard deviation for the following data:

X	45	50	55	60	65	70	75	80
f	3	5	8	7	9	7	4	7

DISCRETE SERIES – ASSUMED MEAN METHOD

X	f	d = X – A	d ²	fd	fd ²
---	---	-----------	----------------	----	-----------------

45	3	–15	225	–45	675
50	5	–10	100	–50	500
55	8	–5	25	–40	200
60	7	0	0	0	0
65	9	5	25	45	225

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

70	7	10	100	70	700
75	4	15	225	60	900
80	7	20	400	140	2800

$$N = 50$$

$$\Sigma f d = 180$$

$$\Sigma f d^2 = 6000$$

Standard Deviation through Assumed Mean Method

Formula:

$$\sigma = \sqrt{(\Sigma f d^2 / N) - (\Sigma f d / N)^2}$$

Substitution:

$$\sigma = \sqrt{(6000 / 50 - (180 / 50)^2)}$$

$$\sigma = \sqrt{(120 - 12.96)}$$

$$\sigma = \sqrt{107.04}$$

$$\sigma = 10.346$$

✓ Standard Deviation = 10.346

3. STANDARD DEVIATION FOR CONTINUOUS SERIES

Direct Method Formula

$$\sigma = \sqrt{(\Sigma f X^2 / N - (\bar{X})^2)}$$

Assumed Mean Method Formula

$$\sigma = \sqrt{(\Sigma f d^2 / N) - (\Sigma f d / N)^2} \times C$$

Where:

- $d = (X - A) / C$
- A = Assumed Mean
- C = Class Interval Width

Example:

Calculate standard deviation for the following data:

CLASS INTERVALS 0–10, 10–20, 20–30, 30–40, 40–50, 50–60, 60–70
Frequencies (f): 1, 5, 17, 45, 26, 5, 2

Solution Table

C.I	f	X	f × X	d = (X-A)/C	d ²	f d	f d ²
-----	---	---	-------	-------------	----------------	-----	------------------

0-10	1	5	5	-3	9	-3	9	
10-20	4	15	60	-2	4	-8	16	
20-30	17	25	425	-1	1	-17	17	
30-40	45	35	1575	0	0	0	0	
40-50	26	45	1170	1	1	26	26	
50-60	5	55	275	2	4	10	20	
60-70	2	65	130	3	9	6	18	

$$N = 100$$

$$\Sigma fX = 3630$$

$$\Sigma f d = 14$$

$$\Sigma f d^2 = 106$$

Standard Deviation (Assumed Mean Method)

Formula:

$$\sigma = \sqrt{(\Sigma f d^2 / N - (\Sigma f d / N)^2) \times C}$$

Substitution:

$$\sigma = \sqrt{(106 / 100 - (14 / 100)^2) \times 10}$$

$$\sigma = \sqrt{(1.06 - 0.0196) \times 10}$$

$$\sigma = \sqrt{1.0404 \times 10}$$

$$\sigma = 1.02 \times 10$$

$$\sigma = 10.2$$

$$\text{Standard Deviation} = 10.2$$

PROBLEMS ON VARIANCE

1. VARIANCE FOR INDIVIDUAL SERIES

Formula

$$\text{Variance} = (\sigma / \bar{X}) \times 100$$

Where:

• σ = Standard Deviation

• $\bar{X}$ = Mean

$$\sigma = 16.3979$$

$$\bar{X} = 264.1$$

$$\text{Variance} = (16.3979 / 264.1) \times 100$$

$$\text{Variance} = \mathbf{6.21}$$

2. VARIANCE FOR DISCRETE SERIES

Variance Formula

$$\text{Variance} = (\sigma / \bar{X}) \times 100$$

Where:

- σ = Standard deviation
- $\bar{X}$ = Mean

Example:

From standard deviation (Discrete Series)

Given:

$$\sigma = 10.346$$

Table

X	f	fX
45	3	135
50	5	250
55	8	440
60	7	420
65	9	585
70	7	490
75	4	300
80	7	560
Total	50	3180

Mean

$$\text{Mean } (\bar{X}) = \Sigma fX / N$$

$$= 3180 / 50$$

$$= \mathbf{63.6}$$

Variance Calculation

$$\text{Variance} = (\sigma / \bar{X}) \times 100$$

$$= (10.346 / 63.6) \times 100$$

$$= 16.2673$$

Therefore,

$$\text{Variance} = 16.2673$$

3. VARIANCE FOR CONTINUOUS SERIES

Variance Formula

$$\text{Variance} = (\sigma / \bar{X}) \times 100$$

Example: From Standard Deviation (Continuous Series)

$$\text{Mean } (\bar{X}) = \Sigma fX / N$$

$$= 3630 / 100$$

$$= 36.3$$

$$\text{Standard deviation } (\sigma) = 10.2$$

Variance Calculation

$$\text{Variance} = (\sigma / \bar{X}) \times 100$$

$$= (10.2 / 36.3) \times 100$$

$$= 28.0991$$

$$\approx 28.1$$

Using Step-Deviation Method

$$\text{Mean } (\bar{X}) = A + (\Sigma f \cdot dx / N) \times C$$

$$A = 35$$

$$\Sigma f \cdot dx = 14$$

$$N = 100$$

$$C = 10$$

Mean:

$$\bar{X} = 35 + (14 / 100) \times 10$$

$$\bar{X} = 36.4$$

Variance Calculation (Step-Deviation Method)

$$\text{Variance} = (\sigma / \bar{X}) \times 100$$

$$= (10.2 / 36.4) \times 100$$

$$= 28.0219$$

$$\text{Variance} = 28.0219$$

CORRELATION AND REGRESSION

Correlation and regression are essential statistical tools used to analyze the relationship between variables. **Correlation** measures the strength and direction of a linear relationship between two variables, indicating how one variable changes in response to another. **Regression**, on the other hand, goes a step further by not only measuring this relationship but also predicting the value of a dependent variable based on one or more independent variables.

CORRELATION

Correlation quantifies how strong and in which direction the two variables are related. It is being measured using the correlation coefficient, which goes to the extent of negative one and positive one.

- **Positive Correlation:** In this case, if the value of one of them goes up, then the value of the other one also goes up.
- **Negative Correlation:** That is, when one variable is high, the other is low, and vice versa.
- **No Correlation:** Not statistically significantly correlated: partners, financial wealth, and life satisfaction.

Types of Correlation Coefficients

Some of the common correlation coefficients are:

- **Pearson Correlation Coefficient:** Calculates the form of a straight line that describes the direction of the relationship between two variables that are both of the interval or ratio scale.
- **Spearman's Rank Correlation:** Used to establish the level of connection between two variables that have been ranked.
- **Kendall's Tau:** Measures the extent to which one variable is ranked relative to another.

CORRELATION PROBLEMS

1. Direct Method

Formula:

$$r = \frac{N \sum xy - \sum x \sum y}{\sqrt{N \sum x^2 - (\sum x)^2} \sqrt{N \sum y^2 - (\sum y)^2}}$$

2. Assumed Mean Method

Formula:

$$r = \frac{N \sum d_x d_y - \sum d_x \sum d_y}{\sqrt{N \sum d_x^2 - (\sum d_x)^2} \sqrt{N \sum d_y^2 - (\sum d_y)^2}}$$

where

$$d_x = x - A$$

$$d_y = y - A$$

Example: Find correlation for the following data

Temperature (X):

57, 42, 40, 38, 42, 45, 42, 44, 40, 46, 44, 43

Time (Y):

10, 26, 30, 41, 29, 27, 27, 19, 18, 19, 31, 29

Solution

Correlation Table

Temp (x)	Time (y)	dx	Dx^2	dy	Dy^2	dx dy
57	10	13	169	-16	256	-208
42	26	-2	4	0	0	0
40	30	-4	16	4	16	-16
38	41	-6	36	15	225	-90
42	29	-2	4	3	9	-6

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Temp (x)	Time (y)	dx	Dx^2	dy	Dy^2	dxdy
45	27	1	1	1	1	1
42	27	-2	4	1	1	-2
44	19	0	0	-7	49	0
40	18	4	16	-8	64	-32
46	19	2	4	-7	49	-14
44	31	0	0	5	25	0
43	29	-1	1	3	9	-3

Totals

$$\sum d_x = -5, \quad \sum d_x^2 = 255, \quad \sum d_y = -6, \quad \sum d_y^2 = 704, \quad \sum d_x d_y = -306$$

$N = 12$ observations

Correlation Formula (Assumed Mean Method)

$$r = \frac{N \sum d_x d_y - (\sum d_x)(\sum d_y)}{\sqrt{N \sum d_x^2 - (\sum d_x)^2} \sqrt{N \sum d_y^2 - (\sum d_y)^2}}$$

Substitution

$$\begin{aligned} r &= \frac{12(-306) - (-5)(-6)}{\sqrt{12(255) - (-5)^2} \sqrt{12(704) - (-6)^2}} \\ &= \frac{-3702}{\sqrt{3035} \sqrt{8412}} \end{aligned}$$

Therefore, Correlation (r) = -0.7327.

REGRESSION

Regression analysis can be defined as a statistical tool that is applied in an endeavor to identify the relationship between a dependent variable and one or many independent variables. Of benefit when trying to forecast the position of the dependent variable, given the position of the independent variables.

Types of Regression

Common types of regression are:

- **Simple Linear Regression:** Looks at two variables and tries to find out if one makes some contribution to the other by putting in place a linear equation.
- **Multiple Linear Regression:** Looks at the effect of more than one independent variable on one dependent variable.
- **Logistic Regression:** Applied where the dependent variable is qualitative.
- **Polynomial Regression:** Plans the connection between the dependent variable and the independent variable by an nth-degree polynomial.

PROBLEMS ON REGRESSION

There are **two types** of regression equations:

1. **y on x**
2. **x on y**

1. Regression of y on x

General form:

$$\hat{y} - \bar{y} = b_{yx}(x - \bar{x})$$

Where,

$$b_{yx} = \frac{n \sum d_x d_y - (\sum d_x)(\sum d_y)}{n \sum d_x^2 - (\sum d_x)^2}$$

OR

$$\hat{y} - \bar{y} = r \cdot \frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

And,

$$r \cdot \frac{\sigma_y}{\sigma_x} = \frac{\sum d_x d_y}{\sum d_x^2}$$

2. Regression of x on y

General form:

$$\hat{x} - \bar{x} = b_{xy}(y - \bar{y})$$

Where,

$$b_{xy} = \frac{n \sum d_x d_y - (\sum d_x)(\sum d_y)}{n \sum d_y^2 - (\sum d_y)^2}$$

OR



$$\hat{x} - \bar{x} = r \cdot \frac{\sigma_x}{\sigma_y} (y - \bar{y})$$

And,

$$r \cdot \frac{\sigma_x}{\sigma_y} = \frac{\sum d_x d_y}{\sum d_y^2}$$

Problem: Find regression equations of both y on x and x on y for the following data

X: 91, 97, 108, 121, 67, 124, 51, 73, 111, 57

Y: 71, 75, 69, 97, 70, 91, 39, 61, 80, 47

Solution

X	Y	dx=x-x̄	dx ²	dy=y-ȳ	dy ²	dx dy
91	71	1	1	1	1	1
97	75	7	49	5	25	35
108	69	18	324	-1	1	-18
121	97	31	961	27	729	837
67	70	-23	529	0	0	0
124	91	34	1156	21	441	714
51	39	-39	1521	-31	961	1209
73	61	-17	289	-9	81	153
111	80	21	441	10	100	210
57	47	-33	1089	-23	529	759

Totals

$$\begin{aligned}\sum x &= 900, & \sum y &= 700 \\ \sum d_x &= 0, & \sum d_y &= 0 \\ \sum d_x^2 &= 6360, & \sum d_y^2 &= 2868 \\ \sum d_x d_y &= 3900\end{aligned}$$

Means

$$\begin{aligned}\bar{x} &= \frac{\sum x}{10} = \frac{900}{10} = 90 \\ \bar{y} &= \frac{\sum y}{10} = \frac{700}{10} = 70\end{aligned}$$

1. Regression of y on x

General form:

$$\hat{y} - \bar{y} = r \cdot \frac{\sigma_y}{\sigma_x} (x - \bar{x})$$

Now,

$$\begin{aligned}r \cdot \frac{\sigma_y}{\sigma_x} &= \frac{\sum d_x d_y}{\sum d_x^2} \\ &= \frac{3900}{6360} \\ &= 0.6132\end{aligned}$$

So,

$$(\hat{y} - 70) = 0.6132(x - 90)$$

$$\hat{y} = 0.6132x - 55.188 + 70$$

$$\boxed{\hat{y} = 0.6132x + 14.812}$$

2. Regression of x on y

General form:

$$\hat{x} - \bar{x} = r \cdot \frac{\sigma_x}{\sigma_y} (y - \bar{y})$$

Now,

$$\begin{aligned} r \cdot \frac{\sigma_x}{\sigma_y} &= \frac{\sum d_x d_y}{\sum d_y^2} \\ &= \frac{3900}{2868} \\ &= 1.3598 \end{aligned}$$

So,

$$(\hat{x} - 90) = 1.3598(y - 70)$$

$$\hat{x} = 1.3598y - 95.186 + 90$$

$$\boxed{\hat{x} = 1.3598y - 5.186}$$



ASSUMPTIONS ABOUT PARAMETRIC AND NON- PARAMETRIC TEST

Statistical analysis plays a crucial role in understanding and interpreting data across various disciplines. Two prominent approaches in statistical analysis are **Parametric and Non-Parametric Methods**. While both aim to draw inferences from data, they differ in their assumptions and underlying principles. This article delves into the differences between these two methods, highlighting their respective strengths and weaknesses, and providing guidance on choosing the appropriate method for different scenarios.

PARAMETRIC METHODS

Parametric methods are statistical techniques that rely on specific assumptions about the underlying distribution of the population being studied. These methods typically assume that the data follows a known Probability distribution, such as the normal distribution, and estimate the parameters of this distribution using the available data.

The basic idea behind the **Parametric method** is that there is a set of fixed parameters that are used to determine a probability model that is used in Machine Learning as well. Parametric methods are those methods for which we priory know that the population is normal, or if not then we can easily approximate it using a Normal Distribution which is possible by invoking the Central Limit Theorem.

Parameters for using the normal distribution are as follows:

- Mean
- Standard Deviation

Eventually, the classification of a method to be parametric completely depends on the presumptions that are made about a population.

Assumptions for Parametric Methods

Parametric methods require several assumptions about the data:

- **Normality:** The data follows a normal (Gaussian) distribution.
- **Homogeneity of variance:** The variance of the population is the same across all groups.
- **Independence:** Observations are independent of each other.

What are Parametric Methods?

- **Statistical Tests:**
 - **t-test:** Tests for the difference between the means of two independent groups.

