Master of Library & Information Science
II - Semester
323 23

RESEARCH METHODOLOGY
### SYLLABI-BOOK MAPPING TABLE

**Research Methodology**

<table>
<thead>
<tr>
<th>Syllabi</th>
<th>Mapping in Book</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BLOCK I: FUNDAMENTALS OF RESEARCH</strong></td>
<td><strong>Unit 1:</strong> Foundations of Research (Pages 3-24)</td>
</tr>
<tr>
<td><strong>Unit-1:</strong> Foundations of Research; Nature, Definition and Objectives; Types of Research, Basis Concepts of Research; Scientific Method; Ethical Consideration of Research.</td>
<td><strong>Unit 2:</strong> Library and Information Science (LIS) as an Interdisciplinary Subject, Significance of Research in LIS; Areas of Research in Library and Information Science.</td>
</tr>
<tr>
<td><strong>Unit-2:</strong> Library and Information Science (LIS) as an Interdisciplinary Subject, Significance of Research in LIS; Areas of Research in Library and Information Science.</td>
<td><strong>Unit 3:</strong> Scientific Method–Nature of Research in Library &amp; Information Science.</td>
</tr>
<tr>
<td><strong>Unit-3:</strong> Scientific Method–Nature of Research in Library &amp; Information Science.</td>
<td><strong>Unit 4:</strong> Research Methods–Definitions–Sources–Advantages–Limitations Historical Method, Case Study Method, Survey Method, Experimental Method and other Methods (Field Investigation Research, Evaluation Research, Action Research, Ex Post Facto).</td>
</tr>
<tr>
<td><strong>BLOCK II: RESEARCH PROBLEMS &amp; LITERATURE SEARCH</strong></td>
<td><strong>Unit 5:</strong> Research Problem: Sources of Research Problem–Locating the Problem–Formulation of the Research Problem–Criteria in Selecting a Problem–Defining and Delimiting Problems.</td>
</tr>
<tr>
<td><strong>Unit-6:</strong> Literature Search–Importance of Surveying Related Literature–Library Sources, Research Reviews, Catalogue, Indices, Abstracts, Bibliographies, Microforms, Computerized Information Retrieval Systems.</td>
<td><strong>Unit 6:</strong> Literature Search (Pages 87-128)</td>
</tr>
<tr>
<td><strong>BLOCK III: RESEARCH PROBLEMS AND HYPOTHESIS</strong></td>
<td><strong>Unit 7:</strong> Hypothesis: Meaning, Importance, Types, Sources, Characteristics-Formulation of Hypothesis, Different Forms of Hypothesis–Difficulties in Formulation–Testing the Hypothesis.</td>
</tr>
<tr>
<td><strong>Unit-8:</strong> Planning of Research; Planning Process; Review of Literature, Election of a Problem-Problems, Process of Identification, Criteria of Selection, Formulation of Problem-Research Design-Essentials of Good Research Design &amp; Its Importance, Preparation of the Research Design/ Writing the Research Proposal.</td>
<td><strong>Unit 8:</strong> Planning of Research and Research Design (Pages 143-156)</td>
</tr>
</tbody>
</table>
BLOCK IV: TYPES OF RESEARCH AND DATA COLLECTION METHODS
Unit-9: Types: Descriptive, Diagnostic, Explanatory, and Experimental.
Unit-10: Data Collection, Primary and Secondary Data, Methods of Data Collection, Research Technique and Tools; Questionnaire, Schedule, Interview, Sampling, Scale and Check List, Library Records and Reports.
Unit-11: Questionnaire Construction & Design, Types of Questionnaire–Secondary Data, Sources and Precautions in the Use of Secondary Data.

Unit 9: Types of Research Designs
(Pages 157-172)
Unit 10: Data Collection
(Pages 173-206)
Unit 11: Questionnaire Design
(Pages 207-232)

BLOCK V: STATISTICAL SOFTWARE & STYLE MANUAL
Unit-12: Data Analysis, Interpretation and Presentation–Research Reporting.
Unit-13: Statistical Analysis Software SPSS, PSP and SOFA.
Unit-15: Bibliometrics, Scientometrics, and Informetrics, Webometrics.

Unit 12: Analysis of Data
(Pages 233-250)
Unit 13: Statistical Analysis Software
(Pages 251-264)
Unit 14: Report Writing
(Pages 265-284)
Unit 15: Informetrics and Webometrics
(Pages 285-300)
INTRODUCTION

BLOCK I:  FUNDAMENTALS OF RESEARCH

UNIT 1  FOUNDATIONS OF RESEARCH 1-24

1.0 Introduction
1.1 Objectives
1.2 Basic Concepts of Research: Nature, Definition and Objectives
   1.2.1 Research: Significance and Approach
   1.2.2 Applications of Research in Business Decisions
   1.2.3 Scientific Method
1.3 Types of Research
1.4 Ethical Consideration of Research
   1.4.1 Client's Ethical Code
   1.4.2 Researcher's Ethical Code
   1.4.3 Ethical Codes Related to Respondents
   1.4.4 Responsibility of Ethics in Research
1.5 Answers to Check Your Progress Questions
1.6 Summary
1.7 Key Words
1.8 Self Assessment Questions and Exercises
1.9 Further Readings

UNIT 2  LIBRARY AND INFORMATION SCIENCE (LIS) 25-36

2.0 Introduction
2.1 Objectives
2.2 Library and Information Science (LIS) as an Interdisciplinary Subject
2.3 Significance of Research in LIS
   2.3.1 Definition of Research
   2.3.2 Objectives of Research
2.4 Areas of Research in Library and Information Science
2.5 Answers to Check Your Progress Questions
2.6 Summary
2.7 Key Words
2.8 Self Assessment Questions and Exercises
2.9 Further Readings
UNIT 3  SCIENTIFIC METHOD  37-42
3.0 Introduction
3.1 Objectives
3.2 Understanding Scientific Method of Research
   3.2.1 Nature of Research in Library and Information Science
3.3 Answers to Check Your Progress Questions
3.4 Summary
3.5 Key Words
3.6 Self Assessment Questions and Exercises
3.7 Further Readings

UNIT 4  RESEARCH METHODS  43-76
4.0 Introduction
4.1 Objectives
4.2 Definition
4.3 Research Methods and its Advantages and Limitations
   4.3.1 Survey
   4.3.2 Case Study
   4.3.3 Census
   4.3.4 Experimental Research
   4.3.5 Focused Groups
4.4 Answers to Check Your Progress Questions
4.5 Summary
4.6 Key Words
4.7 Self Assessment Questions and Exercises
4.8 Further Readings

BLOCK II: RESEARCH PROBLEMS & LITERATURE SEARCH

UNIT 5  RESEARCH PROBLEM  77-86
5.0 Introduction
5.1 Objectives
5.2 Formulating the Research Problem
5.3 Answers to Check Your Progress Questions
5.4 Summary
5.5 Key Words
5.6 Self Assessment Questions and Exercises
5.7 Further Readings
UNIT 6  LITERATURE SEARCH  87-128

6.0 Introduction
6.1 Objectives
6.2 Importance of Surveying Related Literature
6.3 Uses of Library and Internet in Research
   6.3.1 Uses of Library in Research
   6.3.2 Uses of Internet in Research
6.4 Research Reviews, Abstracts and Microforms
6.5 Computerized Information Retrieval System
6.6 Answers to Check Your Progress Questions
6.7 Summary
6.8 Key Words
6.9 Self Assessment Questions and Exercises
6.10 Further Readings

BLOCK III: RESEARCH PROBLEM AND HYPOTHESIS

UNIT 7  HYPOTHESIS  129-142

7.0 Introduction
7.1 Objectives
7.2 Meaning, Importance, Sources and Characteristics
   7.2.1 Criteria for Hypothesis Construction
   7.2.2 Types of Statistical Hypothesis
7.3 Formulation of Hypothesis and Testing
   7.3.1 Procedure for Hypothesis Testing
   7.3.2 Committing Errors: Type I and Type II
   7.3.3 Null and Alternative Hypotheses
7.4 Answers to Check Your Progress Questions
7.5 Summary
7.6 Key Words
7.7 Self Assessment Questions and Exercises
7.8 Further Readings

UNIT 8  PLANNING OF RESEARCH AND RESEARCH DESIGN  143-156

8.0 Introduction
8.1 Objectives
8.2 Planning of Research and its Process
8.3 Research Design: Essentials of a Good Design and its Importance
  8.3.1 Preparation of the Research Design
8.4 Writing the Research Proposal
  8.4.1 Contents of a Research Proposal
  8.4.2 Types of Research Proposals
8.5 Answers to Check Your Progress Questions
8.6 Summary
8.7 Key Words
8.8 Self Assessment Questions and Exercises
8.9 Further Readings

BLOCK IV: TYPES OF RESEARCH AND DATA COLLECTION METHODS

UNIT 9 TYPES OF RESEARCH DESIGNS 157-172
  9.0 Introduction
  9.1 Objectives
  9.2 Exploratory Research Design
    9.2.1 Secondary Resource Analysis
    9.2.2 Case Study Method
    9.2.3 Expert Opinion Survey
    9.2.4 Focus Group Discussions
  9.3 Descriptive Research Design
    9.3.1 Cross-sectional Studies
    9.3.2 Longitudinal Studies
  9.4 Experimental Designs
  9.5 Diagnostic Research Design
  9.6 Answers to Check Your Progress Questions
  9.7 Summary
  9.8 Key Words
  9.9 Self Assessment Questions and Exercises
  9.10 Further Readings

UNIT 10 DATA COLLECTION 173-206
  10.0 Introduction
  10.1 Objectives
  10.2 Primary and Secondary Data
10.3 Methods of Data Collection
   10.3.1 Observation
   10.3.2 Interview
   10.3.3 Questionnaire
10.4 Sampling: Introduction
10.5 Scaling and Checklist
10.6 Library Records and Reports
10.7 Answers to Check Your Progress Questions
10.8 Summary
10.9 Key Words
10.10 Self Assessment Questions and Exercises
10.11 Further Readings

UNIT 11 QUESTIONNAIRE DESIGN 207-232

11.0 Introduction
11.1 Objectives
11.2 Questionnaire Construction, Design and Types
   11.2.1 Types of Questionnaire
11.3 Process of Questionnaire Designing
11.4 Secondary Data Sources and Precautions
   11.4.1 Uses of Secondary Data
   11.4.2 Advantages and Disadvantages of Secondary Data
   11.4.3 Types and Sources of Secondary Data
11.5 Answers to Check Your Progress Questions
11.6 Summary
11.7 Key Words
11.8 Self Assessment Questions and Exercises
11.9 Further Readings

BLOCK V: STATISTICAL SOFTWARE AND STYLE MANUAL

UNIT 12 ANALYSIS OF DATA 233-250

12.0 Introduction
12.1 Objectives
12.2 Data Analysis
15.3 Bibliometrics
15.4 Scientometrics
15.5 Webometrics
15.6 Answers to Check Your Progress Questions
15.7 Summary
15.8 Key Words
15.9 Self Assessment Questions and Exercises
15.10 Further Readings
Introduction

Research is the quest for knowledge or a systematic investigation in order to establish facts. It helps to solve problems and increase knowledge. The basic aim of research is to discover, interpret and develop methods and systems to advance human knowledge on diverse scientific matters. Thus, research is a process of enquiry and investigation. It helps to solve problems and increase knowledge. One of the main purposes of research is to review the existing knowledge and provide solutions to problems. There are different types of research, such as exploratory, descriptive, experimental and analytical. Exploratory research is done when few or no previous studies of the subject exist. Descriptive research is used to classify and identify the characteristics of a subject. Experimental research suggests or explains why or how something happens. Analytical research suggests or explains why or how something happens. Thus, one of the primary aims of research is to explain new phenomena and generate new knowledge.

Before conducting any research, a specific approach should be decided upon, called research methodology. Research methodology refers to the way research can be conducted. It is also known as the process of collecting data for various research projects. It helps to understand both the products as well as the process of scientific enquiry. A research process involves selection and formulation of a research problem, research design, sample strategy or sample design, as well as the interpretation and preparation of research report.

Research methodology is a very important function in today’s business environment. There are many new trends in research methodology through which an organization can function in this dynamic environment. There are two basic types of research approaches, namely quantitative and qualitative. The main emphasis of quantitative research is on collecting numerical data. It also concentrates on measuring the scale, range and frequency of a phenomenon. Qualitative research is more subjective in nature than quantitative research and involves analysis of data. Quantitative research involves examining the tangible aspects of research, such as values, attitudes and perceptions.

This book, Research Methodology has been divided into fifteen units. The book has been written in keeping with the self-instructional mode or the SIM format wherein each Unit begins with an Introduction to the topic, followed by an outline of the Unit Objectives. The detailed content is then presented in a simple and organized manner, interspersed with Check Your Progress questions to test the student’s understanding of the topics covered. A Summary along with a list of Key Words, set of Self-Assessment Questions and Exercises and Further Readings is provided at the end of each Unit for effective recapitulation.
UNIT 1  FOUNDATIONS OF RESEARCH

1.0  INTRODUCTION

Research is a watchful process of systematic inquiry that involves the collection of data; documentation of information; analysis and interpretation of information, in accordance with suitable approaches and practices. In this unit, you will study about the basics of research along with the definition, nature, objectives and scientific method of research. The unit also discusses about the different types of research and complete process of research. In addition to this, you will also study about ethical standards of research.

1.1  OBJECTIVES

After going through this unit, you will be able to:

- Understand the basic concept of research
- Discuss the nature, definition and objectives of research
- Describe the different approaches to research
1.2 BASIC CONCEPTS OF RESEARCH: NATURE, DEFINITION AND OBJECTIVES

Research in common parlance refers to search for knowledge. One can also define research as a scientific and systematic search for pertinent information on a specific topic. In fact, research is an art of scientific investigation. According to Advanced Learner’s Dictionary of Current English, ‘research is a careful investigation or enquiry, especially a thorough search for new facts in any branch of knowledge.’ Well-known authors on research methodology, L.V. Redman and A.V.H. Mory (1923) defined research as a ‘systematized effort to gain new knowledge.’ Some people consider research as a voyage of discovery that involves movement from the known to the unknown.

Research in a technical sense is an academic activity. Renowned author Clifford Woody defined research as an activity that comprises defining and redefining problems, formulating a hypothesis; collecting, organizing and evaluating data; making deductions and reaching conclusions; and carefully testing the conclusions to determine if they support the formulated hypothesis. Eminent authors D. Slesinger and M. Stephenson, in the Encyclopaedia of Social Sciences, defined research as ‘the manipulation of things, concepts or symbols for the purpose of generalizing, extending, correcting or verifying the knowledge, whether that knowledge aids in the construction of theory or in the practice of an art’. Research is, thus, an original contribution to the existing stock of knowledge making for its advancement.