- **ANOVA:** Tests for the difference between the means of three or more groups.
- **F-test:** Compares the variances of two groups.
- **Chi-square test:** Tests for relationships between categorical variables.
- **Correlation analysis:** Measures the strength and direction of the linear relationship between two continuous variables.
- **Machine Learning Models:**
 - **Linear regression:** Predicts a continuous outcome based on a linear relationship with one or more independent variables.
 - **Logistic regression:** Predicts a binary outcome (e.g., yes/no) based on a set of independent variables.
 - **Naive Bayes:** Classifies data points based on Bayes' theorem and assuming independence between features.
 - **Hidden Markov Models:** Models sequential data with hidden states and observable outputs.

some more common parametric methods available some of them are:

- Confidence interval used for - population mean along with known standard deviation.
- The confidence interval is used for - population means along with the unknown standard deviation.
- The confidence interval for population variance.
- The confidence interval for the difference of two means, with unknown standard deviation.

Advantages of Parametric Methods

- **More powerful:** When the assumptions are met, parametric tests are generally more powerful than non-parametric tests, meaning they are more likely to detect a real effect when it exists.
- **More efficient:** Parametric tests require smaller sample sizes than non-parametric tests to achieve the same level of power.
- **Provide estimates of population parameters:** Parametric methods provide estimates of the population mean, variance, and other parameters, which can be used for further analysis.

Disadvantages of Parametric Methods

- **Sensitive to assumptions:** If the assumptions of normality, homogeneity of variance, and independence are not met, parametric tests can be invalid and produce misleading results.
- **Limited flexibility:** Parametric methods are limited to the specific probability distribution they are based on.
- **May not capture complex relationships:** Parametric methods are not well-suited for capturing complex non-linear relationships between variables.

Applications of Parametric Methods

Parametric methods are widely used in various fields, including:

- **Biostatistics:** Comparing the effectiveness of different treatments.
- **Social sciences:** Investigating relationships between variables.
- **Finance:** Estimating risk and return of investments.
- **Engineering:** Analyzing the performance of systems.

II. NON-PARAMETRIC METHODS

Non-parametric methods are statistical techniques that do not rely on specific assumptions about the underlying distribution of the population being studied. These methods are often referred to as "distribution-free" methods because they make no assumptions about the shape of the distribution.

The basic idea behind the non-parametric method is no need to make any assumption of parameters for the given population or the population we are studying. In fact, the methods don't depend on the population. Here there is no fixed set of parameters available, and also there is no distribution (normal distribution, etc.) of any kind is available for use. This is also the reason that nonparametric methods are also referred to as distribution-free methods. Nowadays Non-parametric methods are gaining popularity and an impact of influence some reasons behind this fame is:

- The main reason is that there is no need to be mannered while using non-parametric methods.
- The second important reason is that we do not need to make more and more assumptions about the population given (or taken) on which we are working on.
- Most of the nonparametric methods available are very easy to apply and to understand also i.e. the complexity is very low.

Assumptions of Non-Parametric Methods

Non Parametric methods require several assumptions about the data:

1. **Independence:** Data points are independent and not influenced by others.
2. **Random Sampling:** Data represents a random sample from the population.
3. **Homogeneity of Measurement:** Measurements are consistent across all data points.

What is Non-Parametric Methods?

- **Statistical Tests:**
 - **Mann-Whitney U test:** Tests for the difference between the medians of two independent groups.
 - **Kruskal-Wallis test:** Tests for the difference between the medians of three or more groups.
 - **Spearman's rank correlation:** Measures the strength and direction of the monotonic relationship between two variables.
 - **Wilcoxon signed-rank test:** Tests for the difference between the medians of two paired samples.
- **Machine Learning Models:**
 - **K-Nearest Neighbors (KNN):** Classifies data points based on the k nearest neighbors.
 - **Decision Trees:** Makes classifications based on a series of yes/no questions about the features.
 - **Support Vector Machines (SVM):** Creates a decision boundary that maximizes the margin between different classes.
 - **Neural networks:** Can be designed with specific architectures to handle non-parametric data, such as convolutional neural networks for image data and recurrent neural networks for sequential data.

Advantages of Non-Parametric Methods

- **Robust to outliers:** Non-parametric methods are not affected by outliers in the data, making them more reliable in situations where the data is noisy.
- **Widely applicable:** Non-parametric methods can be used with a variety of data types, including ordinal, nominal, and continuous data.
- **Easy to implement:** Non-parametric methods are often computationally simple and easy to implement, making them suitable for a wide range of users.

Disadvantages of Non-Parametric Methods

- **Less powerful:** When the assumptions of parametric methods are met, non-parametric tests are generally less powerful, meaning they are less likely to detect a real effect when it exists.
- **May require larger sample sizes:** Non-parametric tests may require larger sample sizes than parametric tests to achieve the same level of power.
- **Less information about the population:** Non-parametric methods provide less information about the population parameters than parametric methods.

Applications of Non-Parametric Methods

Non-parametric methods are widely used in various fields, including:

- **Medicine:** Comparing the effectiveness of different treatments.
- **Psychology:** Investigating relationships between variables.
- **Ecology:** Analyzing environmental data.
- **Computer science:** Developing machine learning algorithms.

Difference between Parametric and Non-Parametric

There are several Difference between Parametric and Non-Parametric Methods are as follows:

Parametric Methods	Non-Parametric Methods
Parametric Methods uses a fixed number of parameters to build the model.	Non-Parametric Methods use the flexible number of parameters to build the model.
Parametric analysis is to test group means.	A non-parametric analysis is to test medians.
It is applicable only for variables.	It is applicable for both – Variable and Attribute.
It always considers strong assumptions	It generally fewer assumptions about data.

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Parametric Methods	Non-Parametric Methods
about data.	
Parametric Methods require lesser data than Non-Parametric Methods.	Non-Parametric Methods requires much more data than Parametric Methods.
Parametric methods assumed to be a normal distribution.	There is no assumed distribution in non-parametric methods.
Parametric data handles – Intervals data or ratio data.	But non-parametric methods handle original data.
Here when we use parametric methods then the result or outputs generated can be easily affected by outliers.	When we use non-parametric methods then the result or outputs generated cannot be seriously affected by outliers.
Parametric Methods can perform well in many situations but its performance is at peak (top) when the spread of each group is different.	Similarly, Non-Parametric Methods can perform well in many situations but its performance is at peak (top) when the spread of each group is the same.
Parametric methods have more statistical power than Non-Parametric methods.	Non-parametric methods have less statistical power than Parametric methods.
As far as the computation is considered these methods are computationally faster	As far as the computation is considered these methods are computationally slower than the

Parametric Methods	Non-Parametric Methods
than the Non-Parametric methods.	Parametric methods.
Examples: Logistic Regression, Naïve Bayes Model, etc.	Examples: KNN, Decision Tree Model, etc.

Conclusion

Parametric and non-parametric methods offer distinct advantages and limitations. Understanding these differences is crucial for selecting the most suitable method for a specific analysis. Choosing the appropriate method ensures valid and reliable inferences, enabling researchers to draw insightful conclusions from their data. As statistical analysis continues to evolve, both parametric and non-parametric methods will play crucial roles in advancing knowledge across various fields.

I. PARAMETRIC TEST

1. t- TEST PROBLEMS

PROBLEM 1: The average breaking strength of steel rods is specified to be **18.5 thousand kg**. To test the breaking strength, a sample of **14 rods** was taken. The mean and standard deviation obtained were **17.85 thousand kg** and **1.955 thousand kg**, respectively. Is there any significant deviation in the breaking strength of the rods?

Solution:

Let us take the null hypothesis that there is **no significant deviation** in the breaking strength of the rods.

Null Hypothesis (H_0): $\mu = 18.5$

Alternative Hypothesis (H_1): $\mu \neq 18.5$

- **Given:**
 - $n = 14$
 - Sample mean ($\bar{x}$) = 17.85
 - Sample standard deviation (s) = 1.955
 - Degrees of freedom (df) = $n - 1 = 13$
 - Level of significance (α) = 0.05

- The critical value of t at:
 $df = 13$ and $\alpha/2 = 0.025$ is:
 $t(\alpha/2) = 2.16$

Using the t -test statistic, we get

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{17.85 - 18.5}{1.955/\sqrt{14}} = \frac{-0.65}{0.522} = -1.24$$

Since $t_{cal} = -1.24$ is greater than its critical value $t_{\alpha/2} = -2.16$ at level of significance $\alpha/2 = 0.025$ and $df = 13$, the null hypothesis H_0 is accepted. Hence, we conclude that there is **no significant deviation** of sample mean from the population mean.

PROBLEM 2: An automobile tyre manufacturer claims that the average life of a particular grade of tyre is **more than 20,000 km** under normal conditions. A random sample of **16 tyres** was tested, and a mean and standard deviation of **22,000 km** and **5,000 km**, respectively, were computed. Assuming the tyre life is approximately normally distributed, decide whether the manufacturer's claim is valid.

Solution:

Let us take the null hypothesis that the manufacturer's claim is valid, that is:

$$H_0 : \mu \geq 20,000 \quad \text{and} \quad H_1 : \mu < 20,000$$

Given:

- $n = 16$
- $\bar{x} = 22,000$
- $s = 5,000$
- $df = 15$
- $\alpha = 0.05$

The critical value of t at $df = 15$ and $\alpha = 0.05$ is:

$$t_{\alpha} = 1.753$$

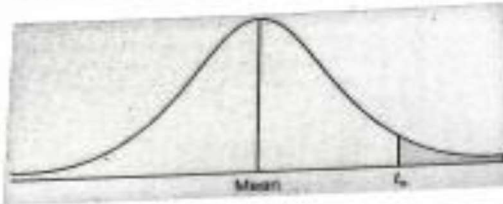
Using the t -test statistic, we get:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{22,000 - 20,000}{5000/\sqrt{16}} = \frac{2,000}{1250} = 1.60$$

Since $t_{cal} = 1.60$ is less than the critical value $t_{\alpha} = 1.753$ at 0.05 significance level and $df = 15$, the null hypothesis H_0 is accepted. Hence, we conclude that the manufacturer's claim is **valid**.

t-TEST TABLE VALUES

Table A4: Critical Values of Student's t-Distribution



Level of Significance for One-tailed Test					
d.f.	0.10	0.05	0.025	0.01	0.005
Level of Significance for One-tailed Test					
d.f.	0.20	0.10	0.05	0.02	0.01
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750

2. Z- TEST PROBLEMS

Problem1: A packaging device is set to fill detergent powder packets with a mean weight of **5 kg** and a standard deviation of **0.21 kg**. The weight of packets can be assumed to be normally distributed. The weight of packets is known to increase over a period of time due to machine fault, which is not tolerable. A random sample of **100 packets** is taken and weighed. This sample has a mean weight of **5.03 kg**. Can we conclude that the mean weight produced by the machine has increased? Use a **5 per cent** level of significance.

Solution:

Let the null hypothesis be that the mean weight of packets produced by the machine has not increased, i.e.,

$$H_0: \mu \geq 5 \text{ and}$$

$$H_1: \mu < 5$$

Given:

- $n = 100$
- $\bar{x} = 5.03 \text{ kg}$
- $\sigma = 0.21 \text{ kg}$
- $\alpha = 5\%$

Using the z-test statistic, we have:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{5.03 - 5}{0.21 / \sqrt{100}} = \frac{0.03}{0.021} = 1.428$$

Since the calculated value

$$z_{cal} = 1.428$$

is less than the critical value

$$z_{\alpha} = 1.645 \quad (\alpha = 0.05),$$

the null hypothesis H_0 is accepted. Hence, we conclude $\downarrow$ that the mean weight produced by the machine **has not increased** beyond 5 kg.

Problem 2: The mean life time of a sample of **400 fluorescent light bulbs** produced by a company is found to be **1600 hours** with a standard deviation of **150 hours**. Test the hypothesis that the mean life time of the bulbs produced in general is **higher than the mean life of 1570 hours** at $\alpha=0.01$ level of significance.

Solution:

Let the null hypothesis be that the mean life time of bulbs is higher than 1570 hours, i.e.,

$$H_0 : \mu \geq 1570 \quad \text{and} \quad H_1 : \mu < 1570$$

Given:

- $n = 400$
- $\bar{x} = 1600$ hours
- $s = 150$ hours
- $\alpha = 0.01$

Using the z-test statistic:

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{1600 - 1570}{150/\sqrt{400}} = \frac{30}{7.5} = 4$$

Since the calculated value

$$z_{cal} = 4$$

is more than the critical value

$$z_{\alpha} = \pm 2.33,$$

the null hypothesis H_0 is rejected.

Hence, we conclude that the mean lifetime of bulbs produced by the company is **not higher than 1570 hours**.

Problem 3: A continuous manufacturing process of steel rods is said to be in *state of control* and produces acceptable rods if the mean diameter of all rods produced is **2 inches**. Although the process standard deviation exhibits stability over time with standard deviation $\sigma=0.01$ inch the process mean may vary due to operator error or problems of process adjustment. Periodically, random samples of **100 rods** are selected to determine whether the process is producing acceptable rods. If result of a test indicates that the process is out of control it is stopped and source of trouble is identified. Otherwise it is allowed to continue operating a random sample of 100 rods is selected resulting in a mean of 2.1 inches test the hypothesis to determine whether the process is be continued ?

Solution: Since rods that are either too narrow or too wide are unacceptable, the low values and high values of the sample mean lead to the rejection of the null hypothesis. Consider the null hypothesis H_0 that the process may be allowed to continue when the diameter is **2**

inches. Consequently, rejection region is on both tails of the sampling distribution. The null and alternative hypotheses are stated as follows:

$$H_0 : \mu = 2 \text{ inches (continue process)}$$

$$H_1 : \mu \neq 2 \text{ inches (stop the process)}$$

Given:

- $n = 100$
- $\bar{x} = 2.1$
- $\sigma = 0.01$
- $\alpha = 0.01$

Using the z-test statistic, we have:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{2.1 - 2}{0.01/\sqrt{100}} = \frac{0.1}{0.001} = 100$$

Since the calculated value

$$z_{cal} = 100$$

is more than its critical value

$$z_{\alpha/2} = 2.58 \quad (\alpha = 0.01),$$

the null hypothesis H_0 is rejected.

Thus, **stop the process** in order to identify the source of trouble.

PROBLEM 4: An ambulance service claims that it takes on the average **8.9 minutes** to reach its destination in emergency calls. To check on this claim, the agency which licenses ambulance services has then timed **50 emergency calls**, getting a mean of **9.3 minutes** with a standard deviation of **1.8 minutes**. Does this indicate that the average time claimed is too low at the **one per cent significance level**?

Solution:

Let the null hypothesis H_0 = Average time claimed and observed are the same.

The null and alternative hypotheses are stated as follows:

Given:

- $n = 50$
- $\bar{x} = 9.3$
- $s = 1.8$

Using the z-test statistic, we get:

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{9.3 - 8.9}{1.8/\sqrt{50}} = \frac{0.4}{0.254} = 1.574$$

Since the calculated value

$$z_{cal} = 1.574$$

is less than the critical value

$$z_{\alpha/2} = \pm 2.58 \quad \text{at } \alpha = 0.01,$$

the null hypothesis H_0 is accepted.

Hence, we conclude that there is no difference between the average time observed and claimed.

Z-TABLE VALUES

Critical values of z

Critical value Z_{α}	Level of Significance		
	1%	5%	10%
Two Tailed Test	$ Z_{\alpha/2} = 2.58$	$ Z_{\alpha/2} = 1.96$	$ Z_{\alpha/2} = 1.645$
Right Tailed Test	$Z_{\alpha} = 2.33$	$Z_{\alpha} = 1.645$	$Z_{\alpha} = 1.28$
Left Tailed Test	$-Z_{\alpha} = -2.33$	$-Z_{\alpha} = -1.645$	$-Z_{\alpha} = -1.28$

3.ANOVA – ANALYSIS OF VARIANCE

ANOVA ONE WAY CLAASIFICATION PROBLEM

ONE WAY ANOVA TABLE

Source of Variation	Sum of Squares	Degree of Freedom	Mean Square	F-value / Test Statistic
Between Samples	SSTR	$r - 1$	$MSTR = SSTR / (r - 1)$	$F = MSTR / MSE$
Within Samples	SSE	$n - r$	$MSE = SSE / (n - r)$	
Total	SST	$n - 1$		

Problem:

A study was conducted to investigate the perception of corporate ethical values among individuals specializing in marketing. Using $\alpha = 0.05$ and the following data (higher scores indicate higher ethical values), test for significant difference in perception among three groups.

Marketing Manager	Marketing Research	Advertisement
6	5	6
5	5	7
4	4	6
5	4	5
6	5	6
4	4	6

Solution:

Let us take the null hypothesis that there is **no significant difference** in ethical values among individuals specializing in marketing. Let us take H_1 : There **is** significant difference in ethical values among individuals specializing in marketing.

Data Table

x_1	x_1^2	x_2	x_2^2	x_3	x_3^2
6	36	5	25	6	36
5	25	5	25	7	49
4	16	4	16	6	36
5	25	4	16	5	25
6	36	5	25	6	36
4	16	4	16	6	36

At the bottom:

$$x_1 = 30$$

$$\Sigma x_1^2 = 154$$

$$x_2 = 27$$

$$\Sigma x_2^2 = 123$$

$$x_3 = 36$$

$$\Sigma x_3^2 = 218$$

T = Sum of all observations in all samples

$$T = 30 + 27 + 36 = 93$$

Thus, T = 93

Correction Factor (C.F.)

$$C.F. = \frac{T^2}{N} = \frac{93^2}{18} = 480.5$$

SST = Total Sum of Squares

$$SST = \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 - C.F.$$

$$SST = 154 + 123 + 218 - 480.5$$

$$= 495 - 480.5$$

$$\boxed{SST = 14.5}$$

SSTR – Sum of Squares Between the Samples

$$SSTR = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} - C.F.$$

$$= \frac{(30)^2}{6} + \frac{(27)^2}{6} + \frac{(36)^2}{6} - 480.5$$

$$= \frac{900}{6} + \frac{729}{6} + \frac{1296}{6} - 480.5$$

$$= 487.5 - 480.5$$

$$\boxed{SSTR = 7}$$

SSE

$$SSE = SST - SSTR$$

$$SSE = 14.5 - 7$$

$$\boxed{SSE = 7.5}$$

Degrees of Freedom

$$df_1 = r - 1 = 3 - 1 = 2$$

$$df_2 = n - r = 18 - 3 = 15$$

$$df_3 = n - 1 = 18 - 1 = 17$$

ANOVA Table Substitution

Source of Variance	S.S	D.F	M.S.S	F-value
Between Samples	7	2	3.5	7
Within Samples	7.5	15	0.5	
Total	14.5	17		

Decision

$$F_{\text{cal}}(7) > F_{\text{tab}}(3.68)$$

$$\Rightarrow H_0 \downarrow \text{rejected}$$

4.ANOVA – ANALYSIS OF VARIANCE
ANOVA TWO WAY CLASSIFICATION PROBLEM

ANOVA – Two Way Classification

Table

Source of Variation	Sum of Square	Degree of Freedom	Mean Square	F-value
Between columns	SSTR	$C - 1$	$MSTR = SSTR / (C - 1)$	$F = MSTR / MSE$
Between rows	SSR	$r - 1$	$MSR = SSR / (r - 1)$	$F = MSR / MSE$
Residual Error	SSE	$(C - 1)(r - 1)$	$MSE = SSE / (C - 1)(r - 1)$	
Total	SST	$Cr - 1$		

Problem:

To study the performance of three detergents and three different water temperatures, the following whiteness readings were obtained with specially designed equipment.

Data Table

Water Temperature	A	B	C
Cold water	57	55	67
Warm water	49	52	88
Hot water	54	46	58

Perform a two-way analysis of variance using 5% level of significance.

Solution

Let us take the null hypothesis that there is no significant difference in the performance of three detergents due to water temperature, and vice versa.

The data coded by subtracting 50 from each observation.

Water temperature table (coded values)

Water Temp	x_1	x_1^2	x_2	x_2^2	x_3	x_3^2	G.T.
Cold water	7	49	5	25	17	289	29
Warm water	-1	1	2	4	18	324	19
Hot water	4	16	-4	16	8	64	8

Below the table:

$$\Sigma x_1 = 10$$

$$\Sigma x_1^2 = 66$$

$$\Sigma x_2 = 3$$

$$\Sigma x_2^2 = 45$$

$$\Sigma x_3 = 43$$

$$\Sigma x_3^2 = 667$$

$$G.T. = 56$$

$$T = 56$$

C.F. (Correction Factor)

$$C.F = \frac{56^2}{9} = \frac{3136}{9} = 348.44444$$

SSTR (Columns/Detergent Effect)

$$\begin{aligned}
 SSTR &= \frac{(10)^2}{3} + \frac{(3)^2}{3} + \frac{(43)^2}{3} - 348.44444 \\
 &= \frac{100}{3} + \frac{9}{3} + \frac{1849}{3} - 348.44444 \\
 &= 33.3333 + 3 + 616.3333 - 348.44444 \\
 &= 652.6666 - 348.44444 \\
 \boxed{SSTR} &= 304.2222
 \end{aligned}$$

SSR (Rows/Temperature Effect)

$$\begin{aligned} SSR &= \frac{(29)^2}{3} + \frac{(19)^2}{3} + \frac{(8)^2}{3} - 348.44444 \\ &= 280.3333 + 120.3333 + 21.3333 - 348.44444 \\ \boxed{SSR} &= \boxed{73.5555} \end{aligned}$$

SST (Total Sum of Squares)

$$\begin{aligned} SST &= 66 + 45 + 667 - 348.44444 \\ &= 788 - 348.44444 \\ &= 439.5556 \end{aligned}$$

SSE (Error Sum of Squares)

$$\begin{aligned} SSE &= SST - (SSTR + SSR) \\ &= 439.5556 - (304.2222 + 73.5555) \\ &= 439.5556 - 377.7777 \\ \boxed{SSE} &= \boxed{61.7779} \end{aligned}$$



DEGREES OF FREEDOM

$$df1 = 3 - 1 = 2$$

$$df1 = 3 - 1 = 2$$

$$df3 = (2)(2) = 4$$

$$df1 = 8$$

ANOVA Table

Source of variation	SS	df	MSS	F-value
B/W column	304.2222	2	$304.2222 \div 2 = 152.1111$	$152.1111 \div 15.44464 = 9.8489$
B/W rows	73.5555	2	$73.5555 \div 2 = 36.7777$	$36.7777 \div 15.44464 = 2.3812$
Error	61.7777	4	$61.7777 \div 4 = 15.44464$	
Total	439.5556	8		

Conclusions

1. The calculated value of F for columns

$$F_{cal} (9.8489) > F_{tab} (2, 4, 0.05) (6.94)$$

→ H_0 is rejected

2. The calculated value of F for rows

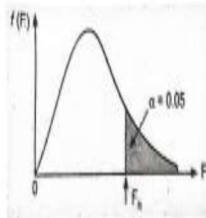
$$F_{cal} (2.3813) < F_{tab} (2, 4, 0.05) (6.94)$$

→ H_0 is accepted

∴ There is no significant change (difference) of water.

TABLE VALUES FOR ANOVA

Table A6: Critical Values of the F-Distribution at a 5 Percent Level of Significance



Degrees of freedom for the denominator	Degrees of freedom for the Numerator													
	1	2	3	4	5	6	7	8	9	10	12	15	20	30
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	250.09
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385	19.396	19.413	19.420	19.446	19.462
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8868	8.8552	8.8123	8.7855	8.7446	8.7029	8.6602	8.6166
4	7.7086	6.9443	6.5914	6.3883	6.2560	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117	5.8578	5.8025	5.7459
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8753	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.4957
6	5.9875	5.1433	4.7571	4.5337	4.3874	4.2839	4.2066	4.1468	4.0990	4.0600	3.9999	3.9381	3.8742	3.8082
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747	3.5108	3.4445	3.3758
8	5.3177	4.4590	4.0662	3.8378	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2840	3.2184	3.1503	3.0794
9	5.1174	4.2565	3.8626	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.8687
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130	2.8450	2.7740	2.6996
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962	2.8536	2.7876	2.7186	2.6464	2.5705
12	4.7272	3.8653	3.4703	3.2392	3.0859	2.9761	2.8934	2.8286	2.7764	2.7334	2.6666	2.5969	2.5236	2.4463
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.6710	2.6037	2.5331	2.4579	2.3803
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6021	2.5342	2.4630	2.3879	2.3082
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4035	2.3275	2.2468
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4247	2.3522	2.2756	2.1938
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943	2.4499	2.3807	2.3077	2.2304	2.1477
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3421	2.2686	2.1906	2.1071
19	4.3808	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.3080	2.2341	2.1555	2.0712
20	4.3513	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0391
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3661	2.3210	2.2504	2.1757	2.0960	2.0102
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.2967	2.2258	2.1508	2.0707	1.9842
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201	2.2747	2.2036	2.1282	2.0476	1.9605
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.1834	2.1077	2.0267	1.9390
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821	2.2365	2.1649	2.0889	2.0075	1.9192
26	4.2252	3.3690	2.9751	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2197	2.1479	2.0716	1.9898	1.9010
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2043	2.1323	2.0558	1.9736	1.8842
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360	2.1900	2.1179	2.0411	1.9586	1.8687
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2782	2.2229	2.1768	2.1045	2.0275	1.9446	1.8543
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921	2.0148	1.9317	1.8409
31	4.0848	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240	2.0772	2.0035	1.9245	1.8389	1.7444
32	4.0012	3.1504	2.7581	2.5252	2.3688	2.2540	2.1665	2.0970	2.0401	1.9926	1.9174	1.8364	1.7480	1.6491
33	3.9201	3.0718	2.6802	2.4472	2.2900	2.1750	2.0867	2.0164	1.9588	1.9105	1.8337	1.7505	1.6587	1.5543
34	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7522	1.6664	1.5705	1.4591

II. NON- PARAAMETRIC TESTS

I.MANN-WHITNEY U- TEST PROBLEM

PROBLEM 1: It is generally believed that as people grow older, they find it harder to get to sleep. To test if there was a difference in time in minutes before people actually went to sleep lying on the bed, a sample of **10 young persons (ages 21–25)** and **10 old persons (ages 65–70)** was randomly selected and their sleeping habits were monitored.

The data shows the number of minutes these 20 persons were awake in bed before getting to sleep:

Young men

10, 12, 20, 16, 25, 20, 30, 40, 47, 28

Old men

98, 62, 148, 70, 140, 95, 68, 90, 112, 156

Is there evidence that young men are significantly more able than old men to get to sleep than old men? Use $\alpha = 0.05$ level of significance.

Solution

Step 1: Arrange the data in ascending order and assign ranks

Young men:

10, 12, 16, 20, 20, 25, 28, 30, 40, 47

Old men:

62, 68, 70, 90, 95, 98, 112, 140, 148, 156

Then assign ranks to each observation:

Young men: 1, 2, 3, 4.5, 4.5, 6, 7, 8, 9, 10

Old men: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20

Sum of ranks:

$R_1=64.5, R_2=148.5, R_1 = 64.5$

Step 2: Compute U-statistics

$$U_1 = n_1n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_1 = 100 + 55 - 64.5 = 90.5$$

$$U_2 = n_1n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

$$U_2 = 100 + 55 - 148.5 = 6.5$$

Lower value = U = 6.5

Step 3: Compute mean and standard deviation of U

$$\mu_U = \frac{n_1n_2}{2} = \frac{10 \times 10}{2} = 50$$

$$\sigma_U = \sqrt{\frac{n_1n_2(n_1 + n_2 + 1)}{12}} = \sqrt{\frac{100 \times 21}{12}} = \sqrt{175} = 13.23$$

Step 4: Compute z-value

$$z = \frac{U - \mu_U}{\sigma_U} = \frac{6.5 - 50}{13.23} = -3.287$$

Since



$$z_{cal} = -3.287 < z_{critical} = -1.96$$

at $\alpha = 0.05$, **Reject the null hypothesis.**

Conclusion:

There is a **significant difference** in time to get to sleep between young men and old men.

PROBLEM 2: Twenty-four applicants are interviewed for a position by three administrators and rated on a scale of 5 as to suitability for the position. Each applicant is given a “suitability” score which is the sum of the three. Although college education is not a requirement for the position, an HR director felt that it might have some bearing on suitability for the position. Raters made their ratings on the basis of individual interviews and were not told the educational background of the applicant. Twelve of the applicants had completed the process and have differences in the scores of the two groups.

Group A : 7 11 9 4 8 6 12 11 9 10 11 11
Group B : 8 9 13 14 11 10 12 14 13 9 10 8

Use a 0.05 level of significance. Group A had an educational background of less than two years of college, while group B had completed at least two years of college.