Principles of Research

The basic principles of research include a systematic process to identify a question or problem, set forth a plan of action to answer the question or resolve the problem, and meticulously collect and analyse data. In conducting any research, it is crucial to choose the right method and design for a specific researchable problem. All research is different. However, the following factors are common to all good pieces of research:

- It is based on empirical data.
- It involves precise observations and measurements.
- It is aimed at developing theories, principles and generalizations.
- There are systematic, logical procedures involved.
- It is replicable.
- The findings of the research need to be reported.
Objectives of Research

The objective of any research is to find answers to questions through the application of scientific procedures. The main aim of any research is to explore the hidden or undiscovered truth. Even though each research study has a specific objective, the research objectives in general can be categorized into the following broad categories:

- **Exploratory or Formulative Research Studies**: These are aimed at gaining familiarity with a particular phenomenon or at gaining new insights into it.
- **Descriptive Research Studies**: These are aimed at accurately portraying the characteristics of a particular event, phenomenon, individual or situation.
- **Diagnostic Research Studies**: These studies try to determine the frequency with which something occurs.
- **Hypothesis Testing Research Studies**: These studies test a hypothesis and determine a causal relationship between the variables.

Characteristics and purpose of research

The following are the characteristics of research:

- It is a systematic and critical investigation into a phenomenon.
- It uses scientific methods.
- It is goal-orientated and logical.
- It requires empirical evidence.
- It focuses on finding answers to questions and solution to problems.

The following points will help in understanding the purpose of research:

- Research helps in extending the knowledge of human beings, the environment and natural phenomenon to others.
- It brings out the information which is not developed fully during the ordinary course of life.
- It verifies the existing facts and identifies the changes in these existing facts.
- It helps in developing facts for critical evaluation.
- It analyses the interrelationship between variables and derives casual explanations.
- It develops new tools and techniques for those who study unknown phenomenon.
- It helps in planning and development.

1.2.1 Research: Significance and Approach

Research involves developing a scientific temperament and logical thinking. The significance of research-based answers can never be underestimated. The role of research is especially important in the fields of economics, business, governance, and so on. Here research helps in finding solutions to problems.
Foundations of Research

NOTES

encountered in real life. Decision-making is facilitated by applied research. Research is also of special significance in the operational and planning processes of business and industry. Here logical and analytical techniques are applied to business problems in order to maximize profits and minimize costs. Motivational research is another key tool in understanding consumer behaviour and health related issues. Responsible citizenship concerns can all be addressed through good research findings. Social relationships involving issues such as attitudes, interpersonal helping behaviour; environmental concerns such as crowding, crime, fatigue, productivity and; other practical issues are all capable of being addressed well by scientific research.

Social science research is extremely significant in terms of providing practical guidance in solving human problems of immediate nature.

Research is also important as a career for those in the field of academics. It could be a career option for professionals who wish to undertake research to gain new insights and idea generation. Research also fosters creative thinking and new theorizations.

Research for its own sake and for the sake of knowledge, and for solving different problems, all that is required is formal training in scientific methodology.

Approaches to Research

Quantitative approach and qualitative approach are the two basic approaches to research. These two paradigms are based on two different and competing ways of understanding the world. These competing ways of comprehending the world are reflected in the way research data is collected (for example, words versus numbers) and the perspective of the researcher (Perspectival versus Objective). The perspectives of the participants are very critical.

(i) Quantitative Approach: If there has been one overwhelming consensus among academic psychologists on a single point over the past few decades, it is that the best empirical research in the field is firmly grounded in quantitative methods. In this approach, data is generated in quantitative form, and then that data is subjected to rigorous quantitative analysis in a rigid and formal fashion. Inferential, experimental and simulation approaches are the sub-classifications of quantitative approach. Inferential approach to research focuses on survey research where databases are built studying samples of population and then these databases are used to infer characteristics or relationships in populations. In experimental approach, greater control is exercised over the research environment and often, some independent variables are controlled or manipulated to record their effects on dependent variables. In simulation approach, an artificial environment is constructed within which relevant data and information is generated. This way, the dynamic behaviours of a system are observed under controlled conditions.

(ii) Qualitative Approach: This approach to research is concerned with subjective assessment of attitudes, opinions and behaviour. Research in such a situation is a function of researcher’s insight and impressions. Such
Table 1.1 provides us with types of research, methods employed and techniques used by these types of research.

<table>
<thead>
<tr>
<th>Type</th>
<th>Methods</th>
<th>Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Library Research</td>
<td>(i) Analysis of historical records</td>
<td>Recording of notes, content analysis, tape and film listening and manipulations, reference and abstract guides, content analysis.</td>
</tr>
<tr>
<td></td>
<td>(ii) Analysis of documents</td>
<td></td>
</tr>
<tr>
<td>2. Field Research</td>
<td>(i) Non-participant direct observation</td>
<td>Observational behavioural scales, use of score cards, etc.</td>
</tr>
<tr>
<td></td>
<td>(ii) Participant observation</td>
<td>Interational recording, possible use of tape recorders, photographic techniques.</td>
</tr>
<tr>
<td></td>
<td>(iii) Mass observation</td>
<td>Recording mass behaviour, interview using independent observers in public places.</td>
</tr>
<tr>
<td></td>
<td>(iv) Mail questionnaire</td>
<td>Identification of social and economic background of respondents.</td>
</tr>
<tr>
<td></td>
<td>(v) Opinionnaire</td>
<td>Use of attitude scales, projective techniques, use of goniometric scales.</td>
</tr>
<tr>
<td></td>
<td>(vi) Personal interview</td>
<td>Interviewer uses a detailed schedule with open and closed questions.</td>
</tr>
<tr>
<td></td>
<td>(vii) Focused interview</td>
<td>Interviewer focuses attention upon a given experience and its effects.</td>
</tr>
<tr>
<td></td>
<td>(viii) Group interview</td>
<td>Small groups of respondents are interviewed simultaneously.</td>
</tr>
<tr>
<td></td>
<td>(ix) Telephone survey</td>
<td>Used as a survey technique for information and for discerning opinion; may also be used as followup questionnaire.</td>
</tr>
<tr>
<td></td>
<td>(x) Case study and life history</td>
<td>Cross-sectional collection of data for intensive analysis, longitudinal collection of data of intensive character.</td>
</tr>
<tr>
<td>3. Laboratory Research</td>
<td>Small group study of random behaviour, play and role analysis</td>
<td>Use of audio-visual recording devices, use of observers, etc.</td>
</tr>
</tbody>
</table>

Methods versus Methodology

**Research Methods:** They refer to all the methods the researchers use while studying the research problems and while conducting research operations. In general, research methods can be categorized into the following three groups:

(i) The first group includes the methods that are concerned with the data collection.

(ii) The second group includes the statistical techniques needed for mapping relationships between the unknowns and the data.

(iii) The third group contains the methods necessary to evaluate the accuracy of the results obtained.

**Research Methodology:** It is the procedure that helps to systematically proceed in steps to solve a research problem. Research methodology is a broader concept.
that includes not the research methods but also the logic behind the research methods in the context of a particular research study; and it explains the reasons for using particular research methods and statistical techniques. Research mythology also defines how the data should be evaluated to get the appropriate results.

1.2.2 Applications of Research in Business Decisions

The discussion so far points out the role and significance of research in aiding business decisions. The question one might ask here is about the critical importance of research in different areas of management. Is it most relevant in marketing? Do financial and production decisions really need research assistance? Does the method or process of research change with the functional area? Figure 1.1 explains the complete research process.

![Fig 1.1 The Process of Research](image-url)
The answer to all the above questions is NO. Business managers in each field—whether human resources or production, marketing or finance—are constantly being confronted by problem situations that require effective and actionable decision-making. Most of these decisions require additional information or information evaluation, which can be best addressed by research. While the nature of the decision problem might be singularly unique to the manager, organization and situation, broadly for the sake of understanding, it is possible to categorize them under different heads.

Marketing Function

This is one area of business where research is the lifeline and is carried out on a vast array of topics, and is conducted both in-house by the organization itself and outsourced to external agencies. Broader industry- or product-category-specific studies are also carried out by market research agencies and sold as reports for assisting in business decisions. Studies like these could be:

- Market potential analysis; market segmentation analysis and demand estimation
- Market structure analysis which includes market size, players and market share of the key players
- Sales and retail audits of product categories by players and regions as well as national sales; consumer and business trend analysis—sometimes including short- and long-term forecasting.

However, it is to be understood that the above-mentioned areas need not always be outsourced; sometimes they might be handled by a dedicated research or new product development department in the organizations. Other than these, an organization also carries out researches related to all four functions of marketing, such as:

- **Product Research:** This would include new product research; product testing and development; product differentiation and positioning; testing and evaluating new products, and packaging research; and brand research—including equity to tracks and imaging studies.
- **Pricing Research:** Price determination research; evaluating customer value; competitor pricing strategies, and alternative pricing models and implications.
- **Promotional Research:** Includes everything from designing of the communication mix to design of advertisements, copy testing, measuring the impact of alternative media vehicles and impact of competitors’ strategy.
- **Place Research:** Includes locational analysis, design and planning of distribution channels, and measuring the effectiveness of the distribution network.
These days, with the onset of increased competition and the need to convert customers into committed customers, Customer Relationship Management (CRM), customer satisfaction, loyalty studies and lead user analysis are also areas in which significant research is being carried out.

**Personnel and Human Resource Management**

Human Resources (HR) and organizational behaviour is an area which involves basic or fundamental research as a lot of academic, macro level research may be adapted and implemented by organizations into their policies and programmes. Applied HR research by contrast is more predictive and solution oriented. Though there are a number of academic and organizational areas in which research is conducted, yet some key contemporary areas which seem to attract more research are as follows:

- **Performance Management**: Leadership analysis development and evaluation; organizational climate and work environment studies; talent and aptitude analysis and management; and organizational change implementation, management and effectiveness analysis.
- **Employee Selection and Staffing**: This includes pre and on-the-job employee assessment and analysis; and staffing studies.
- **Organizational Planning and Development**: Culture assessment—either organization specific or the study of individual, and merged culture analysis for mergers and acquisitions; and manpower planning and development.
- **Incentive and Benefit Studies**: These include job analysis and performance appraisal studies, recognition and reward studies, hierarchical compensation analysis; and employee benefits and reward analysis, both within the organization and industry best practices.
- **Training and Development**: These include training need gap analysis; training development modules; monitoring and assessing impact; and effectiveness of training.
- **Other Areas**: Other areas include employee relationship analysis; labour studies; negotiation and wage settlement studies; absenteeism and accident analysis; turnover and attrition studies; and work-life balance analysis.

Critical success factor analysis and employer branding are some emerging areas in which HR research is being carried out. The first is a participative form of management technique, developed by American Organizational theorist John F. Rockart (1981) in which the employees of an organization identify their critical success factors, and help in customizing and incorporating them in developing the mission and vision of their organization. The idea is that a synchronized objective will benefit both the individual and the organization, and which will lead to a commitment and ownership on the part of the employees. Employer branding is another area which is being actively investigated as the customer perception (in this case, it is the internal customer, i.e., the employee) about the employer or the
employing organization has a strong and direct impact on his intentions to stay or leave. Thus, this is a subjective qualitative construct which can have hazardous effect on organizational effectiveness and efficiency.

Financial and Accounting Research

The area of financial and accounting research is so vast that it is difficult to provide a pen sketch of the research areas. In this section, we are providing just a brief overview of some research topics:

- **Asset Pricing, Corporate Finance and Capital Markets**: The focus here is on stock market response to corporate actions (IPOs or Initial Public Offerings, takeovers and mergers), financial reporting (earnings and firm specific announcements) and the impact of factors on returns, e.g., liquidity and volume.

- **Financial Derivatives and Interest Rate and Credit Risk Modelling**: This includes analysing interest rate derivatives, development and validation of corporate credit rating models and associated derivatives; analysing corporate-decision making and investment risk appraisal.

- **Market Based Accounting Research**: Analysis of corporate financial reporting behaviour; accounting-based valuations; evaluation and usage of accounting information by investors and evaluation of management compensation schemes.

- **Auditing and Accountability**: This includes both private and public sector accounting studies, analysis of audit regulations; analysis of different audit methodologies; governance and accountability of audit committees.

- **Financial Econometrics**: This includes modelling and forecasting in volatility, risk estimation and analysis.

- **Other Areas**: Other related areas of investigation are in merchant banking and insurance sector, and business policy and economics areas.

Considering the nature of the decision required in this area; the research is a mix of historical and empirical research. Behavioural finance is a new and contemporary area in which, probably, for the first time, subjective and perceptual variables are being studied for their predictive value in determining consumer sentiments.

Production and Operation Management

This area of management is one in which quantifiable implementation of the research results takes on huge cost and process implications. Research in this area is highly focused and problem specific. The decision areas in which research studies are carried out are as follows:

- **Operation Planning**: These include product/service design and development; resource allocation and capacity planning.

- **Demand forecasting and decision analysis**.
NOTES

- **Process Planning**: Production scheduling and material requirement management; work design planning and monitoring.
- Project management and maintenance management studies.
- Logistics and supply chain, and inventory management analysis.
- **Quality Estimation and Assurance Studies**: These include Total Quality Management (TQM) and quality certification analysis.

This area of management also invites academic research which might be macro and general but helps in developing technologies, such as JIT (Just-In-Time) technology and EOQ (Economy Order Quantity)—an inventory management model, which are then adapted by organizations for optimizing operations.

**Cross-Functional Research**

Business management being an integrated amalgamation of all these and other areas sometimes requires a unified thought and approach to research. These studies require an open orientation where experts from across the disciplines contribute to and gain from the study. For example, an area, such as new product development, requires the commitment of the marketing, production and consumer insights team to exploit new opportunities. Other areas requiring cross functional efforts are as follows:

- Corporate governance and ethics—the role of social values and ethics, and their integration into a company’s working is an area that is of critical significance to any organization.
- Technical support systems, enterprise resource planning systems, knowledge management, and data mining and warehousing are integrated areas requiring research on managing coordinated efforts across divisions.
- Ecological and environmental analysis; legal analysis of managerial actions; human rights and discrimination studies.

**1.2.3 Scientific Method**

To distinguish between non-scientific and scientific method, we would like to consider a few definitions of research.

One of the earliest distinctions was made by Lundberg (1942) who stated ‘Scientific methods consist of systematic observation, classification, and interpretation of data. Now obviously, this process is one in which nearly all people engage in their daily life. The main difference between our day-to-day generalizations and the conclusions usually recognized as the scientific method lies in the degree of formality, rigorousness, verifiability, and general validity of the latter.’

Fred Kerlinger (1986) also validated the thought and stated that ‘Scientific research is a systematic, controlled and critical investigation of propositions about various phenomena.’ Grinnell (1993) has simplified the debate and stated ‘The
The word *research* is composed of two syllables, *re* and *search*. The dictionary defines the former as a prefix meaning again, anew or over again and the latter as a verb meaning to examine closely and carefully, to test and try, or to probe. Together they form a noun describing a careful, systematic, patient study and investigation in some field of knowledge, undertaken to establish facts or principles.