Solution: Ranking the given figures in group A and B as follows:

Scores		Ranks	Aggregate Ranks	Ranks	
A	B			A = (R ₁)	B = (R ₁)
4		1	1	1	—
6		2	2	2	—
7		3	3	3	—
8	8, 8	4, 5, 6	5	5	10
9, 9	9, 9	7, 8, 9, 10	8.5	17	17
10	10, 10	11, 12, 13	12	12	24
11, 11, 11, 11	11	14, 15, 16, 17, 18	15.5	46.5	15.5
12	12	19, 20	18.5	18.5	18.5
	13, 13	21, 22	20.5		41
	14, 14	23, 24	22.5		45
				105	171

Sum of the two ranks = $n(n + 1)/2 = 23 \times 24/2 = 276$, where n is the total number of observations. This value justifies internal data consistency. Computing the u -test statistic as follows:

$$U_1 = n_1 n_2 + \frac{n_1 (n_1 + 1)}{2} - R_1$$

$$= 11 \times 12 + \frac{11 (11 + 1)}{2} - 105 = 132 + 66 - 105 = 93$$

$$U_2 = n_1 n_2 + \frac{n_2 (n_2 + 1)}{2} - R_2$$

$$= 11 \times 12 + \frac{12 (12 + 1)}{2} - 171 = 152 + 78 - 171 = 39.$$

Consider the lower value between U_1 and U_2 , i.e., $U_2 = 39$ so that

$$\text{Mean, } \mu_U = \frac{n_1 n_2}{2} = \frac{11 \times 12}{2} = 66 \text{ and}$$

$$\text{Standard deviation, } \sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} = \sqrt{\frac{11 \times 12 (11 + 12 + 1)}{12}}$$

$$= \sqrt{\frac{132 \times 24}{12}} = 16.248.$$

Since both n_1 and n_2 are greater than 8, applying z -test statistic, we get

$$z = \frac{U - \mu_U}{\sigma_U} = \frac{39 - 66}{16.248} = -1.661.$$

Since the calculated value $z_{\text{cal}} = -1.661$ is less than its critical value, $z_{\alpha} = -1.96$ at 0.05 level of significance, the null hypothesis is accepted. Hence, there is no significant difference in the scores to the two groups.

2.KRUSKAL-WALLIS TEST

PROBLEM:

Use Kruskal–Wallis test to determine whether there is a significant difference in the following populations. Use $\alpha = 0.05$ level of significance.

Population 1: 17, 19, 20, 27, 35, 40
Population 2: 28, 36, 33, 12, 36
Population 3: 37, 30, 39, 42, 28, 25, 31

Solution

For three populations, $k = 3$ and $n = 18$. The observations in the three populations are combined and ranked. The smallest value is given rank 1.

POPULATION 1		POPULATION 2		POPULATION 3	
VALUE	RANK	VALUE	RANK	VALUE	RANK
17	1	22	4	25	5
19	2	27	6.5	28	8.5
20	3	28	8.5	30	10
27	6.5	33	12	31	11
35	13	36	14	37	15
40	17	$n_2=5$	$t_2=45$	39	16
$n_1=6$	$t_1=42.5$			42	18
				$n_3=7$	$t_3=83.5$

Hypotheses

H_0 : Three populations are identical ($\mu_1=\mu_2=\mu_3$)

H_1 : $\mu_1 \neq \mu_2 \neq \mu_3$

Kruskal–Wallis Test Statistic

$$H = \frac{12}{n(n+1)} \left[\sum_{i=1}^k \frac{t_i^2}{n_i} \right] - 3(n+1)$$

Substituting values:

$$\begin{aligned} H &= \frac{12}{18(19)} \left[\frac{(42.5)^2}{6} + \frac{45^2}{5} + \frac{83.5^2}{7} \right] - 3(19) \\ &= \frac{12}{342} (301.4 + 405 + 996.031) - 57 \\ &= 0.0350877 \times 1702.431 - 57 \\ &= 59.98 - 57 = 2.572 \end{aligned}$$

Correction Factor (C)

$$C = 1 - \frac{\sum (t_j^3 - t_j)}{n^3 - n}$$

$$C = 1 - \frac{[(2^3 - 2) + (2^3 - 2)]}{(18^3 - 18)}$$

$$C = 1 - \frac{12 - 2}{5832 - 18}$$

$$C = 1 - \frac{10}{5814} = 0.998$$

Adjusted H Statistic

$$H' = \frac{H}{C} = \frac{2.572}{0.998} = 2.577$$

Decision

Degrees of freedom:

$$df = k - 1 = 2$$

Critical chi-square value at $\alpha = 0.05$, $df = 2$:

$$\chi_{0.05,2}^2 = 5.991$$

Since:

$$H' = 2.577 < 5.991$$

Fail to reject H_0 .

Conclusion

There is **no significant difference** among the three populations.

Hence, the populations may be considered **identical** at the 5% level of significance.

3. SIGN TEST

PROBLEM 1: The median age of tourists who come to India is claimed to be **40 years**. The ages of tourists in a random sample of **18 tourists** were recorded as follows: **24, 18, 37, 51, 56, 38, 45, 45, 29, 48, 39, 26, 38, 43, 62, 30, 66, 41** Test the hypothesis using $\alpha = 0.05$ level of significance.

Solution

Let us take the null hypothesis that the median age of tourists who come to India is 40 years:

$H_0: \mu = 40$ and $H_1: \mu \neq 40$

Arranging data in ascending order and comparing with the median age of 40:

Tourist Age	18	24	26	29	30	37	38	38	39	41	43	45	45	48	51	56	62	66
Median Age	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40
Sign	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-

SOLUTION:

Given $n = 18$, number of plus (+) signs $x = 9$, and $p = 0.5$.

Thus,

$$\mu = np = 18(0.5) = 9,$$

and

$$\sigma = \sqrt{npq} = \sqrt{18 \times 0.5 \times 0.5} = \sqrt{4.5} = 1.060.$$

Applying the z-test statistic, we get:

$$z = \frac{\bar{x} - \mu}{\sigma} = \frac{8.5 - 9}{1.060} = -1.533$$

where $\bar{x} = 9 \approx 8.5$ because of approximation from discrete distribution to normal distribution.

Since $z_{\text{cal}} = -1.533$ is **greater than** its critical value $z_{\alpha} = -1.96$ at $\alpha = 0.05$ level of significance, **null hypothesis H_0 is accepted**.

Hence, it may be concluded that the **median age of tourists who come to India is 40 years**.

4. CHI-SQUARE TEST

ER 11

11.5.2 Chi-Square Test for Goodness-of-fit

On several occasions, a decision maker needs to understand whether an actual sample distribution matches or coincides with a known theoretical probability distribution such as binomial, Poisson, normal, and so on. The χ^2 test for *goodness-of-fit* determines how closely given data support an assumption about the distribution of a population or random variable of interest. To apply this test, a particular theoretical probability distribution is first hypothesized for the population of interest and then the test is carried out to determine how closely sample data obtained from the population of interest match with the hypothesized theoretical distribution. The observed frequencies come from the sample and the expected frequencies come from the theoretical hypothesized probability distribution. The goodness-of-fit test determines the difference between the observed values and the expected values. A significant difference between sampling distribution and theoretical distribution points towards doubt on the assumption that the hypothesized theoretical distribution is correct. But a small difference between sampling distribution and theoretical distribution is assumed to be due to sampling error.

Procedure

The general steps to conduct a goodness-of-fit test for any hypothesized population distribution are summarized as follows:

1. State the null and alternative hypotheses.
 H_0 : No difference between the observed and expected sets of frequencies.
 H_1 : There is a difference.
2. Select a random sample and record the observed frequencies (O values) for each category.
3. Calculate expected frequencies (E values) in each category by multiplying the category probability by the sample size.
4. Compute the value of χ^2 test statistic

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

5. Find the critical (table) value of χ^2 (See Appendix) at certain level of significance α and $df = n - 1$ provided number of expected frequencies are 5 or more for all categories.
6. Compare the calculated and table value of χ^2 , and use the following decision rule:
 - If χ^2_{cal} is less than its critical value $\chi^2_{\alpha, n-1}$, then accept null hypothesis, H_0 .
 - Otherwise reject H_0 .

Example 11.7: An HR manager is interested to determine whether absenteeism is greater on one day of the week than on another. His records for the past year show the following sample distribution:

Day of the week :	Monday	Tuesday	Wednesday	Thursday	Friday
No. of absentees :	66	56	54	48	75

Test whether the absence is uniformly distributed over the week.

[Madras Univ., M.Com., 2004]

Solution: Let us take the null hypothesis that the absence is uniformly distributed over the week.

If absenteeism is equally probable on all days and number of absentees during a week are 300, then $300/5 = 60$ absentees are expected on each day of the week. Arranging the data as follows:

CHI-SQUARE AND OTHER NON-PARAMETR

Category	O	E	O - E	(O - E) ²	(O - E) ² /E
Monday	66	60	6	36	0.60
Tuesday	57	60	-3	9	0.15
Wednesday	54	60	-6	36	0.60
Thursday	48	60	-12	144	2.40
Friday	75	60	-15	225	3.75
					7.50

Since calculated value $\chi_{cal} = 7.50$ is less than its critical value, $\chi^2 = 9.49$ at $\alpha = 0.05$ and $df = 5 - 1 = 4$, the null hypothesis is accepted.

Example 11.8: A survey of 800 families with 4 children each revealed following distribution:

No. of boys	:	0	1	2	3	4
No. of girls	:	4	3	2	1	0
No. of families	:	32	178	290	236	64

Is this result consistent with the hypothesis that male and female births are equally probable?

Solution: Let us take the null hypothesis that male and female births are equally probable. The probability of a male or female is $p = q = 1/2$. Since birth of male and female is mutually exclusive and exhaustive, the expected number of families having different combinations of boys and girls can be calculated using binomial probability distribution as follows:

$$P(x = r) = {}^4C_r p^r q^{4-r} \quad r = 0, 1, \dots, 4$$

$$= {}^4C_r (1/2)^4 \text{ since } p = q = 1/2$$

The calculations for expected frequencies for each combination (category) of boy or girl are as shown below:

Category x	$P(x = r)$	Expected Frequency, $n P(x)$
0	${}^4C_0 (1/2)^4 = 1/16$	$800 \times (1/16) = 50$
1	${}^4C_1 (1/2)^4 = 4/16$	$800 \times (4/16) = 200$
2	${}^4C_2 (1/2)^4 = 6/16$	$800 \times (6/16) = 300$
3	${}^4C_3 (1/2)^4 = 4/16$	$800 \times (4/16) = 200$
4	${}^4C_4 (1/2)^4 = 1/16$	$800 \times (1/16) = 50$

Arranging observed and expected frequencies in the table below:

Category	O	E	O - E	(O - E) ²	(O - E) ² /E
0	32	50	-18	324	6.480
1	178	200	-22	484	2.420
2	290	300	-10	100	0.333
3	236	200	36	1296	6.480
4	64	50	14	196	3.920
					19.633

Since calculated value of χ^2 is greater than its critical value, $\chi^2 = 9.488$ at $\alpha = 0.05$ and $df = 5 - 1 = 4$, the hypothesis is rejected. Hence, we conclude that male and female births do not seem to be equally probable.

11.5.6 Chi-Square Test of Homogeneity

The test of homogeneity is useful in testing a null hypothesis that several populations are homogeneous with respect to a characteristic. For example, we may like to know that the milk supplied by various companies has a particular ingredient in common or not.

This test is different from the test of independence on account of the following reasons:

- The decision maker may like to know whether different samples come from the same population, instead of knowing whether two attributes are independent or not.
- Two or more independent samples are drawn from each population, instead of only one sample.
- This test is similar to the test of hypothesis for difference between two population means or proportions in case characteristics to be compared consist of two categories.

The procedure starts by selecting a random sample from each population, and then in each sample the proportion falling into each category or strata is determined. The sample data so obtained is arranged in a contingency table. The rest of the procedure to test a null is same as discussed for test of independence

Example 11.20: A movie producer is in the process of releasing a new movie. In order to develop an advertising strategy, he wants to determine whether the movie will appeal to people of particular age group or will appeal equally to people of all age groups. The producer takes a random sample from the group of people attending the preview of the movie, and obtains the following results:

Opinion	Age Groups				Total
	Under 20	20-39	40-59	60 and over	
Liked the movie	146	78	48	28	300
Disliked the movie	54	22	42	22	140
Indifferent	20	10	10	20	60
Total	220	110	100	70	500

What inference will you draw from this data?

Solution: Let us take the null hypothesis that the movie is appreciated by the people of all age group.

The calculations for expected frequencies are as follow:

$$E_{11} = \frac{300 \times 220}{500} = 132,$$

$$E_{12} = \frac{300 \times 110}{500} = 66$$

$$E_{13} = \frac{300 \times 100}{500} = 60$$

$$E_{14} = 300 - (132 + 66 + 60)$$

$$E_{21} = \frac{140 \times 220}{500} = 61.6,$$

$$E_{22} = \frac{140 \times 110}{500} = 30.8$$

$$E_{23} = \frac{140 \times 100}{500} = 28$$

$$E_{24} = 140 - (61.6 + 30.8 + 28) = 42$$

$$E_{31} = 220 - (132 + 61.6) = 26.4$$

$$E_{32} = 110 - (66 + 30.8) = 13.2$$

$$E_{33} = 100 - (60 + 28) = 12$$

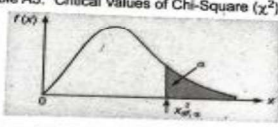
$$E_{34} = 70 - (42 + 19.6) = 8.4 = 19.6$$

O	E	O - E	(O - E) ²	(O - E) ² /E
146	132.0	14.0	196.00	1.485
54	61.6	-7.6	57.76	0.938
20	26.4	-6.4	40.96	1.552
78	66.0	12.0	144.00	2.182
22	30.8	-8.8	77.44	2.514
10	13.2	-3.2	10.24	0.776
48	60.0	-12.0	144.00	2.400
42	28.0	14.0	196.00	7.000
10	12.0	-2.0	4.00	0.333
28	42.0	-14.0	196.00	4.667
22	19.6	2.4	5.76	0.294
20	8.4	11.6	134.56	16.019
				40.16

Since the calculated value of $\chi^2 = 40.16$ is greater than its critical value, $\chi^2 = 12.59$ for $df = (r - 1)(c - 1) = (3 - 1)(4 - 1) = 6$ and $\alpha = 0.05$, the null hypothesis is rejected. Hence, we conclude that the opinion of all age groups is not the same about the new movie.

TABLE VALUES FOR CHI SQUARE TEST

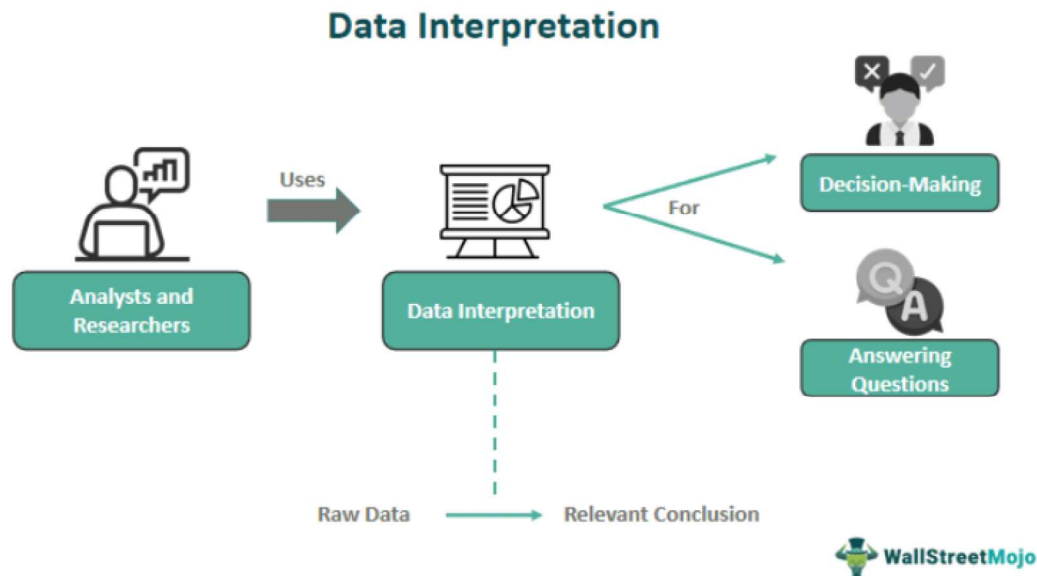
Table A5: Critical Values of Chi-Square (χ^2)



Degree of freedom	0.100	0.050	0.025	0.010	0.005	0.001
1	2.71	3.84	5.02	6.63	7.88	10.8
2	4.61	5.99	7.38	9.21	10.6	13.8
3	6.25	7.81	9.35	11.3	12.8	16.3
4	7.78	9.49	11.1	13.3	14.9	18.5
5	9.24	11.1	12.8	15.1	16.7	20.5
6	10.6	12.6	14.4	16.8	18.5	22.5
7	12.0	14.1	16.0	18.5	20.3	24.3
8	13.4	15.5	17.5	20.1	22.0	26.1
9	14.7	16.9	19.0	21.7	23.6	27.9
10	16.0	18.3	20.5	23.2	25.2	29.6
11	17.3	19.7	21.9	24.7	26.8	31.3
12	18.5	21.0	23.3	26.2	28.3	32.9
13	19.8	22.4	24.7	27.7	29.8	34.5
14	21.1	23.7	26.1	29.1	31.3	36.1
15	22.3	25.0	27.5	30.6	32.8	37.7
16	23.5	26.3	28.8	32.0	34.3	39.3
17	24.8	27.6	30.2	33.4	35.7	40.8
18	26.0	28.9	31.5	34.8	37.2	42.3
19	27.2	30.1	32.9	36.2	38.6	43.8
20	28.4	31.4	34.2	37.6	40.0	45.3
21	29.6	32.7	35.5	38.9	41.4	46.8
22	30.8	33.9	36.8	40.3	42.8	48.3
23	32.0	35.2	38.1	41.6	44.2	49.7
24	33.2	36.4	39.4	43.0	45.6	51.2
25	34.4	37.7	40.6	44.3	46.9	52.6
26	35.6	38.9	41.9	45.6	48.3	54.1
27	36.7	40.1	43.2	47.0	49.6	55.5
28	37.9	41.3	44.5	48.3	51.0	56.9
29	39.1	42.6	45.7	49.6	52.3	58.3
30	40.3	43.8	47.0	50.9	53.7	59.7
35	46.1	49.8	53.2	57.3	60.3	66.6
40	51.8	55.8	59.3	63.7	66.8	73.4
45	57.5	61.7	65.4	70.0	73.2	80.1
50	63.2	67.5	71.4	76.2	79.5	86.7

UNIT-V
INTERPRETATION & REPORT WRITING

INTERPRETATION



Definition: Data interpretation is the process of reviewing data and arriving at relevant conclusions using various analytical methods. It transforms raw numbers into meaningful insights.

Data Interpretation is a process where **analysed data** is used to make conclusions on the **meaning** and **implications** of some particular study and decide on how insights will be applied in a practical environment. It is simply translating **numerical, graphical**, or even text results from data analysis into meaningful stories that can **drive decision-making**. **Interpretation** is necessary to bridge raw data and actionable knowledge because it simply focuses on the meaning of the findings in relation to the problem at hand and what they suggest for future action or strategy.

Key Features of Data Interpretation:

1. **Contextual Understanding:** Requires understanding the context in which data exists; interprets results in terms of what they mean for a specific question or decision.
2. **Decision-Making:** Focuses on translating analyzed data into information that can be used to make decisions.

3. **Subjectivity:** Involves a certain level of subjectivity as different interpreters might conclude differently based on the same data set, depending on their perspective or background.
4. **Critical Thinking:** Requires critical thinking skills to question and consider the limitations of the data, including biases and anomalies.
5. **Narrative Formulation:** Often results in the formulation of a narrative or story that explains what the data shows and why it matters in a given context.

TYPES OF DATA INTERPRETATION (DI)

The various types of Data Interpretation are classified as:

1. Tabular Data

A **table chart** is a structured arrangement of data in **rows and columns**, making it easy to compare and analyze numerical information. It is one of the most common and straightforward formats in Data Interpretation.

Below is an example image to understand Tabular Data

The data below shows A, B, C, D, and E school and their marks in subjects.

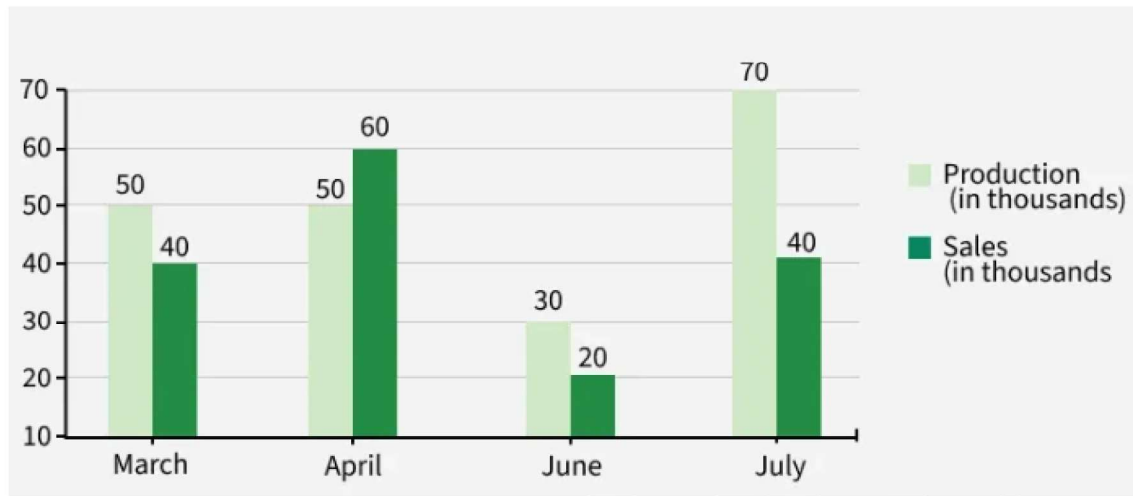
Subject /School	Science	Hindi	English	Maths
A	360	280	590	510
B	310	280	230	470
C	430	350	370	210
D	270	300	340	190
E	290	340	450	280

2. Bar Graphs

A bar chart_(or bar graph) is a visual representation of data using **rectangular bars** of varying lengths or heights. Each bar represents a **category**, and its length corresponds to the **value or frequency** of that category.

Below is an example image to understand about Bar Graphs

The below image shows the production and sale in the month from march to july in terms of a Bar Graph.

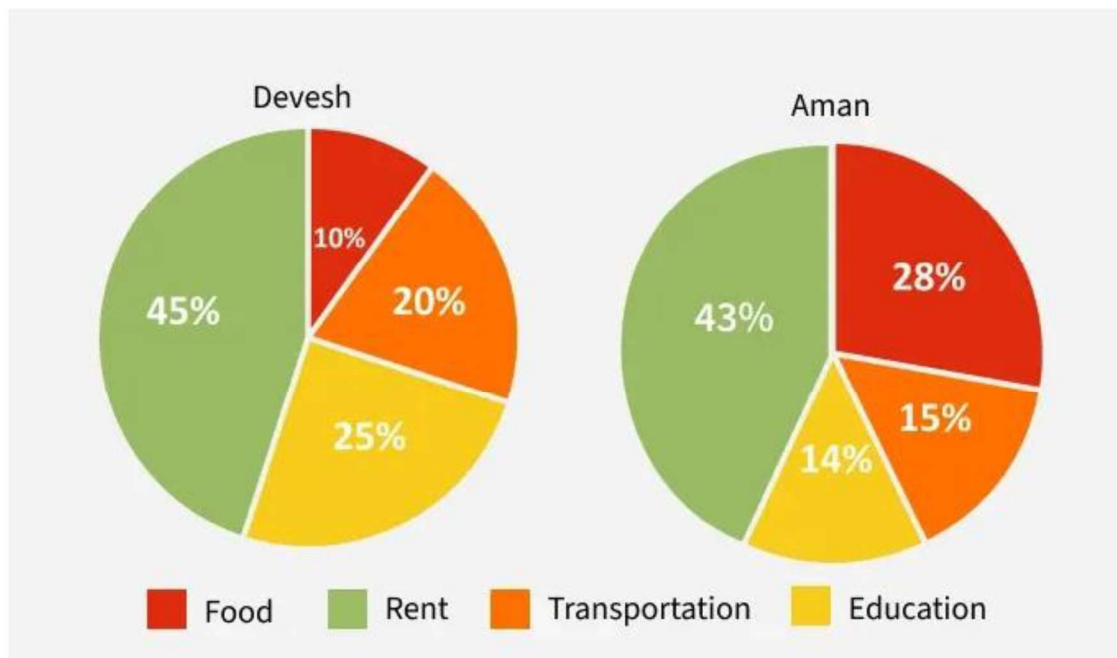


3. Pie Charts

A pie chart is a circular statistical graphic divided into slices to illustrate **proportional distribution** or **percentage composition** of a whole dataset. Each slice represents a category's contribution to the total (100%).

Below is an example image to understand about Pie Charts

The below image shows the expenditure of Devesh and Aman in terms of a Pie chart.

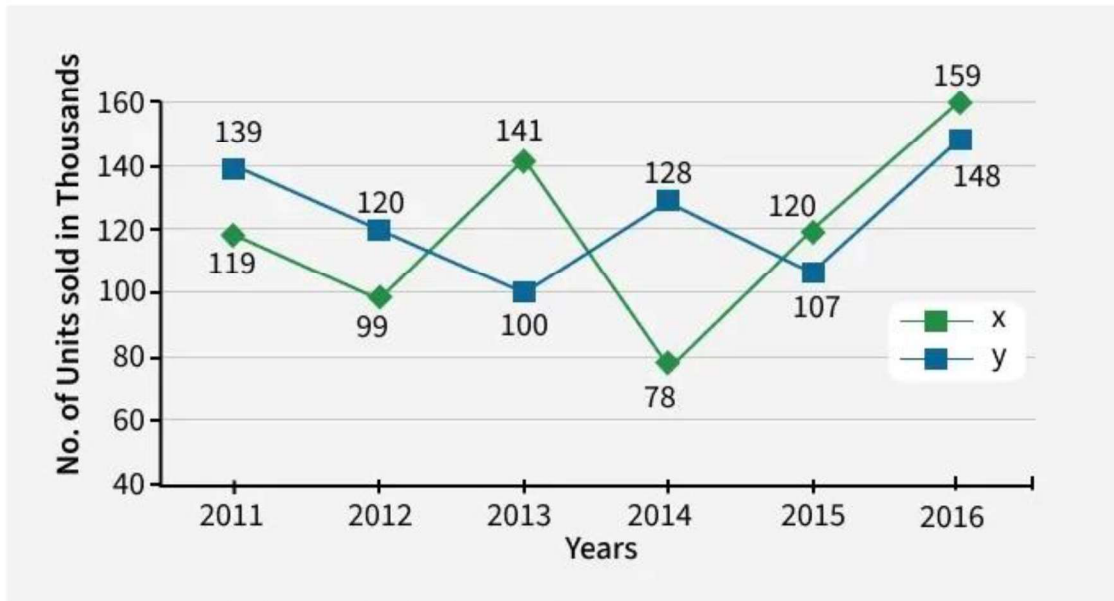


4. Line Graphs

A line chart (or line graph) displays data points connected by straight lines to show **trends over time** or ordered categories. It highlights patterns like growth, decline, or stability in numerical data.

Below is an example image to understand about Line Graphs

The below image shows the No. of unit sold by X and Y in terms of line Graph.

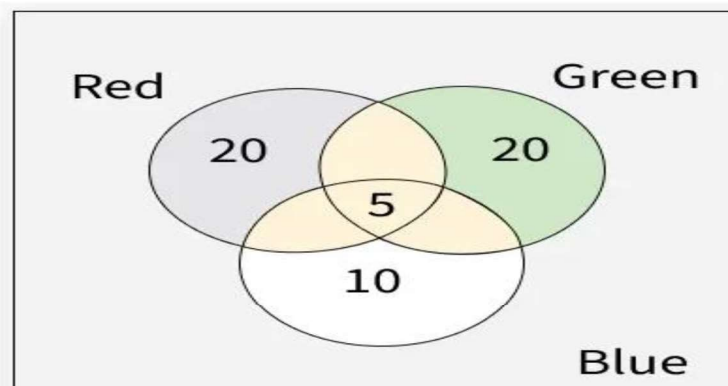


5. Caselet DI

A Caselet DI (or Paragraph-Based DI) presents data in textual form (a short paragraph or story) instead of tables/charts. It requires extracting numerical details, organizing them logically, and solving questions based on the hidden data.

Below is an example image to understand about Caselet

The below image shows the distribution of Red, Green and Blue in terms of Caaselet DI



Process of data interpretation:

1. Baseline Establishment

Just like in a competitive analysis, the first thing you need to do when interpreting data is establish a clear starting point. This means setting specific objectives and mapping out both your short-term and long-term goals that could be influenced by what you find in the data.

For instance, if you're an investor, you might focus on setting targets related to the return on investment (ROI) for the companies you're assessing. Don't forget, at this stage, it's equally crucial to decide which kind of data you'll be diving into for analysis and interpretation.

2. Data Collection

Once you have established your baseline, it's time to collect data. A good way to start is by visualizing it—maybe through a bar chart, graph, or pie chart.

The goal here is to ensure you're analyzing the data clearly and without any bias. At this point, think back to how you conducted your research.

Here are a couple of key questions to ask yourself:

Did anything change or go wrong during the data collection process?

Do you have any notes or observations saved from that time?

Once you've gathered all your data and answered these questions, you're ready to move on to the next step.

3. Data Preparation

After gathering your data, the next step is getting it ready for analysis. This often means cleaning it up by correcting any errors, filling in missing values, or removing outliers. You might also need to reformat the data to make it easier and more effective to analyze.

4. Interpretation

This step is crucial in the data interpretation process—it's where you actually analyze the data you've gathered. This is also where you'll decide whether to take a qualitative or quantitative approach.

If you choose qualitative analysis, you'll be looking at the data from a more subjective angle. When using AI-driven tools for this, you'll need to "code" the data extensively to capture personal sentiments that can't be easily represented by numbers.

On the other hand, quantitative analysis involves a more objective, numbers-based approach. You'll work with raw numerical data, calculating things like mean, standard deviation, and ratios. From there, you can use statistical models to dive deeper into behaviors, helping to predict trends and outcomes.