Thus, drawing from the common threads of the above definitions, we derive that management research is an unbiased, structured, and sequential method of enquiry, directed towards a clear implicit or explicit business objective. This enquiry might lead to validating existing postulates or arriving at new theories and models.

### Check Your Progress

1. What do you mean by research?
2. What is the importance of good research?
3. Name the two basic approaches to research.
4. What are the sub-classifications of quantitative approach?
5. What is critical success factor analysis?

### 1.3 TYPES OF RESEARCH

Although research is a vast subject and is difficult to categorize, it can be classified according to its intent or as per the methods of study.

On the basis of intent, research can be classified as follows:

- **Pure research:** It is done only for the sake of knowledge. The intention is not to apply it in regular practice. Pure research is also called basic or fundamental research. It is not focused on specific problems, but instead it focuses on the extension of knowledge. New theory or refinements of an existing theory are developed with the help of pure research. It lays the foundation for applied research. It helps in finding the critical factors in a problem. It helps in generating alternative solutions and choosing the best one amongst them.

- **Applied research:** When real-life problems require some solution and decision-making, applied research is carried out. This means that applied research is problem oriented and action directed. It brings immediate and practical results; for example, marketing research carried on for identifying customer habits to purchase something. Though it is problem oriented and action directed, it can contribute to the development of theoretical knowledge by leading to the discovery of new facts.

- **Exploratory research:** It is also called formulative research. When a researcher has no knowledge or little knowledge about an unfamiliar
problem, they do a preliminary study. The objective of this research is to
generate new ideas, gather new facts, precise formulation of problem and
increasing familiarity of the researcher to the unfamiliar problem. Renowned
theorist Katz conceptualizes two levels of exploratory research. At the first
level is the discovery of significant variables in particular situations; at the
second, the discovery of relationship between variables.

- **Descriptive research**: In this research, facts are analysed in detail for
clear understanding. This research is simple in nature and in its application.
It is more specific than exploratory research. It focuses on the problem
under study and also aims at a classification of the range of elements
comprising the subject matter of study. Empirical observations are used to
conceptualize the problems and facts. It highlights methods of data collection
and interpretation.

- **Diagnostic research**: It is just like descriptive research but with a different
focus. It is aimed towards in depth approaches to reach the basic casual
relations of a problem and possible solutions for it. Prior knowledge of the
problem is required for this type of research. Problem formulation, defining
the population correctly for study purposes, proper methods for collecting
accurate information, correct measurement of variables, statistical analysis
and tests of significance are essential in diagnostic research.

The classification of research can be done as per methods of study in the following
manner:

- **Fundamental**: This type of research is mainly concerned with identifying
certain important principles in a specific field. It intends to find out information
that has a broad base of application. Examples of fundamental research are
John Robinson’s imperfect competition theory in Economics and Maslow’s
hierarchy of needs theory in motivation.

- **Applied**: This type of research aims at finding a solution to an immediate
problem, faced by a society or an industrial organization. It is supposed to
discover a solution to some basic practical problems. Applied research
suggests corrective methods to minimize a social or business problem.

- **Historical**: Historical research studies the social effects of the past that
may have given rise to current situations, i.e., past incidents are used to
analyse the present as well as the future conditions. The study of the current
state of Indian labour based on past labour union movements in the Indian
economy to formulate the Indian Labour Policy is an example of this type
of research.

- **Formulative or exploratory**: It helps examine a problem with suitable
hypothesis. This research, on social science, is mainly significant for clarifying
concepts and innovations for further researches. The researchers are mainly
concerned with the principles of developing hypothesis and testing with
statistical tools.
• **Experimental**: The experimental type of research enables a person to
calculate the findings, employ the statistical and mathematical devices, and
measure the results, thus, quantified.

• **Ex post facto**: This type of research is the same as experimental research,
which is conducted to deal with the situations that occur in or around an
organization. Examples of such a research are market failure of an
organization’s product being researched later, and research into the causes
for a landslide in the country.

• **Case study**: This method undertakes intensive research that requires a
thorough study of a particular unit.

### 1.4 ETHICAL CONSIDERATION OF RESEARCH

Ethical standards are extremely important no matter what be the field of study.
This takes a special meaning in the conduction of research. Rowley (2004) has
put it very simply as 'conducting research ethically is concerned with respecting
privacy and confidentiality, and being transparent in the use of research data. Ethical
practices hinge on respect and trust and approaches that seek to build rather than
demolish relationships.’ Russeft et al. (1999) advocated that while conducting
business research, the approach must be professional and responsible, the data
collection must be attempted with the respondent's consent under appropriate
and ethically correct methods; and, last but not the least, the interpretation has to
be done in a careful unbiased manner.

A number of corporations have developed their own code of ethics regarding
the conduct of research. While this practice of defining business ethics, which
includes research ethics, is common in most organizations in the West, in India this
is spelt out and documented in the pharmaceutical sector and some banks like
HSBC. Besides this, there are also well established and detailed ethical guidelines
available from international bodies, for example, the Social Research Association’s
(SRA’s) ethical guidelines, the American Psychological Association (APA) code
of ethics, code of standards and ethics for survey research designed by the Council
of American Survey Research Organizations (CASRO), American Marketing
Association (AMA) and Business Marketing Association (BMA) code of conduct
and ethics.

To understand the code of ethics involved in research, one needs to
understand the three significant stakeholders involved in any research, namely:

- The sponsoring clients or decision-makers.
- The respondents from whom one seeks the information.
- The researcher himself/herself while administering and compiling the study.

Each one of these entities has their own specific interests and needs and,
thus, the ethical concerns regarding each one would be unique. Thus, the following
sections present brief guidelines on the ethical issues and their management.
14.4.1 Client’s Ethical Code

Similar to any other business transaction, research is also an exchange process between various people. The first of these is the one between the sponsoring client and the investigator. Thus both parties have an ethical obligation towards the other. In case the study is being conducted for a business client, complete transparency in terms of data gathering and interpreting is a must. It has been observed that the client might be a business manager who because of his own personal interests might steer the results in a specific direction in order to fulfil a hidden agenda. For example, in case a warehousing organization is looking at business expansion and hires a research agency to conduct a research study in order to provide directions. It might so happen that the business manager from the client side, who is dealing with the research agency owns a transport fleet and thus wants the researcher to recommend courier and transit warehousing services as business opportunities that the company can go into.

It has been commonly found amongst small and relatively younger firms to ask for proposals from research agencies for the conduct of a study. However, once they obtain the details of the intended methodology, they usually get the study conducted by their own team or by trainees at a low to minimal cost to the company. And since the proposals are the first stage of a research bid, the company is under no obligation to pay for the research methodology collected by them in an underhand manner.

Another instance could be that even though the initial exploratory research and literature review indicate the nature of the respondent population, the client might, based on his own notions, force the researcher to undertake the study on a specific population. For example, if a new technology is being introduced in the company and the use requires computer literacy, the client might ask the researcher to measure the acceptability of the product amongst only the computer-savvy population. Thus the results would automatically be bent towards acceptance.

Sometimes, the interpretation and recommendations might be beyond the scope of a study. For example, in the organic food study, which was conducted amongst retailers and consumers, the client might ask the researcher to suggest strategies for educating and building usage and recommendations amongst dieticians and doctors.

It is recommended in this instance that the researcher must conduct a comprehensive exploratory research and develop clearly stated objectives that do not leave any scope for unethical intervention. Secondly, he must tell business manager that unless the results are unbiased the study will not contribute to informed decision making. In case of an unethical manager or client, it is best to avoid making recommendations and formulating strategies and leave the use or non-use of the data to the manager. And if nothing works it is best to terminate the research study, as unethical reporting and compilation is bound to spoil the researcher’s reputations.
1.4.2 Researcher’s Ethical Code

Since the researcher is the most involved and main person responsible for the study, it is very important that the highest ethical standards be maintained by him. Some specific checks he can look at are as follows:

Quality control

A very important consideration, both short-term and long-term, is to maintain the standards of quality in the conduct of the study. The researcher must be absolutely objective and correct in choosing the research design that would be right for the study. For example, for studying the impact of a mathematics study programme on an experimental group of children, the researcher must have a matched control group of children with a similar understanding of mathematics so that the comparison is correct.

Sometimes, the client might be unaware of the analytical rules and conditions for the result to be valid, thus it is the responsibility of the researcher to be absolutely transparent about the significance of the results obtained and refrain from emphasizing findings that might be of very little strength or value.

Privacy control

The most significant and important ethical concern of a research study is the issue of trust and confidentiality. At no cost must the researcher reveal any aspect of the study without the consent of the client. This could be in terms of not revealing the name of the company. For example, if the client is interested in finding out the comparative standing of their product with the competitor’s product, it becomes critical to conduct the study amongst users of the product category rather than only the company brand in order to get an unbiased evaluation.

The researcher might also need to guard the reason or purpose of the study. For example if the client wants to measure a new product potential, then revealing the reason for the study might lead to the concept or idea being adopted and converted into a product prototype by someone else before the client is out with the offering. The third level of confidentiality that the researcher must ensure is the complete confidentiality of the findings till the research outcome has been converted into a business decision. For example, based on the organizational health index of its workers and the attrition rate, the correlation between the two variables might be alarming enough to require a major restructuring of the existing employee benefits and work policy. Or the research study might involve a comprehensive and detailed study of potential candidates being considered for the role of the CEO, as the existing leader is due for retirement. Thus, revelation of the findings of such research might lead to turbulence and divided opinion in the organization. Thus the results should not be made available to all till they have been brought into action.
1.4.3 Ethical Codes Related to Respondents

The most important and vulnerable person in the research study is the respondent from whom the data is to be collected. Every association and organization that is directly or indirectly involved with research has made clear and detailed guidelines for ensuring that unethical treatment of the respondent does not happen.

The American Association for Public Opinion Research has formulated the following code of ethics for survey researchers, with reference to the respondent:

- We shall strive to avoid the use of practices or methods that may harm, humiliate or seriously mislead survey respondents.
- Unless the respondent waives confidentiality for specific uses, we shall hold as privileged and confidential all information that might identify a respondent with his or her responses. We shall also not disclose or use the names of respondents for non-research purposes unless the respondent grants us permission to do so.

Study disclosure

The researcher needs to have complete and transparent information regarding the purpose of collecting data and what sort of information would be required from the respondent. The person must know what kind of questioning would be done, so that he is able to understand what the researcher is looking for and whether he has the information, whether he wants to share all or part of it and also how much time and effort would be needed. For example, for a new concept test or a segmentation analysis or an organizational climate survey the administration would require considerable time and commitment from the respondent. Secondly, if it is a before-and-after product acceptability or usage study, again the person would be contacted twice to assess the experience Thus the researcher needs to be absolutely truthful about the nature and objectives of the study.

Coercion and influence

The researcher should not at any stage, either before or during the data collection stage, try to pressurize the respondent through persuasive influence or by forcing him to share information. For example, if the respondent has been through some traumatic experience, he/she might not want to share all details with a stranger, even if it is for an objective study. Schinke and Gilchrist (1993) state that under standards set by the National Commission for the protection of human subjects, all informed-consent procedures must meet three criteria:

- Participants must be competent to give consent
- Sufficient information must be provided to allow for a reasonable decision
- Consent must be voluntary and uncoerced

Sometimes, it may so happen that the respondent is too young or too old or not literate and thus, unable to understand when the researcher might be either
leading him/her to give certain preset answers or trying to force the person to share information that he does not want to reveal or which once shared might be misinterpreted.

**Sensitivity and respect**

There are certain issues like shoplifting or sexual orientation, which are not topics that can be managed in a structured, impersonal manner. The researcher should devote more time here and also keep the questions more open-ended, and usually such situations need a considerable rapport formation and a non-threatening atmosphere. The researcher, at all times, would need to treat the respondent with due respect and be transparent about the nature and objective of the questioning.

**Experimentation and implication**

In case the respondent is going to be part of the experimental group subjected to any sort of treatment, for example, a new shampoo trial or an intervention programme that may involve some behavioural change, complete information must be given regarding the course of the experiment and any risk, even minimal, which might be involved. The researcher, thus, must ensure minimal risk to the respondent and should in no way cause any harm to the person, even if it is for the quest of knowledge. Bailey (1978) describes this ‘harm’ as not only hazardous or medical experiments but also any social research that might involve such things as discomfort, anxiety, harassment, invasion of privacy or demeaning or dehumanizing procedures.

**Agreement or consent**

Once the researcher has clearly communicated the purpose, the nature and likely outcome of the study, it is advisable to make a mutual written or unwritten contract. This ensures that there is no unpleasantness or legal confrontation on either side. Another advantage of this is that in case a point was not very clear the issue gets clarified. For example, for a personal care usage study, the consumer might be under the impression that a questionnaire on usage would be filled in when actually the researcher wants to observe/record the usage ritual. This might call for some invasion of privacy of the respondent by the researcher, and thus taking the consent beforehand would make things clear for both the parties.

Sometimes, the nature of the study might require that the name of the company be disguised. For example, one cannot start a study by saying, ‘We are conducting a survey for Mother Dairy milk; which do you think is the best milk in the city?’

Thus, here the debriefing about the company sponsoring the research can be revealed after the data has been collected, and the purpose of the disguise can be revealed. This ensures respondents’ goodwill and cooperation.

**1.4.4 Responsibility of Ethics in Research**

Besides ensuring that specific protocols and codes be followed for the two benefactors (client) and contributors (respondents), there are some basic tenets
that the researcher must not forego. These are significant not only for the body of knowledge that the researcher is contributing to but also for the society in which we exist.

**NOTES**

**Professional creed**

We have already discussed this in detail in both the sections above. However, here for professional creed, we refer to the overall conduct of the researcher, who has to be truthful during all phases of the study, whether in the conceptualization, conduction or presentation of the research study.

- At no stage should the researcher exaggerate or underplay the expense or effort incurred in the conduct of the study. Thus, sometimes the investigator might overclaim the expense incurred in travel or field visit. On the other hand, he might underpay the field investigators that he has kept for data collection by hiring undergraduate students rather than professional investigators.

- The respondent group being studied should be a true representative of the identified respondent population studied and not a skewed and biased sample. Another unethical practice observed is that the researcher might conduct the study with a professional group of respondents who are well versed in the response technique and thus give ‘good’ or predictable answers.

- The data and the questionnaire completed should be on authentic, real-time conduction, with actual respondents representative of the population under study and not fake completion done by the field investigators themselves.

- The findings and results should be presented as they were found based on actual conduction and under no circumstances must the researcher attempt to fudge or manipulate the results of the study.

**Professional confidentiality**

The researcher must bear the responsibility to maintain the confidentiality of the research findings and not making public any aspects of the study, in an apparent or camouflaged manner. This code of ethics applies both to the sponsoring client, as well as the respondent. The anonymity and privacy of the respondent is to be respected and not violated. Also, recording private or personal behaviour with hidden devices is considered a monumental violation of an individual’s right to privacy (e.g., observing people in a fitting room with a hidden camera).

The right to privacy and confidentiality takes on a new meaning in cyberspace, where the respondent’s personal and demographic details are made available to the researching company and this could be compiled and collated and sold as databases to various service providers as authentic locational details for tapping potential customers. Thus, maintaining anonymity and confidentiality of information shared is a professional norm that any ethical researcher should follow. In case the data is to be shared, it must be done with the consent of the respondent.
Professional objectivity

As a true researcher and contributor to the existing body of knowledge, the researcher must maintain the objectivity of an absolutely neutral reporter of facts. He must maintain objectivity in all phases of the study while:

- Designing the research objectives which must be based on facts and sound analysis rather than simple opinion.
- Collecting information by using a standard and not differential set of instructions. For example, in the intervention study quoted earlier, the researcher must give the instructions in the same way to both the experimental and control group and in no way try to exaggerate the actual impact of the treatment.
- Interpreting and presenting the findings as they are and not in a particular direction based on the researcher’s own gut feel or liking. For example, a researcher who is a consumer of organic food will attempt to exaggerate the health benefits of the products not because that is what was found but because as a consumer of the category, that is what he believes.

Thus, as stated earlier, just like any other business function a code of ethics for conducting research is well structured and laid out by almost every business association. At all times, the researcher must remember that besides aiding in business decision-making, research also contributes to the huge domain of management knowledge. Thus, an authentic, transparent and objective reporting and compilation of the research becomes that much more critical.

Check Your Progress

6. Name the different types of research on the basis of intent.
7. What is ethical consideration in research?
8. Mention some of the well-established and detailed ethical guidelines from international bodies.
9. What are some specific checks that a researcher look at?

1.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Research is basically a scientific and systematic search for pertinent information on a specific topic. It is an art of scientific investigation. As stated in the Advanced Learner’s Dictionary of Current English, ‘research is a careful investigation or enquiry, especially a thorough search for new facts in any branch of knowledge.’

2. Good research plays a vital role in the fields of economics, governance and business etc. Research also plays a significant role in the operational and
planning processes of business and industry as businesses apply logical and analytical techniques to business problems in order to maximize profits and minimize costs.

3. The two basic approaches to research are quantitative approach and qualitative approach.

4. Inferential, experimental and simulation approaches are the sub-classifications of quantitative approach.

5. Critical success factor analysis is one of the emerging areas in which HR research is being carried out. This is a participative form of management technique which was developed by American Organizational theorist John F. Rockart (1981) wherein the employees of an organization identify their critical success factor, and help in customizing and incorporating them in developing the mission and vision of their organization.

6. Research can be classified on the basis of intent as follows:
   (a) Pure research
   (b) Applied research
   (c) Exploratory research
   (d) Descriptive research
   (e) Diagnostic research

7. Ethical considerations are very much vital in the field of research. In the words of Rowley (2004), 'conducting research ethically is concerned with respecting privacy and confidentiality, and being transparent in the use of research data. Ethical practices hinge on respect and trust and approaches that seek to build rather than demolish relationships.'

8. The Social Research Association’s (SRA’s) ethical guidelines, the American Psychological Association (APA) code of ethics, code of standards and ethics for survey research designed by the Council of American Survey Research Organizations (CASRO), American Marketing Association (AMA) and Business Marketing Association (BMA) code of conduct and ethics.

9. A researcher is the most involved and key person responsible for the study, hence it is very important that the highest ethical standards maintained by him. Quality control and privacy control are the specific checks that a researcher look at in order to maintain the highest ethical standards.

**1.6 SUMMARY**

- Research in common parlance refers to search for knowledge. One can also define research as a scientific and systematic search for pertinent information on a specific topic.
• According to the *Advanced Learner’s Dictionary of Current English*, ‘research is a careful investigation or enquiry, especially a thorough search for new facts in any branch of knowledge.’

• Renowned author Clifford Woody defined research as an activity that comprises defining and redefining problems, formulating a hypothesis; collecting, organizing and evaluating data; making deductions and reaching conclusions; and carefully testing the conclusions to determine if they support the formulated hypothesis.

• The basic principles of research include a systematic process to identify a question or problem, set forth a plan of action to answer the question or resolve the problem, and meticulously collect and analyse data.

• The objective of any research is to find answers to questions through the application of scientific procedures. The main aim of any research is exploring the hidden or undiscovered truth.

• Research involves developing a scientific temperament and logical thinking.

• The significance of research-based answers can never be underestimated. The role of research is especially important in the fields of Economics, Business, Governance, and so on. Here research helps in finding solutions to problems encountered in real life.

• Research is also important as a career for those in the field of academics. It could be a career option for professionals who wish to undertake research to gain new insights and idea generation. Research also fosters creative thinking and new theorizations.

• Quantitative approach and qualitative approach are the two basic approaches to research.

• In quantitative approach, data is generated in quantitative or calculable form, and then that data is subjected to rigorous quantitative analysis in a rigid and formal fashion. Inferential, experimental and simulation approaches are the sub-classifications of quantitative approach.

• Qualitative approach to research is concerned with subjective assessment of attitudes, opinions and behaviour. Research in such a situation is a function of researcher’s insight and impressions.

• Research Methods refer to all the methods the researchers use while studying the research problems and while conducting research operations.

• Research methodology is a broader concept that includes not the research methods but also the logic behind the research methods in the context of a particular research study; and it explains the reasons for using particular research methods and statistical techniques. Research mythology also defines how the data should be evaluated to get the appropriate results.
Human Resources (HR) and organizational behaviour is an area which involves basic or fundamental research as a lot of academic, macro level research may be adapted and implemented by organizations into their policies and programmes.

Critical success factor analysis and employer branding are some emerging areas in which HR research is being carried out. The first is a participative form of management technique, developed by American Organizational theorist John F. Rockart (1981) in which the employees of an organization identify their critical success factors, and help in customizing and incorporating them in developing the mission and vision of their organization. Employer branding is another area which is being actively investigated as the customer perception (in this case, it is the internal customer, i.e., the employee) about the employer or the employing organization has a strong and direct impact on his intentions to stay or leave.

‘Scientific research is a systematic, controlled and critical investigation of propositions about various phenomena’, as stated by Fred Kerlinger (1986).

On the basis of intent, research can be classified as pure research, applied research, exploratory research, descriptive research, diagnostic research. However, the classification if research can also be done as per the methods of study such as fundamental research, applied research, historical research, formulative or exploratory research, experimental research, ex post facto research, case study research.

Ethical standards are extremely important no matter what be the field of study. This takes a special meaning in the conduction of research. Rowley (2004) has put it very simply as ‘conducting research ethically is concerned with respecting privacy and confidentiality, and being transparent in the use of research data. Ethical practices hinge on respect and trust and approaches that seek to build rather than demolish relationships.’

Research is also an exchange process between various people. The first of these is the one between the sponsoring client and the investigator. Thus both parties have an ethical obligation towards the other. In case the study is being conducted for a business client, complete transparency in terms of data gathering and interpreting is a must.

The most significant and important ethical concern of a research study is the issue of trust and confidentiality. At no cost must the researcher reveal any aspect of the study without the consent of the client.

The researcher might also need to guard the reason or purpose of the study. For example if the client wants to measure a new product potential, then revealing the reason for the study might lead to the concept or idea being adopted and converted into a product prototype by someone else before the client is out with the offering.
The researcher should not at any stage, either before or during the data collection stage, try to pressurize the respondent through persuasive influence or by forcing him to share information.

Professional creed, professional confidentiality, and professional objectivity are some basic tenets that the researcher must not forego. These are significant not only for the body of knowledge that the researcher is contributing to but also for the society in which we exist.

1.7 KEY WORDS

- **Exploratory Research**: It refers to a research used to investigate a problem which is not clearly defined.
- **Inferential approach**: It refers to a survey research where a sample of population is studied to determine its characteristics.
- **Hypothesis Testing**: It refers to a statistical test used to determine whether the hypothesis assumed for the sample of data stands true for the entire population or not.
- **Cross-functional research**: It refers to a research which is used to examine one variable in different groups that are similar in all other characteristics.
- **Ex post facto research**: It refers to a category of research design in which the investigation starts after the fact has occurred with interference from the researcher.

1.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. What are the objectives of research?
2. Define quantitative and qualitative approach to research.
3. What are the three groups under which research methods are categorized?
4. Write a short note on the techniques used in field research.
5. Briefly mention about applied research and ex post facto research.
6. What is client’s ethical code?

**Long Answer Questions**

1. Discuss the characteristics and purposes of research.
2. Examine the significance of research process.
3. Give a detailed account of the research topics involved in financial and accounting research.
4. Discuss the different types of research on the basis of intent and as per methods of study.

5. Describe the ethical codes related to respondents.

1.9 FURTHER READINGS


UNIT 2  LIBRARY AND INFORMATION SCIENCE (LIS)

Structure
2.0 Introduction
2.1 Objectives
2.2 Library and Information Science (LIS) as an Interdisciplinary Subject
2.3 Significance of Research in LIS
   2.3.1 Definition of Research
   2.3.2 Objectives of Research
2.4 Areas of Research in Library and Information Science
2.5 Answers to Check Your Progress Questions
2.6 Summary
2.7 Key Words
2.8 Self Assessment Questions and Exercises
2.9 Further Readings

2.0 INTRODUCTION
Library and Information Science (LIS) as the name implies is the combination of two fields, i.e., library and information science. In practice, the term LIS is occasionally used for an area that is not science (or research, a scholarship or an academic discipline). In suggestive words, the term library and information science research may be used to make focus entirely on research. In this unit, you will study about LIS as an interdisciplinary subject, the significance of research in LIS and the areas of research in Library and Information Science.

2.1 OBJECTIVES
After going through this unit, you will be able to:
- Examine LIS as an interdisciplinary subject
- Analyse the significance of research in LIS
- Identify the areas of research in Library and Information Science

2.2 LIBRARY AND INFORMATION SCIENCE (LIS) AS AN INTERDISCIPLINARY SUBJECT
The exploration in LIS in the past was principally considered to give a hypothetical establishment to proficient practice. The foundations of research in LIS seemed
by all accounts not to be not profound. Research in library science in the twentieth
century in the areas of LIS can be traced back to the Library School of the University
of Chicago, in mid-1920s.

The credit for the formal organization of the doctoral qualification programme
in library science in India goes unquestionably to mathematician and librarian Dr
S.R. Ranganathan (1892–1972). In 1951, he began giving library science instruction
at the University of Delhi, surmounting numerous challenges. The University of
Delhi granted the degree in library science in 1957 to D.B. Krishan Rao who
chipped away at ‘facet order for agriculture’. Doctoral research stayed in the
wild when Ranganathan shook the Delhi soil off his feet in 1955. During the 1970s,
a few doctorates on library-related subjects were earned by the library experts
under the direction and supervision of resources having a place with the orders,
for example, humanism, history, law, financial aspects, the executives and so forth.
The mantle of restoring and assisting doctoral research offices was expected by J.
S. Sharma (1924–1993), the then college curator and leader of the library science
bureau of the Panjab University, Chandigarh. Under his direction, the second by
law Ph.D. in library science was granted in 1977 after two decades. From that
point, there was no thinking back. Doctoral research got a fillip during the 1980s
and continuous enhancement in offices cleared routes for India to keep up its
Third World administration in library research and library writing. Ph.D. programmes
from there on mushroomed even in spite of the absence of offices or adherence to
principles.

2.3 SIGNIFICANCE OF RESEARCH IN LIS

Library and Information Science (LIS) is a very broad discipline, which uses a
wide range of constantly evolving research strategies and techniques.

Research involves ‘imaginative and orderly work attempted to expand the
load of information, including learning of people, culture and society, and the
utilization of this load of learning to devise new applications’. It is utilized to build
or affirm certainties, reaffirm the aftereffects of past work, tackle new or existing
issues, bolster hypotheses, or lead to new speculations.

An examination venture may likewise be an extension on past works in the
field. Research activities can be utilized to further enhance information on a subject,
or in the case of a school enquiry about venture, they can be utilized to facilitate an
understudy’s examination ability to set them up for future occupations or reports.
To test the legitimacy of instruments, methods, or investigations, research may
repeat components of earlier ventures or the undertaking in general. The main
roles of fundamental research (instead of connected research) are documentation,
disclosure, translation, or the innovative work (R&D) of techniques and
frameworks for the progression of human learning. Ways to deal with research
rely upon epistemologies, which shift significantly both inside and among humanities
and sciences.
In the early 1960s, Library and Information Science (LIS), was defined as a merger of librarianship and information science. LIS can be stated as a domain of both professional operations and scientific enquiry. As an area of application, it includes the profession of librarianship along with the number of other information professions, all of which presume the role of the following:

- Information content
- The community who interacts with the content
- Technology used to assist the progress of the entity development, exchanging data and information, depository, or conversion of the content.

LIS research aimed at highlighting the substantial information issues, such as those of ‘information retrieval, information quality and authenticity, policy for access and preservation, the health and security applications of data mining’. In the domain of LIS research, fields have also been motivated in its research introduction and the evolution of methodological tools, techniques and other theoretical perspectives and on the other side, it has a strong network with qualified professionals.

The corrective ground of LIS, which began in the 1920s, targeted at giving a theoretical base to library field. LIS has developed an imminent link with other areas in relation to the various subjects of research, especially computer science, communication studies and cognitive sciences.

In the late 1990s, the development of data-based librarianship has come up with some appreciative ways for research in LIS. The nature of research can best be described as combination of several branches of learning. It is being highly affected by research designs in fields like management sciences, social and behavioural context and to a minor range by the theoretical analysis chosen in the humanistic discipline. The usage of various approaches for deriving the information on research has been influenced from a special domain i.e., computer science. LIS has come up with its own developed methodological approaches, which can best be cited by example of bibliometrics. Bibliometrics can be defined as ‘the use of mathematical and statistical methods to study documents and patterns of publication’. It is a native research methodology commonly used especially in science studies.

The Role of Theory in LIS Research

In order to identify the presence factor of theory in LIS we need to understand the importance of research and integrity including the characteristics of various fields involved therein. According to Smiraglia, ‘theory does not exist in a vacuum but in a system that explains the domains of human actions, the phenomena found in these domains, and the ways in which they are affected’. He believed that theory emerged by methodically experiences, like in two ways; one is Positivism which is a philosophical theory stating that certain (“positive”) knowledge is based on natural phenomena and their properties and relations. Verified data (positive facts) received
from the senses are known as empirical evidence; thus, positivism is based on empiricism, and the other one is Hermeneutics. Hermeneutics is a wider discipline which includes written, verbal and non-verbal communication. Exegesis focuses primarily upon the word and grammar of texts. Hermeneutic, as a count noun in the singular, refers to some particular method of interpretation (see, in contrast, double hermeneutic).