5. Visualization

Once your analysis is finished, it's time to visualize your data and extract insights from different angles. Many companies today use dashboards for this step, offering quick, real-time insights through algorithms.

However, even without a dashboard, visualizing data is fairly simple. You'll just need to format your data in a way that makes it easy to present visually.

Some common formats you can use include:

Bar charts

Tables

Scatter plots

Line graphs

Pie charts

Histograms

ViewMetrics provides an impressive array of customizable charts and graphs, allowing you to effortlessly create the ideal visual for your data. With its interactive dashboards, you can dive into your data from various perspectives and discover insights you might have otherwise overlooked.

6. Reflection

Finally, after you've crafted the right visualizations that align with your initial goals, it's time to take a step back and reflect. While this might feel like a simpler task compared to the earlier stages, it's actually a critical moment that can shape how effectively you interpret your data.

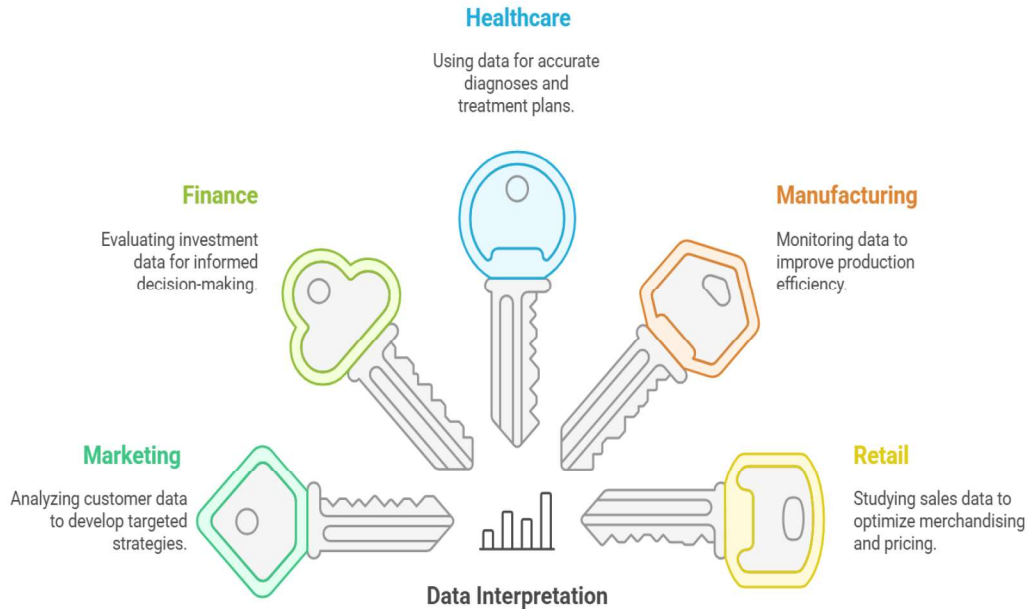
Use this time to review the entire data analysis process, search for any hidden relationships, and spot outliers or mistakes that might have slipped through during data cleaning. It's important here to distinguish between correlation and causation, recognize any biases, and catch insights you may have overlooked.

How Can Data Interpretation Be Used in Different Industries?

There are several use cases for data interpretation in different industries and settings, such as marketing, healthcare, financial services, retail, and manufacturing.

Let's see how data interpretation can be used in these different industries:

Data Interpretation Across Diverse Industries



Marketing – Marketers analyze customer behaviors, preferences, and trends through data to develop targeted strategies and campaigns.

Finance – Financial companies evaluate investment performance through data, helping them make informed decisions on investment strategies.

Healthcare – Healthcare professionals assess patient histories and test results using data, aiding in accurate diagnoses and treatment plans.

Manufacturing – Manufacturers use data to monitor machine performance and inventory levels, improving production efficiency and management.

Retail – Retailers study sales figures and customer habits to optimize merchandising and pricing strategies.

KEY METHODS OF DATA INTERPRETATION

Data interpretation is not a one-size-fits-all process. The method you choose depends on the type of data you have. The two primary categories of data are quantitative and qualitative.

Aspect	Quantitative Interpretation	Qualitative Interpretation
--------	-----------------------------	----------------------------

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Data Type	Numerical data (e.g., sales figures, click-through rates, temperatures).	Non-numerical, descriptive data (e.g., interview transcripts, open-ended survey responses, observations).
Purpose	To measure, test hypotheses, and answer "what" or "how many."	To explore ideas, understand context, and answer "why."
Analysis Method	Statistical analysis, mathematical modeling. Uses tools like mean, standard deviation, regression.	Thematic analysis, content analysis, narrative analysis. Involves identifying themes and patterns in text or images.
Sample Size	Typically involves a large, representative sample to ensure statistical significance.	Often uses a smaller, targeted sample to gain in-depth understanding.
Results	Provides objective, generalizable conclusions presented with numbers, charts, and graphs.	Provides subjective, context-rich insights presented with quotes, stories, and descriptions.
Example Question	What percentage of customers churned last quarter?	Why did customers who churned decide to leave?

1. Quantitative Data Interpretation:

Quantitative data interpretation focuses on statistical analysis of numbers. It's about finding the hard facts in your data. Here are some common methods used.

i. Mean, Median, and Mode

These are measures of central tendency that provide a summary of your dataset.

- **Mean:** The average of all numbers in a set. It's useful for getting a general idea but can be skewed by outliers.
- **Median:** The middle value in a sorted dataset. It is less affected by extreme values and often gives a better sense of the "typical" value.
- **Mode:** The most frequently occurring value in a set. It is useful for categorical data, like identifying the most popular product.

ii. **Standard Deviation and Variance**

These methods measure the dispersion or spread of data points around the mean. A low standard deviation indicates that the data points tend to be close to the mean. A high standard deviation indicates that the data points are spread out over a wider range. This is crucial for understanding the consistency and reliability of your data.

iii. **Regression Analysis**

Regression analysis is used to understand the relationship between a dependent variable and one or more independent variables. For example, you could use it to determine if there is a correlation between advertising spend (independent variable) and monthly sales (dependent variable). This helps in making predictions and forecasting future outcomes.

iv. **Cohort Analysis**

This method involves breaking down users into groups (cohorts) who share common characteristics over time. For example, you could analyze the behavior of all customers who signed up in January versus those who signed up in February. This helps you understand how user behavior changes over time and the long-term impact of your strategies.

2. Qualitative Data Interpretation:

Qualitative interpretation is about understanding the human element behind the numbers. It seeks to uncover motivations, opinions, and experiences through non-numerical data.

i. **Thematic Analysis**

This is one of the most common qualitative methods. It involves carefully reading through textual data (like customer reviews) and identifying recurring themes or ideas.

For example, you might find that many negative reviews mention "slow customer support" or "confusing user interface." These themes point directly to areas for improvement.

ii. **Narrative Analysis**

Narrative analysis focuses on the stories people tell. This could be through testimonials, case studies, or in-depth interviews. By analyzing the structure, language, and content of these stories, researchers can gain a deep understanding of customer experiences and perspectives. It's a powerful tool for developing marketing messages that resonate with your audience.

iii. Content Analysis

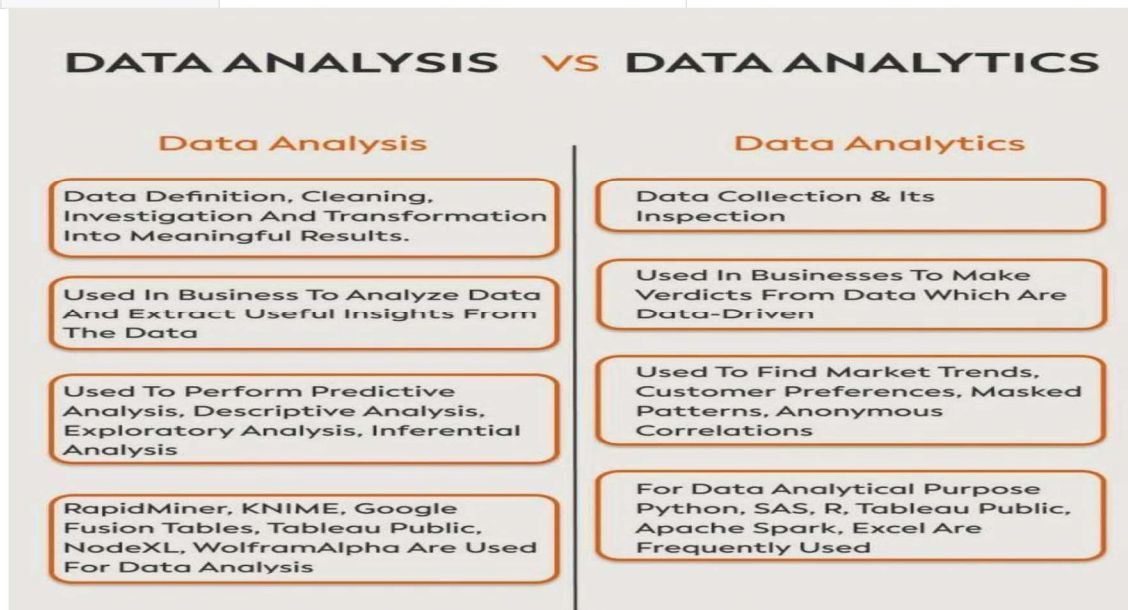
Content analysis is used to quantify the presence of certain words, concepts, or themes within texts or images. For example, a marketing team could analyze social media comments to count the number of times positive words (e.g., "love," "amazing") versus negative words (e.g., "disappointed," "frustrated") are mentioned in relation to a new product launch.

- **Action:** Shift budget from Campaign C to Campaign A. Develop new campaigns with similar messaging and creative, specifically targeting the 18-24 age group.

Key Differences between Analysis and Interpretation

Aspect	Data Analysis	Data Interpretation
Objective	Process and organize raw data to uncover patterns or trends.	Make sense of analyzed data, draw conclusions, and provide context.
Process	Involves data collection, cleaning, transformation, and application of analytical techniques.	Involves evaluating and synthesizing results to explain findings and suggest actions.
Focus	Answers "what" and "how" questions about the data.	Answers "why" and "what next" based on the analysis results.
Nature	More technical and quantitative.	More qualitative and subjective.
Outcome	Produces structured data, statistical outputs, and models.	Produces insights, conclusions, and actionable recommendations.
Role in Decision-Making	Provides data and evidence to support decisions.	Directly informs and influences decision-making.

Aspect	Data Analysis	Data Interpretation
Dependency	Can be performed independently but is limited without interpretation.	Relies on the results of data analysis and cannot occur without it.



REPORT WRITING

Report Writing

“ Report writing is the art of communicating information that has been acquired through research or investigation in a formal, structured manner. It involves synthesizing information, drawing conclusions, and making recommendations based on the findings. ” - American Management Association



Identify



Research



Organise

GEEKTONIGHT.com

Report writing is a structured method of documenting and presenting information in a clear and concise manner. Reports are used for various purposes, including presenting research findings, providing information about a particular topic, or summarizing data and statistics.

Report Writing is the process of presenting information in a structured and organised way. It serves as a means of communicating facts, findings, or recommendations to a specific audience, typically in a written format. The importance of report writing lies in various fields, including academics, business, science, and government, to convey important details and insights.

A Report typically starts with a clear purpose or objective. The Writer gathers relevant information through research, observation, or data collection. This data is then analysed and organised into a coherent document. Reports can vary in length, complexity, and style, depending on the intended audience and purpose.

One of the key aspects of Report Writing is its structure. A typical Report consists of sections such as an introduction, methodology, findings or results, discussion, and a conclusion. These sections help readers understand the context, the process of gathering information, the outcomes, and the significance of the findings.

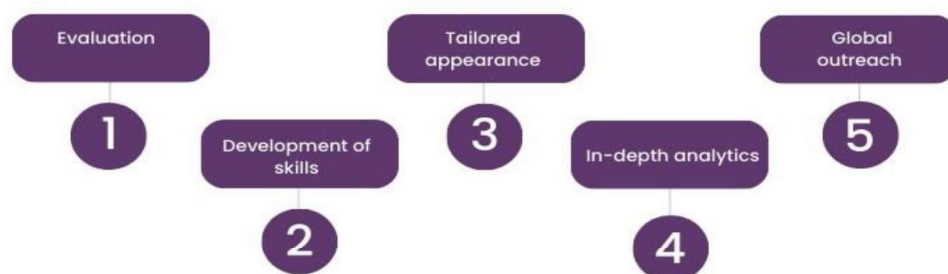
Reports often include visual aids like charts, graphs, and tables to make complex data more accessible. Additionally, citing sources is essential to provide credibility and allow readers to verify the information.

SIGNIFICANCE OF REPORT WRITING

Report Writing is a valuable skill that offers numerous advantages across different situations. Here are some key Features of Report Writing that highlight its significance:

*the*knowledgeacademy

Importance of Report Writing



Evaluation

The importance of Report Writing can be reflected during the evaluation process. This is because Report Writing can help you evaluate your own or others' performance, progress, or outcomes. For example, if you are a student, you can write a Report to assess your learning outcomes, achievements, or challenges in a course or a project. As a teacher, you can write a Report to evaluate your students' learning outcomes, strengths, or weaknesses in a course or an assignment.

If you are a researcher, you can write a Report to evaluate your research methods, results, or implications in a study or an experiment. Report Writing can help you identify the gaps, strengths, or areas for improvement in your own or others' work, following a clear Report Writing Format.

Development of skills

Report Writing can help you develop your skills in various domains and disciplines. For example, if you are a student, you can write a Report to improve your writing, research, analysis, or presentation skills. If you are a teacher, you can write a Report to enhance your teaching, assessment, feedback, or curriculum design skills.

If you are a researcher, you can write a Report to advance your knowledge, Research Methodology, innovation, or contribution skills. Report Writing can help you acquire new knowledge, apply existing knowledge, or create new knowledge in your field of interest.

Investigation

The Importance of Report Writing also lies in investigating a problem or a topic in-depth and in detail. For example, if you are a student, you can write a Report to explore a question or an issue that interests you or relates to your course or project.

At the same time, if you are a teacher, you can write a Report to investigate a pedagogical or educational problem or phenomenon that affects your teaching or learning environment. On the other hand, if you are a researcher, you can write a Report to investigate a scientific or social problem or phenomenon that has significance or relevance for your discipline or society. Report Writing can help you collect, analyse, and present data in an organised and systematic way.

Quick location

Report Writing can help you locate information quickly and easily. For example, students can write a Report to summarise the main points and findings of your course or project for future

reference or revision. If you are a teacher, you can write a Report to document the key aspects and outcomes of your course or assignment for future use or evaluation.

At the same time, researchers can write a Report to record the essential details and implications of a study or experiment for future dissemination or publication. Report Writing can help you create an index, an abstract, or an executive summary that can help you access information at a glance.