A common fact is that theory is always used to plan interpretation in quantitative way and comes out with results in qualitative methods of research.

Some others described theory as ‘a law, a hypothesis, group of hypotheses, proposition, supposition, explanation, model, assumption, conjecture, construct, edifice, structure, opinion, speculation, belief, principle, rule, point of view, generalization, scheme, or idea’. According to Babbie, research is ‘a systematic explanation for the observed facts and laws that related to a particular aspect of life’. It is ‘a multiple level component of the research process, comprising a range of generalizations that move beyond a descriptive level to a more explanatory level’.

We can summarize the definition of theory as ‘a set of interrelated concepts, definitions, and propositions that explains or predicts events or situations by specifying relations among variables.’

According to the data and facts being collected by scholars, Glazier and Grover, they have come up with a model called ‘circuits of theory’ for theory building in LIS. The purpose of this model is to throw light on the technique of classification which has been described previously by some of the authors. The most important point is that research is influenced by psychological and critical social factors. This helps to explain the connection between the concepts of research, theory, paradigms, and phenomena. Phenomena can be defined as ‘events experienced in the empirical world’.

Now the question arises as to how the researchers conceptualize the link between the phenomena and form a hypotheses and questions related to research. So the answer to this question is simple. Researchers by taking the help of symbols (which means digital or iconic representations, usually by words or pictures) gives exact description to phenomena and symbols. It can be summarized as ‘In the taxonomy, empirical research begins with the formation of research questions to be answered about the concepts or hypotheses for testing the concepts within a narrow set of predetermined parameters’.

In Library and Information Science, a study has revealed that there are different stages of theories along with their results in terms of research conducted in LIS. They are as follows:

1. **Substantive Theory:** It is the first level theory which is defined as ‘a set of propositions which furnish an explanation for an applied area of inquiry’. This cannot be seen as a theory, in fact, it can be treated as a research hypothesis been tested or a subject for research finding.
2. **Formal Theory**: A second level described as ‘a set of propositions which furnish an explanation for a formal or conceptual area of inquiry that is a discipline’.

3. **Grand Theory**: It can be defined as ‘a set of theories or generalizations that transcend the borders of disciplines to explain relationships among phenomenon’.

4. **Paradigm**: The best definition is ‘a framework of basic assumptions with which perceptions are evaluated and relationships are delineated and applied to a discipline or profession’.

5. **World View**: It is the most significant theory described as ‘an individual’s accepted knowledge, including values and assumptions, which provide a “filter” for perception of all phenomena’.

The two theories, Substantive and Formal theories are being considered as “middle range” theories in social sciences domain. Hence, in order to distinguish between the two of them, it can be said that the way or method being used to structure generalizations and the potential for explanation and prediction are different from each other.

### 2.3.1 Definition of Research

As per the content available on social networking sites, research can be defined as a ‘detailed study of a subject in order to discover information or achieve a new understanding of it’. Till now no one has been able to present the broad definition of research in Library and Information Science (LIS). As per some of the agreed terminologies, research in LIS is done ‘to solve the professional obstacles, design tools and techniques which can further help in the analysis of organizations, services, and behaviour, to determine costs and benefits of our services, and most importantly, to establish or develop a body of theory on which to base our practice’.

In this modern era, one of the distinguish features is that our great scholars and scientists with respect to their different domains have shown their full cooperation and efforts to find the optimum solution to the biggest problem. The process of research is the oldest, i.e., coming from the ancient stages, which keeps on providing the solution to the issues in order to come up with the best scientific method. Now, the final picture depicts that by introducing the team research organization the level of uncertainty has come to the lowest; thus, also reducing the chances of failure.

The generalized outcome of these actions is being certified as principles, laws, or truths which makes a great impact on one’s understanding of his work and environment. Research is done in an organized way. It can be further defined as to find out more set of facts, or to create a strong bond between various facts and figures and through the formulation of a preliminary explanation or hypothesis which is subjected to an appropriate investigation for validation or disproof and ‘The men (researcher) of experiment are like the ant; they only collect and use: the
researchers resemble spiders, who make cobwebs out of their own substance. But the bee takes a middle course; it gathers its material from the flowers of the garden and of the field, but transforms and digests it by a power of its own’. Clearly, to be an effective research investigator one must resemble the bee-purposeful, industrious, and imaginatively selective in the assembling of evidence.

The research is always based on objectivity. It is a disciplinary course of action because it leaves no place for the subjective. In terms of the research process, the complexity of human beings pursued by them has to be overcome. Reasoning or observation that is diluted with emotion becomes sophistry or dogma. We should keep in mind that librarianship is always a service, as emotions can be a threat to the same. The duty of librarian is to perform well and make the best out of it.

2.3.2 Objectives of Research

The main aim of research is to find out the hidden facts and figures through the scientific method. In a certain way, research is utilized for obtaining the best answers to the questions. It always comes with the representation of data which has not been represented anywhere.

As we are aware that every research study has its own predetermined purpose. Hence, some of the common objectives are as follows:

(i) To decide the recurrence with which something happens or with something different. (It connects with diagnostic research study).
(ii) It depicts the traits of a particular individual, situation or a group in an accurate manner.
(iii) To pick up nature with a marvel or to accomplish new bits of knowledge into it. (It gives the example of exploratory or formulate research study).
(iv) To test the theory of a causal connection between factors (Hypothesis-testing research studies).

Research in Library

Research in librarianship quickly implies the accumulation and investigation of unique information on an issue of librarianship, done inside the library schools as indicated by logical and insightful gauges. Research in this association comprehensively incorporates examination; contemplation; overview; scholastic works; and in real life investigate by rehearsing custodians; data personnels; and document lists.

Importance of research in Library and Information Science has been stated as follows:

1. The fundamental motivation behind training is to prepare the students or individuals to gain exposure at the beginning phase of professional training.

An extensive preparation in the form of advanced learning leads to effective and productive specialization in the areas of scholastic librarianship, extraordinary librarianship, open libraries, library organization, book reference documentation and data and so forth.
2. An important significance of research is to be critical in nature. The profession requires students who can scrutinize, add to and enhance their learning and skill. It involves the basic function and basic ways to deal with librarianship. All the methods, forms, and acts of librarianship which are underestimated at the beginning, such courses ought to be put to critical examination. This kind of basic study empowers the students to accomplish and develop an aptitude for research. The learning acquired in the profession through surveys can be put to use and simple research on the issues of librarianship can be undertaken. This helps in serving the basic capacity not for analysis, but rather for the enhancement of expert practice.

3. It is the representative capacity. This is an aftereffect of developing polished skill in library and information science. Since 1930s, expansion of library profession is in progress and the profession started to see more noteworthy avenues of administration opening before it. The arrival of an era where we manage PCs, online administrations and complexities of data innovation—all these improvements in information technology have impacted library and information science. To adapt up to this ever-expanding multi-dimensional learning, we need to keep ourselves prepared and this requires research in the field of Library and Information Science.

2.4 AREAS OF RESEARCH IN LIBRARY AND INFORMATION SCIENCE

Library and Information Science focuses on how individuals or organizations search, use and gain access to information. Also, conditions are studied at the micro and macro level and include aspects of social strategy and/or informational policy.

Three main areas of research are the following:

(i) Information Use/Information Behaviour
(ii) Digitalization and Digital Libraries
(iii) Libraries, Cultural Policy and Information Policy

Research Methods in Library and Information Science

Let us study the various research methods applied in Library and Information Science briefly here.

1. Historical Method: We understand our current environment keeping the history of mankind in the backdrop. History is a written methodological record of mankind’s evolution and development and it paves the way for the present. History of library and information science is a continuous methodological recounting of the past events pertaining to the establishment, maintenance and utilization of systematically arranged collection of recorded
information. Hence, historical method is adopted in the research of this discipline. It also lays the foundation for improvement on the basis of the past study. It can also help to find out various shortcomings in this field on the basis of the past record.

2. Survey Method: A survey is a systematic collection of data concerning a system, its activities, operations, persons involved in the system. Thus, the library survey is defined as a systematic collection of data concerning a library—its activities, operations, personnel working in the library and its users. This objective of the survey method is to make a specialized type of investigation to improve library services. Hence, library surveys are conducted either to assess an existing situation or to check the library system or to evaluate the area of librarianship in order to removing the shortcomings.

Some of the characteristics of the survey method are the following:
(i) It is specifically worried about public activity such as what is being watched, depicted, gathered or a group of certainties about the present scenario.
(ii) It centres upon the given region or geological zone.
(iii) A huge amount of data is gathered from an expansive populace.
(iv) The data for the most part gathered through study is precise.

3. Case Study Method: The case study method is a technique in which an institution is recognized as a unit of study and various aspects to the unit are studied deeply. In this method, the emphasis is on principles and processes rather than the transfer of factual information. It represents the real situation drawn from practice and provides an opportunity to enquire skills in analysing problem, solving them and taking decisions. This method can be used in library management system analysis, cost benefit analysis, cost effectiveness, library effectiveness, library effectiveness and computer application in libraries.

4. Delphi Method: The Delphi method is basically a technique of obtaining consensus among experts on a given problem. A questionnaire is prepared highlighting the aims and objectives of research. The exact problem is put up to the panel of experts in many rounds till a consensus is achieved. The basic theory behind this technique is that consensus achieved will have greater credibility and authority than only the opinion of the most articulate spokesperson in a group of participating respondents. This method is used in the formulation of library legislation, policy-making for libraries, curriculum design, method of teaching and education, process of decision-making and manpower planning.

5. Statistical Method: The statistical method is used for aggregative analysis and intensive study of individual unit outside its scope. This method is based upon quantitative analysis. By using this method, the researcher can study the problem related to libraries such as budget estimates, library planning,
assessment of library services, evaluation of library services and library forecasting studies.

Some well-known examples of survey research that are applicable to library and information sciences are the following:

(i) **User Study or User Survey**: To decide the use of library assets, fulfillment of the customers and their need.

(ii) **Community Survey**: To know the normal usage of library by the populace being served by a library.

### Research Strategies in LIS

LIS is an exceptionally expansive order, which utilizes a wide scope of continually advancing examination procedures and methods. Different order plans are created to break down strategies utilized in LIS. In the ‘investigate record’ segment of the Journal of Education for Library and Information Science, Kim incorporated past classes and definitions and presented a rundown of research systems, including information gathering and examination strategies. The listing included four general research strategies:

(i) Theoretical/philosophical inquiry (development of conceptual models or frameworks).

(ii) Bibliographic research (descriptive studies of books and their properties as well as bibliographies of various kinds).

(iii) R&D (advancement of capacity and recovery frameworks, programming, interface and so forth.).

(iv) Tackling issues and realizing change in associations.

Methodologies are then bifurcated into quantitative and qualitative types. Primary classification incorporates distinct investigations, prescient/illustrative examinations, bibliometrics, content examination and so forth. Qualitative driven methodologies are viewed as the contextual investigation, true to life technique, authentic strategy, grounded hypothesis, ethnography, phenomenology, emblematic interactionism/semiotics, sociolinguistics/talk examination/ethnographic semantics/ethnography of correspondence and hermeneutics/interpretive interactionism.

Methodologies incorporated concentrate on specific publication years, geographic territories, journal titles, parts of LIS and explicit attributes, for example, subjects, creation and research strategies. The following are a few discoveries that are predictable among all examinations:

- Gathering information, data recovery and library and data benefit exercises are among the most widely recognized subjects examined.
- Descriptive research approaches are dependent on reviews and prevalence of surveys.
- Over the years, there has been a significant increase in the variety of research approaches used to investigate library issues.
Data investigation is generally restricted to distinct measurements, including frequencies, means and standard deviations.

**Data Collection and Analysis**

Strategies for information analysis refer to the procedures utilized by the researchers to investigate information and answer their issues or questions. In qualitative examination, information investigation includes distinguishing basic examples of information and understanding the implications of information. Information investigation for quantitative examination includes measurable examination and elucidation of figures and numbers. The variety of data analysis methods included the following classes:

- Descriptive statistics
- Inferential statistics
- Qualitative data analysis
- Experimental evaluation
- Other methods

---

**Check Your Progress**

1. Who is credited for creating the formal organization of the doctoral qualification programme in library science in India?
2. State any one aim of research in LIS.
3. Name the various stages of theories in terms of research conducted in LIS.

---

**2.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS**

1. Dr S.R. Ranganathan, mathematician and librarian, is credited for creating the formal organization of the doctoral qualification programme in library science in India.
2. LIS research aimed at highlighting the substantial information issues, such as those of ‘information retrieval, information quality and authenticity, policy for access and preservation, the health and security applications of data mining’.
3. The various stages of theories in terms of research conducted in LIS are the following:
   - Substantive Theory
   - Formal Theory
   - Grand Theory
   - Paradigm
   - World View
2.6 SUMMARY

- The exploration in LIS in the past was principally considered to give a hypothetical establishment to proficient practice.
- The credit for the formal organization of the doctoral qualification programme in library science in India goes unquestionably to mathematician and librarian Dr S.R. Ranganathan (1892–1972).
- Library and Information Science (LIS) is a very broad discipline, which uses a wide range of constantly evolving research strategies and techniques.
- In the early 1960s, Library and Information Science (LIS), was defined as a merger of librarianship and information science. LIS can be stated as a domain of both professional operations and scientific enquiry.
- In order to identify the presence factor of theory in LIS we need to understand the importance of research and integrity including the characteristics of various fields involved therein.
- The two theories, Substantive and Formal theories are being considered as “middle range” theories in social sciences domain.
- As per the content available on social networking sites, research can be defined as a ‘detailed study of a subject in order to discover information or achieve a new understanding of it’.
- The main aim of research is to find out the hidden facts and figures through the scientific method. In a certain way, research is being utilized for obtaining the best answers to the questions present.
- The main aim of research is to find out the hidden facts and figures through the scientific method. In a certain way, research is being utilized for obtaining the best answers to the questions present.
- Research in librarianship quickly implies the accumulation and investigation of unique information on an issue of librarianship, done inside the library schools as indicated by logical and insightful gauges.
- Library and information science focuses on how individuals or organizations search, use and gain access to information.
- A survey is a systematic collection of data concerning a system, its activities, operations, persons involved in the system.
- The case study method is a technique in which an institution is recognized as a unit of study and various aspects to the unit are studied deeply.
- The statistical method is used for aggregative analysis and intensive study of individual unit outside its scope. This method is based upon quantitative analysis.
- Strategies for information analysis refer to the procedures utilized by the researchers to investigate information and answers their issues or questions.
2.7 KEY WORDS

- **Epistemology**: It is a branch of philosophy that investigates the origin, nature, methods and limits of human knowledge.
- **Bibliometrics**: It is the application of statistical methods to the study of bibliographic data.
- **Hermeneutics**: It is the science of interpretation, especially of the scriptures.

2.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Write a short note on Library and Information Science (LIS) as an interdisciplinary subject.
2. What is the role of theory in LIS research?
3. Define research in your own words.
4. What are the research methods in Library and Information Science (LIS)?

**Long Answer Questions**

1. Analyse the significance of research in Library and Information Science (LIS).
2. Discuss the objectives of research.
3. Examine the areas of research in Library and Information Science (LIS).
4. Explain the research strategies in Library and Information Science (LIS).

2.9 FURTHER READINGS

UNIT 3 SCIENTIFIC METHOD

Structure
3.0 Introduction
3.1 Objectives
3.2 Understanding Scientific Method of Research
3.2.1 Nature of Research in Library and Information Science
3.3 Answers to Check Your Progress Questions
3.4 Summary
3.5 Key Words
3.6 Self Assessment Questions and Exercises
3.7 Further Readings

3.0 INTRODUCTION

In this unit, we will discuss the scientific method of research in reference to research in library and information sciences. The objective of any research is to find answers to questions through the application of scientific procedures. Thus, we can say that scientific research is a systematic and objective way of seeking answers to certain questions that require inquiry and insight or that have been raised on a particular topic. We have briefly spoken about the scientific method in the first unit. Here, we will explore the concept of scientific research as well as the nature of research in library and information science.

3.1 OBJECTIVES

After going through this unit, you will be able to:
- Discuss the scientific method of research
- Describe the nature of research in Library and Information Science

3.2 UNDERSTANDING SCIENTIFIC METHOD OF RESEARCH

The scientific method encourages doubts and criticism so that what emerges is the real evidence which has stood the test of reasoning. It makes science progressive as it is never too sure about its results. A unique characteristic possessed by this method is self-correction. A scientist does not believe any proposition without testing it. He has a number of built-in checks all along the way to enable him to adhere to the right path and arrive at the ‘truth’. Such checks are free from personal beliefs, attitudes and values.
Karl Pearson in his famous book *The Grammar of Science*, observes that there are three distinct characteristics of the scientific method: (a) careful and accurate classification of facts and observation of their correlation and sequence, (b) the discovery of scientific laws with the aid of the creative imagination, and (c) self-criticism and the final touchstone of equal validity for all normally constituted minds.

The first characteristic shows that the scientific method should enable us to classify facts accurately and carefully, free from the idiosyncrasies of the individual mind. In other words, there must be objectivity in this task. The second characteristic relates to the discovery of scientific laws with the help of the imagination. A mere collection of facts will not be sufficient to bring about scientific discoveries which will be the result of disciplined imagination and painstaking effort of the scientists. Finally, the third characteristic is that of self-criticism, i.e. that the scientist should critically examine his own research in a detached manner.

The author A. Wolfe expresses these characteristics in a different language. According to him, the common characteristics of the science are: (i) critical discrimination, (ii) generality and system, and (iii) empirical verification. Critical discrimination implies that one must not be influenced by mere appearance or prevalent notions but must try to get at the naked facts. Second, science is not interested in individuals but is concerned with generality of the system i.e., groups or classes of objects. Finally, science aims at the testing and verification of facts empirically so that they can be confirmed or rejected.

**Comparison of the Scientific Method and Non-Scientific Method**

Having looked into the different methods of knowing, we may now turn to a comparison of scientific and non-scientific methods.

<table>
<thead>
<tr>
<th>Scientific Method</th>
<th>Non-Scientific Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Objectivity of the Investigator</td>
<td>The scientific method is more objective because it attempts to eliminate the preconception or desires of the investigator from the results, making them unbiased.</td>
</tr>
<tr>
<td>The non-scientific method is more subjective because it bases its judgment on preconceived notions, authoritative statements or intuition, making the results more biased.</td>
<td></td>
</tr>
<tr>
<td>2. Systematic Procedure</td>
<td>The scientific method proceeds in a systematic manner with predesigned steps for this purpose.</td>
</tr>
<tr>
<td>The non-scientific method is carried out in a haphazard manner.</td>
<td></td>
</tr>
<tr>
<td>3. Accuracy of Measurement</td>
<td>The results in a scientific method can be measured more accurately with the help of electronic measuring devices which are available for most of the researches conducted scientifically. Most of the data is quantitative and hence measurable.</td>
</tr>
<tr>
<td>The results in a non-scientific method cannot be measured accurately because accurate measuring devices do not lend themselves owing to unsystematically conducted research and scattered or sporadic presence of data. Also most of the data are qualitative in nature.</td>
<td></td>
</tr>
</tbody>
</table>
4. Continuing and Exhaustive Nature of Investigation

Scientific method considers all the facts that are pertinent to the problem at hand. Evidence is found to support the existing conclusion repeatedly. The scientist is never sure that he has found the ultimate truth.

Non-scientific method does not consider all the facts with regard to problems at hand. It is not continuous like scientific method because the results are not scientific and same cannot be taken up for further researches.

5. Method of Analysis and Interpretation

A number of statistical techniques are applied to analyze the data.

The sophisticated statistical tools may not be applicable because of the sporadic nature of data and its qualitative nature.

One should examine whether marketing research satisfies these criteria, of being science, first, the marketing researcher is expected to be objective in his investigation. However, it is extremely difficult to remain completely objective during the entire research process, as one may be anxious at times to collect information to support and justify one's own position in regard to a certain issue.

Second, the marketing researcher is not so happily placed as the natural scientist in respect of the instruments of measurement. The latter can measure the minutest changes in his laboratory and is sure of the accuracy of his measurement. The marketing researcher has to deal with such aspects as the attitudes of consumers, changes in their preferences and their impact on the consumption of a given product. The instrument of measurement that is often used in such problems is the questionnaire, which is relatively crude and cannot give a very high degree of precision. It is used by several interviewers with varying backgrounds, training, experience and ability. As such, the information collected by them will have different degrees of accuracy.

The third criterion of the scientific method is that it is a continuous and unending process leading to the accumulation of systematic knowledge. Marketing research should ideally satisfy this criterion. However, as it is problem solving and problem-oriented research, the focus of investigation is narrow. The urgent nature of problems handled by marketing researchers, they seldom undertake exhaustive studies as is done by the natural scientists.

3.2.1 Nature of Research in Library and Information Science

A library is the treasury of books and documents which comprise and carry knowledge. A library is a vital part of academic life. One of the first step of the researcher’s is to go to the library to read books and other materials pertaining to his/her own field of interest to formulate a research problem. According to the author P.V. Young, "As a preliminary to field research, or in connection with it, a sustained and high-quality search for data in the library is a most pressing need in the social sciences." There are, therefore, implications about the increasing research activities for libraries.
The research workers visit libraries for their information needs. It is the responsibility of the librarian to give them the necessary information as quickly and exhaustively as possible. Thus, it is important for librarians to know in detail, their information requirements and hence to know reasons for conducting research, how they arrive at generalizations and generate new knowledge. This is why students of library and information science need to know the process of research. A student of library science who has such knowledge will be better equipped to serve the scientist and other research workers.

Scientific research in librarianship is a cautious process through which librarians can acquire more accurate knowledge and comprehension of libraries and librarianship. Knowledge about the different facets of library and information science can be attained by asking questions, thinking of possible answers and testing the possibilities by means of careful inquiry. In particular, a researcher in library and information science must have a clear understanding of the user’s behaviour, nature and demand of the clientele, size and nature of the library, types of collection it possess and above all the type storage and extent of continuous advancement taking place in information technology. The researcher in information science runs the risk of being out-dated if he/she is not dynamic enough to keep track of these fast developments. A more careful and concerned researcher on the other hand may bring in inestimable benefits to a developing country like India if he/she succeeds in laying hands on the most suitable technology for the taming of so many otherwise intractable problems of under development.

---

### Check Your Progress

1. What is the first characteristic of the scientific method according to Karl Pearson?
2. In conducting a research, what is the first step a researcher undertakes?

---

### 3.3 ANSWERS TO CHECK YOUR PROGRESS

1. According to Karl Pearson, the first characteristic of the scientific method shows that the scientific method should enable us to classify facts accurately and carefully, free from the idiosyncrasies of the individual mind.

2. One of the first step of the researcher’s is to go to the library to read books and other materials pertaining to his/her own field of interest to formulate a research problem.

---

### 3.4 SUMMARY

- The scientific method encourages doubts and criticism so that what emerges is the real evidence which has stood the test of reasoning.
• A unique characteristic possessed by this method is self-correction. A scientist does not believe any proposition without testing it. He has a number of built-in checks all along the way to enable him to adhere to the right path and arrive at the ‘truth’.

• The common characteristics of the science are: (i) critical discrimination, (ii) generality and system, and (iii) empirical verification.

• A library is the treasury of books and documents which comprise and carry knowledge. A library is a vital part of academic life.

• The research workers visit libraries for their information needs. It is the responsibility of the librarian to give them the necessary information as quickly and exhaustively as possible.

• Scientific research in librarianship is a cautious process through which librarians can acquire more accurate knowledge and comprehension of libraries and librarianship.

• The researcher in information science runs the risk of being out-dated if he/she is not dynamic enough to keep track of fast developments in the field.

3.5 KEY WORDS

• Empirical Verification: It refers to the use of empirical data, observation, test, or experiment to confirm the truth or rational justification of a hypothesis. Scientific beliefs must be evaluated and supported by empirical data.

• Marketing Research: It refers to the systematic collection, analysis, and interpretation of data pertaining to the marketing conditions.

3.6 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. Define scientific research.

2. Differentiate between scientific and non-scientific research.

Long Answer Questions

1. Discuss the characteristics of scientific research.

2. Describe the nature of research in library and information science.
3.7 FURTHER READINGS


UNIT 4  RESEARCH METHODS

Structure
4.0 Introduction
4.1 Objectives
4.2 Definition
4.3 Research Methods and its Advantages and Limitations
  4.3.1 Survey
  4.3.2 Case Study
  4.3.3 Census
  4.3.4 Experimental Research
  4.3.5 Focused Groups
4.4 Answers to Check Your Progress Questions
4.5 Summary
4.6 Key Words
4.7 Self Assessment Questions and Exercises
4.8 Further Readings

4.0 INTRODUCTION

Research methods are basically strategies, techniques, and processes which are applied in the collection of data for analysis so as to find out new information or create better understanding of a topic. Using research methods gives a better understanding to the area of study. Different research methods are used for different studies. Though the methods are different but these are closely related to each other. Choosing a research method depends on the research goal, on the subjects. While choosing one research method over other, one needs to consider factors like time for data collection and analysis of data.

In this unit, you will study about the definition and sources of research. The unit goes on discussing different research methods like survey, case study, experimental research, census, focus groups etc. Besides, you will also get to study about the merits and demerits of the respective research methods.

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Discuss the different definitions of research methods
- Understand the characteristics and features of different research methods like survey, experimental research, case study, focus groups, census etc.
- Describe the advantages and limitations of the research methods
4.2 DEFINITION

The term ‘research methods’ generally refers to the plan or approach that a researcher has in order to gather data or information.

Sources of Research

There are primary two sources of research. These are primary and secondary sources of research. Primary resources contain first-hand information, meaning that one is reading the author’s own account on a specific topic or event that they have participated in. Some examples of primary resources include scholarly research articles, books, and diaries. On the other hand, secondary sources describe, summarize, or discuss information or details originally presented in another source. Examples of secondary research include publications such as textbooks, magazine articles, book reviews, commentaries, encyclopaedias, almanacs, and so on.

4.3 RESEARCH METHODS AND ITS ADVANTAGES AND LIMITATIONS

A research method refers to the specific tools or ways which are used to collect and analyse data, such as survey, questionnaire, case study, interview etc. Let us study the various methods of research in detail.

4.3.1 Survey

Survey is an important tool in research. No research can be performed without a survey. Survey can be defined in various ways. Some of the common definitions of survey are as follows:

- To view with a scrutinizing eye; to examine
- To inspect, or take a view of; to view with attention, as from a high place; to overlook; as to stand on a hill and survey the surrounding country
- To determine the form, extent, position, and so on, of a tract of land, a coast, harbour, or the like, by means of linear and angular measurements, and the application of the principles of geometry and trigonometry; as to survey land or a coast
- A particular view; an examination, especially an official examination, of all the parts or particulars of a thing, with a design to ascertain the condition, quantity, or quality; as a survey of the stores of a ship; a survey of roads and bridges; a survey of buildings
- To examine with reference to condition, situation, value, and so on to examine and ascertain the state of; as to survey a building in order to determine its value and exposure to loss by fire
Types of Survey

There are basically two types of surveys:

(i) **Descriptive**: These surveys generally collect information on what people think and do.

(ii) **Analytic**: These surveys are generally used to either test hypotheses or to answer particular research questions.

While the most common method of collecting survey data is the ‘questionnaire’, the means by which you gather the information that goes into the survey responses may vary. If the survey makes use of a questionnaire, the measuring instruments must have demonstrability, reliability and validity, especially with regard to sampling, questioning and mode of questioning.

Some examples of collecting survey data include self-administered posted questionnaires, web-based forms, telephonic question and answer interviews or face-to-face interviews. There are advantages and disadvantages of each approach, primarily to do with sample size and open versus closed questions. In order to make a judgement, the key areas for consideration include the cost, co-ordination, size of the sample, rate of return, nature and quality of the data obtained, and the ability to clarify questions or responses. The success of using surveys depends strongly on the design of appropriate body of questions and the skill of the interviewer.

Other important methods of collecting survey data are Interview, Observation and Case Study. All these methods have been discussed in detail later.

Stages in Survey Method

Surveys go through the following seven stages:

(i) **Planning and designing the survey**

In this stage, you must define the goals and objectives of your survey. You should write down the outline of your research and also establish a budget for the project. You are also required to plan your schedule, define the population and estimate the required sample size. The method of data collection and the method for determination of the results should also be decided at this stage. Finally, you must write down the questions and design and pretest the questionnaire.

(ii) **Collecting data**

In this stage, first you have to decide on the survey method that will suit your research needs. There is no best method of collection of data and you must gather the required data keeping your resources in mind. You must also decide what steps to take in case sufficient data is not collected from the respondents.
(iii) **Accessing data**

The only purpose of this stage is to transfer the data into the analytical software for processing it further.

(iv) **Preparing and managing data**

The main aim of this stage is to get the data ready for analysis. In this step, you are required to formulate a ‘codebook’. This codebook must include variable names, variable formats and descriptive variable labels. You should also set up multiple item indices and scales, i.e., multiple variables that have exactly the same answer set. In this step, you should transform your data, which will help to get the data in the form and structure required for analysis. Also, the missing data values should be replaced with estimates so that better summary statistics are obtained.

(v) **Analysing data**

In this stage, you take out all the useful information that you require from the data that you have collected. This helps you make informed decisions.

(vi) **Reporting**

After analysing the data, the results need to be reported. The main aim of reporting is to produce results from the data analysis which can be easily understood by others who can use this information.

(vii) **Deployment**

You must tailor your results according to the needs of the target audience. This will ensure the effectiveness of the results.