Decision making tool

Report Writing acts as a decision-making tool that can assist you in making decisions based on facts and evidence. For example, if you are a student, you can write a Report to support your arguments or opinions with data and sources in an essay or a debate. If you are a teacher, you can write a Report to justify your decisions or recommendations with data and sources in an assessment or feedback.

If you are a researcher, you can write a Report to validate your claims or hypotheses with data and sources in a study or an experiment. Thus, Report Writing can help you use logic, reasoning, and analysis to make informed and rational decisions.

Neutral presentation of facts

Report Writing can help you present facts in a neutral and objective manner. For example, if you are a student, you can write a Report to avoid bias or emotion in your writing and use facts and evidence to support your points. If you are a teacher, you can write a Report to avoid bias or emotion in your assessment and use facts and evidence to evaluate your students.

Researchers can write a Report to avoid bias or emotion in their research and use facts and evidence to demonstrate their findings. Report Writing can help you maintain a professional and ethical tone in your communication.

A managerial tool

Report Writing can help you manage your work or project effectively and efficiently. For example, if you are a student, you can write a Report to plan, organise, and monitor your progress or outcomes in a course or a project. If you are a teacher, you can write a Report to plan, organise, and monitor your activities or objectives in a course or an assignment.

If you are a researcher, you can write a Report to plan, organise, and monitor your methods or results in a study or an experiment. As a result, Report Writing can help you set goals, allocate resources, and measure performance.

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Unleash your creativity and become a master storyteller with our Creative Writing Training - start your writing journey now!

Proper control

Report Writing helps you control your work or project effectively and efficiently. For example, if you are a student, you can write a Report to check, review, and revise your work or project before submission or presentation. If you are a teacher, you can write a Report to check, review, and revise your work or project before delivery or evaluation.

If you are a researcher, you can write a Report to check, review, and revise your work or project before dissemination or publication. Thus, Report Writing can help you ensure quality, accuracy, and consistency in your work or project.

Professional advancement

The importance of Report Writing lies in advancing and developing your professional career. For example, if you are a student, you can write a Report to demonstrate your competence, knowledge, and skills in a course or a project. In contrast, if you are a teacher, you can write a Report to demonstrate your competence, knowledge, and skills in a course or an assignment.

If you are a researcher, you can write a Report to demonstrate your competence, knowledge, and skills in a study or an experiment. Report Writing can help you showcase your achievements, contributions, or innovations in your field of interest.

Encountering advance and complex situation

Report Writing can help you encounter advanced and complex situations in your work or project. For example, if you are a student, you can write a Report to deal with challenging questions or issues that arise in your course or project. If you are a teacher, you can write a Report to deal with challenging questions or issues that arise in your course or assignment.

If you are a researcher, you can write a Report to deal with challenging questions or issues that arise in your study or experiment. Report Writing can help you solve problems, overcome obstacles, or discover new possibilities in your work or project.

Unlock the power of effective communication with our Report Writing Training - start crafting impactful Reports today!

STEPS IN WRITING REPORT WRITING

Writing a report can be a complex and time-consuming process, but by following a structured approach, you can ensure that your report is clear, concise, and effective. Here are the steps involved in writing a report:

Different Steps in Writing Report:

- 1. Define the Purpose and Scope of the Report**
- 2. Gather Data and Information**
- 3. Analyze the Data**
- 4. Outline the Report Structure**
- 5. Write the Report**
- 6. Format and Present the Report**
- 7. Review and Edit the Report**
- 8. Finalize the Report**

1. **Define the Purpose and Scope of the Report:** Determine the reason for writing the report and what you want to achieve with it. This will help you to focus your research and determine what information to include in the report.
2. **Gather Data and Information:** Collect relevant data and information from various sources, such as books, articles, interviews, and surveys. Ensure that the information you gather is accurate and relevant to the purpose of the report.
3. **Analyze the Data:** Organize and analyze the data to identify trends, patterns, and relationships. This will help you to draw meaningful conclusions and make recommendations.
4. **Outline the Report Structure:** Create an outline of the report's structure, including the main sections, sub-sections, and headings. This will help you to organize the information and ensure that the report is easy to read and understand.
5. **Write the Report:** Start by writing the introduction, which should provide background information and explain the purpose of the report. Then, write the main body of the report, including the results, discussion, and conclusion. Finally, write the executive summary and conclusion, which should summarize the main findings and recommendations of the report.
6. **Format and Present the Report:** Format the report to ensure that it is visually appealing and easy to read. Choose appropriate charts, tables, and graphs to present the data and make sure they are easy to understand.
7. **Review and Edit the Report:** Review the report for grammar, spelling, and punctuation errors. Ensure that the report is well-structured and that the information is presented in a clear and concise manner. Make revisions as needed to improve the overall quality of the report.

8. **Finalize the Report:** After you have reviewed and edited the report, it's time to finalize it. This can include adding a title page, table of contents, references, and appendices, if necessary.

LAYOUT OF REPORT

The layout of a research report is an important aspect of the overall report as it affects readability, clarity, and the impact of the report on its intended audience. Here is a typical layout of a research report:

- **Title page:** The title page should include the title of the report, the author's name, the date of submission, and the name of the organization or institution.
- **Abstract:** The abstract is a brief summary of the report, including its purpose, methodology, results, and conclusions. It should be concise and not exceed 200-300 words.
- **Table of Contents:** This section lists the major sections of the report and their corresponding page numbers.
- **Introduction:** The introduction should provide background information on the topic of the research, explain the purpose of the study, and describe the research questions or hypotheses being investigated.
- **Literature Review:** This section reviews previous studies and research on the topic and provides a context for the current research.
- **Methodology:** The methodology section should describe the research design, sampling method, data collection techniques, and data analysis methods used in the study.
- **Results:** The results section presents the findings of the study, including any data collected, tables, graphs, or charts.
- **Discussion:** The discussion section provides an interpretation of the results, draws conclusions, and discusses the implications of the findings.
- **Conclusion:** The conclusion should summarize the main findings of the study and their implications, and provide recommendations for future research.
- **References:** The references section lists all the sources cited in the report, including books, articles, and other sources of information. The references should be formatted according to the appropriate citation style (e.g., APA, MLA, or Chicago).

- **Appendices:** The appendices section includes any additional material that supports the information in the report, such as raw data, questionnaires, or interview transcripts.

TYPES OF REPORTS

There are various types of report writing, each with its own unique purpose and format. Some of the most common types of reports include:

- i. Business Reports:** These reports provide information and analysis to support decision-making in a business setting. They may include financial reports, marketing reports, operational reports, and strategic reports.
- ii. Academic Reports:** Academic reports are written for educational purposes and often take the form of research reports, lab reports, or case studies. They are used to present the results of academic research and provide a detailed analysis of the findings.
- iii. Technical Reports:** Technical reports are written to document the results of scientific or engineering research. They often include detailed descriptions of methods, procedures, and results, as well as graphs, diagrams, and other visual aids.
- iv. Feasibility Reports:** Feasibility reports are written to assess the viability of a proposed project or initiative. They include an analysis of the costs, benefits, and risks associated with the project, as well as a recommendation as to whether or not it should be pursued.
- v. Progress Reports:** Progress reports are written to provide updates on the progress of a project or initiative. They are often used to communicate the status of the project to stakeholders and provide information on any challenges or obstacles that have been encountered.
- vi. Incident Reports:** Incident reports are written to document events or incidents that have taken place, such as accidents, security breaches, or equipment failures. They are used to record what happened, why it happened, and what steps were taken to address the issue.
- vii. Analytical Reports:** Analytical reports are written to provide a comprehensive analysis of a particular issue or topic. They may include an analysis of data, trends, and patterns, as well as recommendations for future action.

The type of report you write will depend on the purpose of the report. It's important to choose the right format and style to ensure that the report is effective in communicating its message.

Conclusion

Report Writing is a crucial skill that can open doors to various opportunities in your academic and professional life. By understanding What is Report Writing, the types of Reports, Report

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Writing formats, and following effective tips, you can become a proficient Report Writer. Moreover, recognising the differences between project writing, article writing, and Report Writing will help you choose the right approach for your communication needs. Finally, with the help of modern Report writing software, you can streamline the process and create impressive Reports that convey your message effectively.

ORAL PRESENTATION



An important aspect to any research project is the oral presentation of the experiment to other people. As with a research report, you want to tell the story of your experiment: why the experiment was done, how it was done, the results, interpretation of the results, and why the experiment matters.

However, a good presentation is different from a good paper. The presentation should not consist of simply reading from a paper that was previously prepared. Care should be taken to not overwhelm the listener with needless detail. Much more detailed information can be presented and understood in a written paper than in an oral presentation.

The style of a presentation is also important. The presenter must try to keep the listener focused on the key information that is being conveyed.

The following are specific things that should be considered when preparing an oral presentation.

Process of oral presentation:

Oral presentations should be organized to have introduction, body and conclusion sections.

❖ Introduction

This section should be brief. It should provide enough background information so that the listener understands the general hypothesis and why the experiments were done. It should also state the specific research question that was studied.

❖ Body

This section is the major portion of the talk. It should include research methods as well as research results. The methods should be briefly stated, providing detail when necessary for understanding a particular result.

❖ Conclusion

This section should also be brief. A clear, concise statement of what the results prove should be made. The data can be related to experiments others have performed, but this should not be overdone. Future experiments to test unanswered questions could be suggested. State why this experiment matters.

Presentation Style

The following are things that should be considered when designing a presentation.

❖ Time

Pay attention to the time. Most research talks are short and no more than 15 minutes.

❖ Pace

Do not talk too quickly. Slow down so that the listener has time to hear you.

If you think you are speaking too slowly, then you probably are going at the right pace.

❖ Volume/Tone

Talk loud enough so that your listener can hear you. Use a variety of voice inflections and pitches so that the listener stays interested.

- Nothing is more boring than a monotone presentation.
- Alterations in volume/tone gives the listener the feeling that the presenter is interested in the topic.

❖ Eye Contact

Try to maintain eye contact with the listener; this helps them stay focused on the talk.

- Do not simply read your paper! Whether or not you are presenting from notes, a fully prepared script, or from memory, eye contact must be made frequently.
- Face the audience: do not read off of your slides! You are talking to the people so look at them.

❖ Poise

The presentation should be made in a formal, professional manner.

- Dress appropriately.
- Maintain good, erect posture
- Refrain from informal speech patterns and actions.
- Minimize unnecessary movements such as excessive walking, hand motions, etc.
- Keep your hands out of your pockets

❖ Visual Aids

In general, all research presentations need some sort of visual aid. This is most often done using PowerPoint.

- Graphs, tables, photographs etc. of data help the listener sort through the material.
- Complex methods can be presented clearly through visuals.
- A list of conclusion statements helps the listener focus on the final statement.
- A clearly stated research question when visually presented helps.
- Be sure that the visuals are not too complicated. Include only the information you will be discussing.
- Be sure the visual is large enough to be clearly seen by the listener.
- Point to the visuals during the presentation
- Leave the visual up long enough so that the listener can assimilate it.

❖ Present Information Clearly

The information in a presentation should be organized logically and clearly in a way that the listener can understand and follow.

- Use of visuals helps here.
- Details should be included when they are important in reaching a particular conclusion. They should be omitted when they get in the way of seeing a particular point.
- Remember: it is not what you say that is important, it is what the listener hears, understands, and takes with him/her that is important.

❖ Subject Knowledge

The presenter should demonstrate that he/she understands the subject being presented. This is done by:

- presenting accurate information,
- by responding to controversies in an appropriate way,

- by answering reasonable questions from the audience.

EXECUTIVE SUMMARY

In a research paper, the executive summary serves as a condensed version of the entire paper. It provides a snapshot of the key elements and findings of the research, allowing readers to quickly grasp the purpose, methodology, and main outcomes without having to read the entire document.

Why Do We Compose An Executive Summary?

We compose an executive summary for several important reasons:

1. Decision-Making Support

Executives, stakeholders, or decision-makers rely on executive summaries to make informed decisions. By presenting key information clearly and concisely, an executive summary helps decision-makers understand the essence of the document and evaluate its significance and potential impact.

2. Accessible To A Wide Audience

Executive summaries are shared with diverse audiences, including those without specialized knowledge. By distilling complex information, executive summaries make the content more accessible to a broader range of readers.

3. Overview Of Research Or Project

An executive summary offers a high-level overview of the research or project, outlining objectives, methodology, and main findings. It provides a snapshot of the work, allowing readers to quickly assess its relevance and determine if they need to explore the full document.

The Executive Summary Length And Placement

The length and placement of an executive summary can vary depending on specific requirements and guidelines. However, there are some general considerations to keep in mind:

- **Length:** Executive summaries are typically kept relatively short, ranging from a few paragraphs to a couple of pages. They aim to provide a concise overview of the main points and findings of the document.
- **Placement:** In most cases, the executive summary is placed at the beginning of the document, before the main body or introduction. This allows readers to quickly access the key information without having to go through the entire document. However, in some cases, it may be placed at the end, serving as a

summary or recap of the main points for those who have already read the document.

It's important to note that the length and placement of the executive summary can be influenced by specific guidelines or requirements set by the organization or publication.

Structure of an Executive Summary For A Research Paper

The structure of an executive summary for a research paper can vary slightly depending on the specific requirements and nature of the research. However, a commonly used structure includes the following key elements:

➤ **Introduction**

Provide a brief introduction that sets the context for the research. Clearly state the purpose, objectives, and significance of the study.

➤ **Research Methodology**

Summarize the research methodology used in the study. Briefly explain the data collection methods, sample size, research design, and any statistical analyses employed. This helps establish the credibility and reliability of the research.

➤ **Key Findings**

Present the most significant findings of the research. Summarize the main results, trends, or patterns that emerged from the data analysis. Focus on the key outcomes that directly address the research objectives.

➤ **Conclusion**

Conclude the executive summary by summarizing the main points and emphasizing the overall significance of the research. Restate the main findings and their implications in a concise manner.

Executive Summary For A Research Paper Formatting

When formatting an executive summary for a research paper, it's important to follow the specific guidelines provided by the target journal or publication. General formatting considerations to keep in mind:

- **Title:** Include a clear and descriptive title for the executive summary at the top of the page. It should reflect the content and focus of the research paper.
- **Length:** The length of the executive summary can vary, but it is typically recommended to be concise, ranging from a few paragraphs to a maximum of one or two pages. Adhere to any specified word count or page limit guidelines.

- **Formatting Style:** Follow the formatting style required by the target journal or publication. This may include font type, font size, line spacing, and margins. Typically, a professional and readable font such as Times New Roman or Arial with a standard font size of 12 points is used.
- **Structure and Subheadings:** Use clear and informative subheadings to structure the content of the executive summary. This helps guide the reader through the main sections, such as Introduction, Methodology, Key Findings, Implications, Recommendations, and Conclusion.
- **Concise Writing:** Write in a concise and focused manner, using clear and simple language. Avoid unnecessary technical jargon or complex explanations. Use bullet points or numbered lists to present key findings or recommendations, making them easy to read and comprehend.
- **Proofreading and Editing:** Before finalizing the executive summary, carefully proofread and edit the content for clarity, coherence, and grammatical correctness. Ensure that the summary is free from errors and presents a professional image.