### 4.3.2 Case Study

We explore and analyse the life of a social entity, whether it be a family, a person, an institution or a community, with the help of a case study. The purpose of case study method is to identify the factors and reasons that account for particular behaviour patterns of a sample chapter and its association with other social or environmental factors. Generally, social researchers use the case study method to understand the complex social phenomenon and to identify the factors related to this phenomenon. The case study provides the clues and ideas to a researcher for further research study. By adopting the case study method, a researcher gets to know about happenings in the past, which could be related to the research studies and analyse the problem with better perspectives.

#### Assumptions of the case study method

The assumptions made in a case study method are as follows:

- Case study depends on the imagination of the investigator who is analysing the case study. The investigator makes up his procedure as he goes along.
History related to the case is complete and as coherent as it could be.

It is advisable to supplement the case data by observational, statistical and historical data, since these provide standards for assessing the reliability and consistency of the case material.

Efforts should be made to ascertain the reliability of life history data by examining the internal consistency of the material.

A judicious combination of data collection techniques is a prerequisite for securing data that is culturally meaningful and scientifically significant.

Advantages and Disadvantages of Case Study Method

Case study ensures several advantages to the researcher for his research work. The key advantages of the case study method are as follows:

- Provides the basis for understanding complex social phenomenon and all related factors affecting the social phenomenon.
- Provides clues and ideas for exploratory research. When the researcher is not able to get a fair idea about the research, past happenings mentioned in a case study help the researcher get clues and ideas.
- Case study helps in generating objectives for exploratory research.
- It suggests the new courses of inquiry.
- Case study helps in formulating research hypothesis.

Some important disadvantages of case study method are as follows:

- Reliability: Data collected through case study may not be reliable or it can be difficult to verify the reliability of data in the current scenario.
- Adequacy: Data collected through case studies may not be adequate for research work as data is not pertinent to the research conditions.
- Representative: Data presented by case studies represents the happenings with unknown circumstances to a researcher. Hence, it cannot be the true representation of events to a researcher.

Making Case Study Effective

The criteria for evaluating the adequacy of case history is of central importance for case study. American psychologist and social scientist John Dollard has proposed seven criteria for evaluating such adequacy. They are as follows:

(i) The subject must be viewed as a specimen in the cultural series, i.e., the case drawn out from its total context for the purpose of study must be considered as a member of the particular cultural group or community. The scrutiny of life histories of people must be done with a view to identify community values, standards and their shared way of life.

(ii) The organic motto of action must be socially relevant, i.e., the action of individual cases must be viewed as a series of reactions to social stimuli or...
situations. In other words, the social meaning of behaviour must be taken into consideration.

(iii) The strategic role of the family group in transmitting the culture must be recognized, i.e., in case of the individual being the member of a family in shaping his behaviour must never be overlooked.

(iv) The specific method of elaboration of organic material onto social behaviour must be clearly shown, i.e., the case history that portrays in detail how basically a biological organism, the man, gradually blossoms forth into a social person, is especially fruitful.

(v) The continuous related character of experience for childhood through adulthood must be stressed. In other words, the life history must be a configuration depicting the interrelationships between the persons with various experiences.

(vi) The social situation must be carefully and continuously specified as a factor. One of the important criteria for the life history is that a person’s life must be shown as unfolding itself in the context of and partly owing to specific social situations.

(vii) The life history material itself must be organized according to some conceptual framework. This, in turn, would facilitate generalizations at a higher level.

Case study as a method of business research

A detailed case study helps the researcher identify the reasons behind business-related problems. As it can be possible that particular incident has happened in past, so the current issues can be sorted out by referring to the same case. In-depth analysis of selected cases is of particular value to business research when a complex set of variables may be at work in generating observed results and intensive study is needed to unravel the complexities. The exploratory investigator should have an active curiosity and willingness to deviate from the initial plan, when the finding suggests a new course of enquiry, which might prove more productive. With the help of case study method, the risk can be minimized in any decision-making process.

4.3.3 Census

Census Survey: Census survey means gathering pertinent information about all the units of population, viz., people, institutions, householders, and so on. As you know, population may consist of persons, institutions, objects, attributes, qualities, families among others. A population is a well defined group of many of these. For instance, the Census Survey of India, which takes place once in ten years, gathers benchmark data about each and every household of India. Since it concentrates on each and every household, it restricts its scope to certain surface level demographic data such as age, sex, income, education, lands possessed, cattle, nature of house, domestic facilities available, and so on. The studies are conducted...
through a quick survey in a stipulated period. However, coverage of units is very
exhaustive. The census survey as a method of research in education can be
employed to understand educational problems and make policy decisions.

**Strength of Census Survey:** The strength of the census survey is associated with
generalized characteristics of data. Description of population data acts as a major
source of identifying several pertinent issues and questions for research. It is very
useful in making a trend analysis of different events. Moreover, hard database
system of the entire population is very useful in the development of strategic planning
and policy-making of education at the micro level as well as at the macro level.

**Limitations of Census Survey:** As discussed, each and every unit of population
is covered under the census survey. However, data is gathered only under limited
headings. Also, this data is only surface level information. Through a census survey,
one can gather nominal data. Thus, the researcher cannot ask questions in depth.

Many times, such data is gathered mechanically where the investigators are
not well trained about cross examining the evidence at the field level. In such
cases, the probability of getting valid data is also minimized. Census surveys involve
employment of huge manpower and monetary resources. This method is also time
consuming. Getting each respondent to cooperate for data collection is very difficult.
Hence, the feasibility of conducting census studies is very limited. Moreover, because
of sample surveys, many questions can be well answered by saving time, money
and human resources; hence, one may look for census studies with limited focus
of research.

### 4.3.4 Experimental Research

Experimental research refers to the research activity wherein the manipulation
of variables takes place and the resultant effect on other variables is studied. It provides
a logical and structured basis for answering questions. The experimental researchers
manipulate the environment, stimuli or applications and observe the impact of this
manipulation on the condition or behaviour of the subject. The manipulation that they
undertake is deliberate and systematic.

Experimentation is the testing of hypotheses. Once the experimenters have
defined a situation or issue, they formulate a preliminary solution or hypothesis.
They then apply their observations of the controlled variable relationships in order to
test, and then confirm or reject the hypothesis.

Experimentation is the classic method of experimenting in a science laboratory
where elements are manipulated and effects observed can be controlled. It is the
most sophisticated, exacting and powerful method for discovering and developing
an organized body of knowledge.

According to J. W. Best, (Butler University, Emeritus), Experimental research
is the description and analysis of what will be or what will occur under carefully
controlled conditions.
Characteristics of Experimental Research

Experimental research is based on highly rigorous procedures, and aims at producing reliable and valid conclusions. By looking at the various designs and procedures used, one can formulate some essential characteristics of experimental research which distinguish it from other types of research methods like survey and historical.

- **Pre-Experimental Statistical Equivalence of Subjects in Different Groups:** This pre-condition is achieved by random selection and assignment of subjects to different groups. This procedure is essential to meet the threat of selection differences to the internal validity of the results.

- **Use of at least Two Groups or Conditions that can be Compared:** An experiment cannot be conducted with one group of subjects or one condition at a time. The intent of the experimenter is to compare the effect of one condition on one group with the effect of a different condition on another equivalent group. An experiment may take the shape of a comparison of the effect of one condition on a group of subjects and the effect of another condition on the same group.

- **Manipulation of the Independent Variable:** It is perhaps the most distinct feature of experimental research. Manipulation stands for the process of assignment of different values or magnitudes or conditions or levels of the independent variable to different groups.

- **Measurement of Dependent Variable in Quantifiable Form:** This distinguishes experimental research from descriptive, qualitative or analytical research.

- **Use of Inferential Statistics:** This is done to make probability statements about the results, and, thus, meet the requirements of imperfect measurements on which the behavioural sciences base their generalization.

- **Control of Extraneous Variables:** Though applicable to any other type of research, control of extraneous variables is the sine qua non of true experimental designs and the experimenter makes a determined effort to achieve it. It helps the experimenter to eliminate the possibility of any other plausible rival hypothesis claiming to explain the result.

Steps in Experimental Research

The steps in experimental research are as follows:

(i) **Survey of the Literature Relating to the Problem:** In experimentation, the researcher needs to acquire up-to-date information relating to the problem.

(ii) **Selection and Definition of the Problem:** It needs a rigorous logical analysis and definition of the problem in precise terms. The variables to be studied are defined in operational terms clearly and unambiguously. It helps the researcher to convert the problem into a hypothesis that can be verified or refuted by the experimental data.
(iii) **Statement of Hypotheses:** Hypotheses are the heart of experimental research. They suggest that an antecedent condition or phenomenon is related to the occurrence of another condition, phenomenon, event or effect. To test a hypothesis, the researcher attempts to control all the conditions except the independent variable. Therefore, he should give sufficient attention to the formulation of hypotheses. The experimental plant and statistical procedures help him in the testing of hypotheses and contribute little in the development of theories or advancement of knowledge. However, the hypotheses developed or derived from existing theories contribute to the development of new theories and knowledge.

(iv) **Construction of Experimental Plan:** Experimental plan refers to the conceptual framework within which the experiment is to be conducted. According to well-known theorist Van Dalen, an experimental plan represents all elements, conditions, phenomena and relations of consequences so as to:

- Identify the non-experimental variables.
- Identify the most appropriate research design.
- Identify a sample of subjects that will suitably represent the target population, form groups of these subjects and decide on the experiments which will be conducted on each group.
- Choose or develop an instrument that can be deployed to measure the results of the experiment.
- Lay out the data collection process and conduct a pilot study to test the instrument and the research design and state the hypotheses.

**Variables**

A **variable** is any feature or aspect of an event, function or process that, with its presence and nature, affects some other event or process which is being studied. According to Professor Fred N. Kerlinger, **variable is a property that takes on different value.**

**Types of Variables**

The various types of variables are as follows:

- **Independent Variables:** These are conditions or characteristics that are manipulated by the researcher in order to identify their relationship to observed phenomena. In the field of educational research, for instance, a specific teaching method or a variety of teaching material are types of independent variables.

  The two kinds of independent variables are as follows:

  (i) **Treatment Variables:** These are variables which can be manipulated by the researcher and to which he assigns subjects.

  (ii) **Organism or Attribute Variables:** These are factors, such as age, sex, race, religion, and so on, which cannot be manipulated.
• **Dependent Variables**: Dependent variables represent characteristics that alter, appear or vanish as a consequence of introduction, change or removal of independent variables. The dependent variable may be a test score or achievement of a student in a test, the number of errors or measured speed in performing a task.

• **Confounding Variables**: A confounding variable is one which is not the subject of the study but is statistically related with the independent variable. Hence, changes in the confounding variable track the changes in the independent variable. This creates a situation wherein subjects in a particular condition differ unintentionally from subjects in another condition. This is not a good result for the experiment which is attempting to create a situation wherein there is no difference between conditions other than the difference in the independent variable. This phenomenon enables us to conclude that the manipulation undertaken directly causes differences in the dependent variable. However, if there is another variable besides the independent variable that is also changing, then the confounding variable is the likely cause of the difference. An example of a common confounding variable is that when the researcher has not randomly assigned participants to groups, and some individual difference such as ability, confidence, shyness, height, looks, and so on, acts as a confounding variable. For instance, any experiment that involves both men and women is naturally afflicted with confounding variables, one of the most apparent being that males and females operate under diverse social environments. This should not be confused to mean that gender comparison studies have no value, or that other studies in which random assignment is not employed have no value; it only means that the researcher must apply more caution in interpreting the results and drawing conclusions.

Let us consider an instance wherein an educational psychologist is keen to measure how effective a new learning strategy that he has developed. He assigns students randomly to two groups and each of the students study materials on a specific topic for a defined time period. One group deploys the new strategy that the psychologist has developed, while the other uses any strategy that they prefer. Subsequently, each participant takes a test on the materials. One of the obvious confounding variables in this study would be advance knowledge of the topic of the study. This variable will affect the test results, no matter which strategy is used. Because of an extraneous variable of this nature, there will be a level of inconsistency within and between the groups. It would obviously be the preferred situation if all students had the exact same level of pre-knowledge. In any event, the experimenter, by randomly assigning the groups, has already taken an important step to ensure the likelihood that the extraneous variable will equivalently affect the two groups.

Let us imagine an experiment being undertaken to measure the effect that noise has on concentration. Assume that there are 50 subjects each in quiet and noisy environments. Table 4.1 below illustrates the ideal or perfect version of this experiment. ‘IV’ and ‘EV’ represent the independent variable and external variables, respectively. Note that (as shown in the table) the only difference between the two conditions is the IV, which indicates that the noise level varies from low to high in...
the two conditions. All the other variables are controlled and are exactly the same for the two conditions. Therefore, any difference in the concentration levels of subjects between the two conditions must have been caused by the independent variable.

### Table 4.1 Determining the Impact of Internal and External Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Quiet Condition N = 50</th>
<th>Noisy Condition N = 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise Level (IV)</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>IQ (EV)</td>
<td>Average</td>
<td>Average</td>
</tr>
<tr>
<td>Room temperature (EV)</td>
<td>68 degrees</td>
<td>82 degrees</td>
</tr>
<tr>
<td>Sex of subjects (EV)</td>
<td>60 per cent F</td>
<td>60 per cent F</td>
</tr>
<tr>
<td>Task difficulty (EV)</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Time of day (EV)</td>
<td>All different times between 9–5</td>
<td>All different times between 9–5</td>
</tr>
<tr>
<td>Etc. (EV)</td>
<td>Same as noisy environ</td>
<td>Same as quiet environ</td>
</tr>
<tr>
<td>Etc. (EV)</td>
<td>Same as noisy environ</td>
<td>Same as quiet environ</td>
</tr>
</tbody>
</table>

An Ideal Experiment

Now consider another version of this experiment wherein some of the other variables differ across conditions. These are confounding variables (highlighted below) and the experiment being conducted is not ideal. In this experiment, if the concentration levels of subjects vary between the two conditions, this may have been caused by the independent variable, but it could also have been caused by one or more of the confounding variables. For instance, if the subjects in the noisy environment have lower concentration levels, is it because it was louder, too hot or because they were tested in the afternoon? It is not possible to tell and, therefore, this is less than ideal.