MECHANISMS FOR WRITING RESEARCH REPORT

When it comes to mechanisms for writing a research report, it's essentially about the structured process and methods used to plan, conduct, organize, and present research findings.

1. Understanding the Research Purpose

Before writing, you must clearly define:

- **Research problem or question:** What exactly are you investigating?
- **Objectives:** What do you want to achieve?
- **Scope and limitations:** What will be included or excluded in your study?

This sets the framework for the report.

2. Collecting and Organizing Data

Mechanisms for this stage include:

- **Primary data collection:** Surveys, experiments, interviews, observations.
- **Secondary data collection:** Books, journals, reports, online databases.
- **Data organization tools:** Spreadsheets, reference management software (e.g., Zotero, EndNote).

3. Planning the Structure

A research report typically follows a standard structure. Mechanisms here involve outlining the report:

1. **Title Page** – Name, title, date, institution.
2. **Abstract** – Brief summary of purpose, methods, results, conclusions.
3. **Introduction** – Research problem, background, objectives.
4. **Literature Review** – Existing studies and theoretical framework.
5. **Methodology** – How research was conducted (tools, procedures, sample).
6. **Results/Findings** – Data presented in tables, charts, or descriptive form.
7. **Discussion** – Interpretation, implications, comparisons with prior studies.
8. **Conclusion** – Summary of findings and recommendations.
9. **References/Bibliography** – All sources cited.
10. **Appendices** – Additional material like questionnaires, raw data.

4. Writing Mechanisms

- **Drafting:** Start with rough drafts, focusing on content over perfection.
- **Clarity and coherence:** Ensure logical flow between sections.
- **Academic style:** Formal tone, objective language, proper citations.
- **Using visuals:** Tables, graphs, and charts to present data effectively.

5. Reviewing and Editing

- **Self-review:** Check for grammar, structure, clarity, and logical flow.
- **Peer review:** Get feedback from colleagues or mentors.
- **Plagiarism check:** Ensure originality.
- **Formatting:** Follow required style (APA, MLA, Chicago, etc.).

6. Finalization

- Compile all sections neatly.
- Ensure all references and appendices are included.
- Submit or present according to guidelines.

Tip: The "mechanism" is essentially a **step-by-step workflow**: planning → data collection → organization → drafting → revising → finalizing. Using a structured approach ensures clarity, reliability, and academic rigor.

PRECAUTIONS FOR WRITING A REPORT

Report writing can be considered as an art which is learned through practicing and experience. While preparing project/research report, following precautions should be taken by the researcher:

1) Accuracy:

A research report should contain accurate information about the research problem and research objectives as inaccurate information may mislead the managers and can cause problems in decision-making.

2) Simplicity:

An ideal research report should be simple in every aspect. The layout, wording, grammar, phrasing, sequencing, tabulation, etc., should be developed with full attention to keep it simple.

3) Clear and Completeness:

A research report should be clear and complete. These two terms are complementary to each other. Researchers should not use ambiguous words or sentences while preparing a research report. The report should be clearly defined in terms of its objective, scope, sources, findings, etc. The concepts and techniques of the report should be explained appropriately, to make the research report complete in every sense.

4) Conciseness:

The research reports should be concise and brief enough to represent all the information needed by the executives as they want the facts and results only. But this feature of research report should not be compromised with the clarity and completeness of the report. The report should be able to highlight the essential points briefly without damaging its quality.

5) Comprehensibility and Readability:

Research report should be comprehensive and easily readable. There should be no confusion in the language or grammar of the contents. Technical language or terms should be avoided in the preparation of questionnaire. The language and content of the report should be in such a way that it can be understood by everyone.

6) Reliability:

The information included in the research report should be reliable and valid. It is also essential that all important facts should be included so that readers may understand it easily. Researchers should prepare the report with special care such that the erroneous information should not be included, as it confuses the readers.

7) Timelines:

The unique feature of research report is that it should be prepared within a stipulated time. Timeframe is the crucial aspect of a report, as data included in the report may become obsolete or irrelevant after certain period time.

8) Logical Content:

All the contents of a research report should be written in a logical way. None of the information should be included without proper investigation and analysis. There should be a sequence applied to all the components of report. The entire report should be categorized into several parts containing distinct facts and information and should facilitate the ease to locate the concerned topic in a report.

9) Original Content:

The contents of a research report should be original and specific. The research report should address some specific problem. The researcher should attempt to provide a solution to the problem being addressed.

10) Free from Errors:

The report should not contain any errors such as spelling errors, missing data, grammatical errors, incorrect calculations, etc. Researchers should ensure the report to be error-free before presenting it finally.

11) Good Appearance:

Research reports should be well typed, neat and clean. It should maintain this consistency throughout the report, to catch the attention of readers.

NORMS FOR USING TABLES, CHARTS, AND DIAGRAMS

1. Purpose and Relevance

- Use visuals only when they add value—to clarify, summarize, compare, or highlight information.
- Ensure each visual has a clear objective (showing trends, relationships, proportions, processes, etc.).

2. Clarity and Simplicity

- Keep designs simple and uncluttered.
- Avoid unnecessary decoration (shading, patterns, 3D effects).
- Use consistent colours and styles across visuals.

3. Correct Labelling

- Every table, chart, or diagram must have:

- A title (clear and descriptive).
- Labelled axes (for charts/graphs).
- Units of measurement (e.g., %, kg, years).
- Legends/keys when necessary.
- Number visuals systematically (e.g., *Table 1*, *Figure 2*).

4. Accuracy

- Ensure data used is correct, complete, and up-to-date.
- Present numbers consistently (e.g., same number of decimal places).
- Do not distort scales—use proportionate axes to avoid misleading interpretation.

5. Appropriate Type of Visual

Choose the visual that best represents the data:

- **Table** → exact values, detailed data.
- **Bar chart** → comparisons.
- **Line graph** → trends over time.
- **Pie chart** → parts of a whole.
- **Flowchart/diagram** → processes, relationships, structures.

6. Adequate Explanation

- Accompany visuals with a brief explanation in the text.
- Refer to them explicitly (e.g., “As shown in Figure 3…”).
- Avoid repeating all data—interpret or highlight key points instead.

7. Readability

- Use readable font sizes and clear color contrasts.
- Ensure data points, labels, and legends are legible even in printed or projected form.

8. Ethical and Professional Presentation

- Do not manipulate visuals to exaggerate or hide information.
- Credit data sources appropriately (e.g., “Source: Ministry of Health, 2024”).

9. Consistency

- Maintain uniform style across the document:
 - Same font type and size
 - Same color palette
 - Same numbering and caption format

10. Accessibility

- Use patterns/textures in addition to color for color-blind readers.
- Provide alt-text when used in digital or online documents.

APPENDIX

What is an appendix in writing?

An appendix is a section of a paper that provides additional information to support the main content. It may contain tables, images, extended explanations that will help readers understand the research without disrupting the flow of the main discussion.



1. What Is an Appendix?

An **appendix** (plural: *appendices* or *appendixes*) is a section placed **at the end** of a research report that contains **supplementary information**.

This information supports your study but is **too detailed, lengthy, or distracting** to include in the main body of the report.

2. Purpose of an Appendix

- To provide **additional evidence** or **background information**.
- To avoid interrupting the **flow of the main text**.
- To offer **transparency**, allowing readers to verify data or methods.
- To document **materials, tools, or calculations** used in the research.

3. What to Include in an Appendix

Common items include:

A. Data and Results

- Raw data tables
- Statistical calculations
- Detailed numerical results
- Graphs not essential to the main text

B. Research Instruments

- Questionnaires
- Interview guides

- Observation checklists
- Survey forms

C. Supplementary Materials

- Maps, photos, or diagrams
- Consent forms
- Coding schemes
- Extended quotations

D. Technical Details

- Formulas and derivations
- Detailed methodology
- Software code (if relevant)

4. How to Format an Appendix

A. Placement

Appendices appear **after the references/bibliography** and before any glossary or index (if used).

B. Naming

Each appendix should be labeled clearly:

- **Appendix A**
- **Appendix B**
- **Appendix C**, etc.

If there is only one, it can simply be titled **Appendix**.

C. Titles

Each appendix should have a descriptive title, such as:

- Appendix A: Survey Questionnaire
- Appendix B: Raw Test Scores

D. Order

Arrange appendices in the **order they are mentioned** in the text.

5. Referencing Appendices in the Text

When discussing material found in an appendix, refer to it explicitly:

- “See **Appendix A** for the full questionnaire.”
- “The raw data are provided in **Appendix B**.”

6. Best Practices

- Only include material that is **relevant and necessary** for transparency.

- Keep appendices **organized and clearly labeled**.
- Ensure appendix materials are **readable**, even if complex.
- Use consistent formatting (font, spacing, numbering).

7. Example Structure

References

Appendix A: Survey Questionnaire

Appendix B: Raw Data Tables

Appendix C: Interview Transcripts

APPENDICES

Appendix A: Survey Questionnaire

Title: Study Habits and Academic Performance Survey

Instructions: Please check (✓) the option that best describes your study habits.

1. How many hours do you study per day?

- Less than 1 hour
- 1–2 hours
- 3–4 hours
- More than 4 hours

2. Which study environment do you prefer?

- Quiet room
- Library
- With background music
- With friends

3. How often do you review your notes?

- Daily
- Weekly
- Before exams only
- Rarely

4. What study method do you use most often?

- Reading
- Highlighting
- Summarizing
- Practice tests

Appendix B: Raw Data Table

Table B1. Students' Study Hours and Corresponding GPA

Student Study Hours/Day GPA

01	1–2 hours	88
02	3–4 hours	92
03	Less than 1	80
04	1–2 hours	85
05	More than 4	94

Appendix C: Interview Guide

Title: Semi-Structured Interview for Selected Students

Sample Questions:

1. How do you usually prepare for your examinations?
2. What challenges do you face when studying?
3. Which study method do you find most effective, and why?
4. How do you balance academic tasks with personal activities?

Appendix D: Sample Informed Consent Form

I agree to participate in the research study titled “The Impact of Study Habits on Academic Performance Among Senior High School Students.”

I understand that participation is voluntary and that my responses will remain confidential.

Name: _____

Signature: _____

Date: _____

Appendices are usually formatted using three guides—APA, Chicago, and MLA, as described below.

APA	Chicago	MLA
Begin with the heading “Appendix” followed by	Times New Roman, 12-point text	Use letters to label multiple appendices,

letters, e.g., Appendix A		e.g., Appendix C
Place the heading at the top of the appendix	Add page numbers (as 1, 2, 3, ...) at the top right of each page; no page number on the front cover	Maintain the order of information
Ensure same order as in main paper	Bibliography should be the final page	Place the appendix before references
Include appendix after reference list		Start each appendix on a new page and number all pages

NORMS FOR USING INDEX

Below are the standard norms for using an Index and Bibliography in research methodology. These guidelines apply to academic theses, dissertations, and research reports across most disciplines, though specific formats (APA, MLA, Chicago, etc.) may differ slightly.

1. Norms for Using an Index in Research

An **Index** is usually placed at the end of a research work (after the bibliography) and is **not mandatory**, but is highly recommended in lengthy documents such as theses and books.

Purpose

- Helps readers locate specific concepts, terms, names, and topics quickly.
- Enhances usability of the document.

A. What to Include in an Index

- **Keywords** from the research
- Important **names, places, theories, methods, instruments**, etc.
- Technical terms the reader may search for
- Core concepts discussed across multiple sections

B. How to Prepare an Index

1. Alphabetical Order

- Entries should be arranged **alphabetically**.
- Subentries may be used for clarity (e.g., “Sampling → Purposive Sampling”).

2. Page Numbers

- Each index entry must include **accurate page numbers**.
- Use page ranges for topics spanning multiple pages (e.g., 45–47).

3. Consistency in Terms

- Use consistent terminology throughout the index.
- Avoid synonyms unless cross-references are required (e.g., Validity. See also Reliability).

4. Cross-References

- Use:
 - **See** → to point the reader to another entry
 - **See also** → for related terms

Example:

- **Correlation**, 34–36
- **Regression**, 40–48
- **Sampling**
 - stratified, 22–23
 - purposive, 24
- **Validity**, see Reliability

C. Formatting Norms

- Font size should match the rest of the document (or be slightly smaller).
- Generally single-spaced with hanging indents.
- Located after Appendices and before Bibliography (or after Bibliography, depending on institutional rules).

NORMS FOR USING A BIBLIOGRAPHY

A **Bibliography** is essential in all research works. It lists all sources consulted or cited

A. Purpose

- To acknowledge the work of other scholars
- To provide readers with sources for further study
- To ensure academic honesty and avoid plagiaris

B. What to Include

Depending on style guidelines:

- All **books, journal articles, websites, reports, theses, conference papers**
- Sometimes only **cited sources** are listed (called a References section, e.g., in APA)

C. General Norms

1. Follow a Standard Citation Style

Use one style consistently:

- **APA** (commonly used in social sciences)
- **MLA** (humanities)
- **Chicago/Turabian** (history, qualitative research)
- **Harvard** (general academic use)

2. Alphabetical Arrangement

Entries arranged alphabetically by **author's last name**.

3. Hanging Indentation

Second and subsequent lines of each entry should be indented.

4. Complete Details

Each entry must include:

- Author(s)
- Year of publication
- Title
- Publisher / Journal name
- Volume, Issue, and Page numbers (for articles)
- DOI or URL if required
- Access date for online sources (depending on style)

D. Examples (in different styles)

APA 7th Edition

Smith, J. A. (2021). Research methods in education. Sage Publications.

MLA 9th Edition

Smith, John. Research Methods in Education. Sage, 2021.

Chicago

Smith, John. Research Methods in Education. London: Sage, 2021.

Difference Between Index and Bibliography

Feature	Index	Bibliography
Purpose	Helps locate terms/topics in the	Lists sources used in research

MBA SEMESTER-I 105- BUSINESS RESEARCH METHODS

Feature	Index	Bibliography
	document	
Placement	End of document	End of document (before/after index)
Arrangement	Alphabetical terms + page numbers	Alphabetical by authors
Mandatory?	Optional	Mandatory

Bibliography Example (APA 7th Edition)

Books

- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage Publications.
- Kothari, C. R. (2004). *Research methodology: Methods and techniques* (2nd ed.). New Age International.

Journal Articles

- Kumar, R. (2018). Sampling techniques in social research. *Journal of Social Sciences*, 10(2), 45–59.
- Patel, S. (2020). Validity and reliability in educational research. *International Review of Education*, 66(1), 15–29.

Web Sources

- World Health Organization. (2021). *Research methods in health*. <https://www.who.int/research-methods>