### Table 4.1 Determining the Impact of Internal and External Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Quiet Condition</th>
<th>Noisy Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise Level (IV)</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>IQ (EV)</td>
<td>Average</td>
<td>Average</td>
</tr>
<tr>
<td>Room temperature (EV)</td>
<td>68 degrees</td>
<td>82 degrees</td>
</tr>
<tr>
<td>Sex of subjects (EV)</td>
<td>60 per cent F</td>
<td>60 per cent F</td>
</tr>
<tr>
<td>Task difficulty (EV)</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Time of day (EV)</td>
<td>Morning</td>
<td>Afternoon</td>
</tr>
<tr>
<td>Etc. (EV)</td>
<td>Same as noisy environ</td>
<td>Same as quiet environ</td>
</tr>
<tr>
<td>Etc. (EV)</td>
<td>Same as noisy environ</td>
<td>Same as quiet environ</td>
</tr>
</tbody>
</table>

A Non-Ideal Experiment

**Controlling the Confounding Variables**

There are ways by which the extraneous variables may be controlled to ensure that they do not become confounding variables. All people-related variables can be controlled through the process of random assignment which will most likely ensure that the subjects will be equally intelligent, outgoing, committed, and so on. Random assignment does not necessarily ensure that this is the case for every extraneous
variable in every experiment. However, when a sample is large, it works very well and the researcher’s motives for using this method will never be questioned.

One of the way in which situation variables or task variables can be controlled is basically by keeping them constant. For instance, in the noise-concentration experiment above, we could adjust the thermostat and, thereby, keep the room temperature constant and test all the subjects in the same room. We would, of course, hold the difficulty of the tasks constant by giving all subjects in both environments the same task. It is common practice for instructions to be written or recorded and presented to each subject in exactly the same way.

At times, the researcher cannot hold a situation or task variable constant. In these situations too, random assignment can be of great help. Consider a situation where the same room is not available for testing the two groups and, in fact, one group is tested on a Monday in Room 1 and the other group on a Tuesday in Room 2. In this situation, we can use random assignment which can result in half the Monday subjects in Condition A and the rest in Condition B, and the same for the Tuesday subjects. Hence, both conditions will have roughly the same percentage of subjects tested in Room 1 and 2. On the other hand, consider what would happen if we did not use random assignment and instead tested the Monday subjects in Condition A and the Tuesday subjects in Condition B. In this situation, we have two confounding variables. Subjects in Condition A were tested on different days of the week and in different rooms from those in Condition B. Any difference in the results could have been caused by one or more of the independent variable, the day of the week, or the room.

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

(a) **Intervening Variables**: In many types of behavioural research, the relationship between independent and dependent variables is not a simple one of stimulus to response. Certain variables that cannot be controlled or measured directly may have an important effect on the outcome. These modifying variables intervene between the cause and the effect. For example, in a classroom language experiment, a researcher is interested in determining the effect of immediate reinforcement on learning the parts of speech. He suspects that certain factors or variables other than the one being studied may be influencing the result, even though they cannot be observed directly. These factors may be anxiety, fatigue or motivation. These factors cannot be ignored. Rather, they must be controlled as much as possible through the use of appropriate design. For example, a variable (as memory) whose effect occurs between the treatment in a psychological experiment (as the presentation of a stimulus) and the outcome (as a response) is difficult to anticipate or is unanticipated, and may confuse the results.

(b) **Extraneous Variables**: These are variables that are not the subject of an experiment but may have an impact on the results. Hence, extraneous variables are uncontrolled and could significantly influence the results of a study. Often
we find that research conclusions need to be questioned further because of the influence of extraneous variables. For instance, a popular study was conducted to compare the effectiveness of three methods of social science teaching. Ongoing, regular classes were used, and the researchers were not able to randomize or control the key variables as teacher quality, enthusiasm or experience. Hence, the influence of these variables could be mistaken for that of an independent variable.

For instance, in a study which attempts to measure the effect of temperature in a classroom on students’ concentration levels, noise coming into the class through doors or windows can influence the results and is, therefore, an extraneous variable. This may be controlled by soundproofing the room, which illustrates how the extraneous variable may be controlled in order to eliminate its influence on the results of the test.

The following are the types of extraneous variables:

- **Subject variables** pertain specifically to the people being studied. These people’s characteristics, such as age, gender, health status, mood, background, and so on, are likely to affect their actions.
- **Experimental variables** pertain to the persons conducting the experiment. Factors, such as gender, racial bias or language influence how a person behaves.
- **Situational variables** represent the environmental factors which were prevalent at the time when the study or research was conducted. These include the temperature, humidity, lighting and the time of day, and could have a bearing on the outcome of the experiment.
- **Continuous variable** is one wherein any value is possible within the range of the limits of the variable. For instance, the variable ‘time taken to run the marathon’ is continuous since it could take 2 hours 30 minutes or 3 hours 15 minutes to run the marathon. On the other hand, the variable ‘number of days in a month that a worker came to office’ is not a continuous variable since it is not possible to come to office on 14.32 days.
- **Discrete variable** is one that does not take on all values within the limits of the variable. For instance, the response to a five-point rating scale must only have the specific values of 1, 2, 3, 4 or 5. It cannot have a decimal value such as 3.6. Similarly, this variable cannot be in the form of 1.3 persons.
- **Quantitative variable** is any variable that can be measured numerically or on a quantitative scale, at an ordinal, interval or ratio scale. For example, a person’s wages, the speed of a car or the person’s waist size are all quantitative variables.
- **Qualitative variables** are also known as categorical variables. These variables vary with no natural sense of ordering. They are, therefore, measured on the quality or characteristic. For example, eye colour (black, brown, or blue) is a qualitative variable, as are a person’s looks (pretty, handsome, ugly, and so on). Qualitative variables may be converted to
appear numeric, but this conversion is meaningless and of no real value (as in male = 1, female = 2).

**Experimental Designs**

The various experimental designs have been discussed in this section.

(a) **Single Group Design**: In this design study is carried out on a single group. Experiments can be conducted in the following ways:

   (i) **One-Shot Case Study**: This is a single group studied only once. A group is introduced to a treatment or condition and then observed for changes which are attributed to the treatment. This is like an ex post facto method in which on the basis of a dependent variable, an independent variable is looked for.

   (ii) **One Group before after Design**: This design entails the inclusion of a pre-test in order to establish base level scores. For instance, to use this design in a study of college performance, we could compare college grades prior to gaining the experience to the grades after completing a semester of work experience. In this design, we subtract the score of pre-test from post-test and see the differences. This difference is seen using a 't' test.

   (iii) **Time Series Designs**: Time series designs refer to the pre-testing and post-testing of one group of subjects at different intervals. In this design, continuous observation is carried out till a clear result is not seen. The purpose is to establish the long-term effects of treatment and can often lead to the number of pre- and post-tests varying from just one each to many. At times, there is a period of interruption between the tests so as to assess the strength of the treatment over a long time frame.

   (iv) **Counterbalanced Design**: Experiments that use counterbalanced design are effective ways to avoid the pitfalls of repeated measures, where the subjects are exposed to treatments one after the other.

Typically in an experiment, the order in which the treatments are administered can affect the behaviour of the subjects. It may also elicit a false response due to fatigue or any other external factors which may have a bearing on the behaviour of the subjects. To control or neutralize this, researchers use a counterbalanced design, which helps to reduce the adverse effects of the order of treatment or other factors on the results.

Counterbalancing helps to avoid confounding among variables. Take for example an experiment in which subjects are tested on both auditory reaction time task and visual reaction time task. If each and every subject were first tested on the auditory reaction time task and then on the visual reaction time task, the type of task and the order of presentation would be confounded. If the visual reaction time was lower, we would not be sure whether reaction time to a visual stimulus is ‘really’ faster to an auditory stimulus, as it is quite likely that the subjects would have learned something while performing the auditory task, which led to an improvement of their performance on the visual task.
(b) Two Equivalent Group Design

(i) Static Group Comparison Study: This design attempts to make up for the lack of a control group but falls short in relation to showing if a change has occurred. In this group, no treatment is given but only observation is carried out in a natural way of two groups, e.g., observation of the monkeys living in a city and observation of other monkeys living in the jungle. It is fair to mention here that in these groups nothing is manipulated as this design does not include any pre-testing and, therefore, any difference between the two groups prior to the study is unknown.

(ii) Post-Test Equivalent Groups Design: Randomization as well as the comparison of both the control and experimental group are used in studies of this nature. Each group is chosen and assigned randomly and presented with either the treatment or a type of control. Post-tests are subsequently administered to each subject to establish whether or not a difference exists between the two groups. While this is close to being the best possible method, it falls short on account of its lack of a pre-test measure. It is not possible to establish if the difference that seems to exist at the end of the study actually represents a change from the difference at the beginning of the study. Hence, while randomization mixes the subjects well, it does not necessarily create an equivalency between the two groups.

(iii) Pre-Test and Post-Test Equivalent Groups Design: This is the most effective as well as the most difficult method in terms of demonstrating cause and effect. The pre-test and post-test equivalent groups design ensures the presence of a control group as well as a measure of change. Importantly, it also adds a pre-test, thereby, assessing any differences that existed between the groups prior to the study taking place. In order to apply this method, we select students at random and then segregate them into one of two groups. We would subsequently evaluate the previous semester’s grades for each group in order to arrive at a mean grade point average. The treatment (work experience) would be applied to one group, whereas a control would be applied to the other. It is critical that the two groups should be treated similarly in order to control for variables, such as socialization, so that the control group may participate in an activity such as a softball league, while the other group participates in the work experience programme. The experiment ends at the end of the semester and the semester’s grades are compared. If it is found that the grade change for the experimental group was significantly different from the grade change of the control group, one could conclude that a semester of work experience results in a significant difference in grades when compared to a semester of non-work related activity programme.

(iv) Counterbalanced Randomized Two Groups Design: In this design, the group is divided in two parts on a random basis. This design is also called ‘rotation design’.
The simplest type of counterbalanced measure design is used when there are two possible conditions, A and B. As with the standard repeated measures design, the researchers want to test every subject for both conditions. They divide the subjects into two groups—one group is treated with condition A, followed by condition B, and the other is tested with condition B, followed by condition A, as shown in Figure 4.1.

![Fig. 4.1 Experiment to Show Counterbalanced Measure Design](image)

(c) **Solomon Four Group Design**: The sample is randomly divided into four groups. Two of the groups are experimental samples, whereas the other two groups experience no experimental manipulation of variables. Two groups receive a pre-test and a post-test. Two groups receive only a post-test. Table 4.2 shows the effect of a particular teaching method on the following groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test</th>
<th>Treatment</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>R</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>(b)</td>
<td>R</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>(c)</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>(d)</td>
<td>R</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 4.3 shows a teaching experiment using the Solomon design where testing before and without treatment have similar results, whilst results after teaching are significantly improved. This indicates that the treatment is effective and not subject to priming or learning effects.

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test</th>
<th>Treatment</th>
<th>Post-test</th>
<th>Pre-result</th>
<th>Post-result</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>R</td>
<td>No</td>
<td>No</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>(b)</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>(c)</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>(d)</td>
<td>R</td>
<td>Yes</td>
<td>No</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

**Internal and External Validity in Experimental Research**

**Internal validity** is considered as a property of scientific studies which indicates the extent to which an underlying conclusion based on a study is warranted. This type of warrant is constituted by the extent to which a study minimizes a systematic error or a ‘bias’. If a causal relation between two variables is properly demonstrated, then the inferences are said to possess internal validity. A fundamental inference may be based on a relation when the following three criteria are satisfied.
1. The ‘cause’ precedes the ‘effect’ in time (temporal precedence).
2. The ‘cause’ and the ‘effect’ are related (covariation).
3. There are no plausible alternative explanations for the observed covariation (non-spuriousness).

**Internal validity** refers to the ability of a research design for providing an adequate test of an hypothesis and the ability to rule out all plausible explanations for the results but the explanation being tested. For example, let us consider that a researcher decides that a particular medication prevents the development of heart disease because he found that research participants who took the medication developed lower rates of heart disease than those who never took the medication. This interpretation of the study’s results is likely to be correct, however, only if the study has high internal validity. In order to have high internal validity, the research design must have controlled the directionality and third-variable problems, as well as for the effects of other extraneous variables. In short, the researcher would have needed to perform an experimental study in which:

- Participants were randomly assigned to the experimental and control groups.
- Participants did not know whether they were taking the medication.

The most internally valid studies are experimental studies because they are better than correlational and case studies at controlling for the directionality and third-variable problems, as well as for the effects of other extraneous variables.

**Threats to Internal Validity**

The following are the various threats to internal validity:

- **Ambiguous Temporal Precedence**: Lack of precision about the occurrence of variable, i.e., which variable occurred first, may yield confusion that which variable is the cause and which is the effect.

- **Confounding**: Confounding is a major threat to the validity of fundamental inferences. Changes in the dependent variable may rather be attributed to the existence or variations in the degree of a third variable which is related to the manipulated variable. Rival hypotheses to the original fundamental inference hypothesis of the researcher may be developed where spurious relationships cannot be ruled out.

- **Selection Bias**: It refers to the problem that, at pre-test, differences between the existing groups may interact with the independent variable and, thus, be ‘responsible’ for the observed outcome. Researchers and participants bring to the experiment a myriad of characteristics, some learned and others inherent; for example, sex, weight, hair, eye and skin color, personality, mental capabilities and physical abilities, and so on. Attitudes like motivation or willingness to participate can also be involved. If an unequal number of test subjects have similar subject-related variables during the selection step of the research study, then there is a threat to the internal validity.
Repeated Testing: It is also referred to as testing effects. Repeatedly measuring or testing the participants may lead to bias. Participants of the testing may remember the correct answers or may be conditioned to know that they are being tested. Repeatedly performing the same or similar intelligence tests usually leads to score gains instead of concluding that the underlying skills have changed for good. This type of threat to internal validity provides good rival hypotheses.

Regression towards the Mean: When subjects are selected on the basis of extreme scores (one far away from the mean) during a test, then this type of threat occurs. For example, in a testing when children with the bad reading scores are selected for participating in a reading course, improvements in the reading at the end of the course might be due to regression toward the mean and not the course’s effectiveness actually. If the children had been tested again before the course started, they would likely have obtained better scores anyway.

External Validity

External Validity is considered as the validity of generalized (causal or fundamental) inferences in scientific studies. It is typically based on experiments as experimental validity. In other words, it is the degree to which the outcomes of a study can be generalized to other situations and people.

If inferences about cause and effect relationships which are based on a particular scientific study may be generalized from the unique and characteristics settings, procedures and participants to other populations and conditions, then they are said to possess external validity. Causal inferences possessing high degrees of external validity can reasonably be expected to apply:

- To the target population of the study, i.e., from which the sample was drawn. It is also referred to as population validity.
- To the universe of other populations, i.e., across time and space.

An experiment using human participants often employ small samples which are obtained from a single geographic location or with characteristic features is considered as the most common threat to external validity. Due to this reason, one cannot be certain that the conclusions drawn about cause and effect relationships do actually apply to people in other geographic locations or without these particular features.

External validity refers to the ability of a research design for providing outcomes that can be generalized to other situations, especially to real-life situations. For instance, if the researcher in the hypothetical heart disease medication study found that the medication, under controlled conditions, prevented the development of heart disease in research participants, he would want to generalize these findings to state that the medication will prevent heart disease in the general population. However, let us consider that the research design required the elimination of many potential participants, such as people who abuse alcohol or other drugs, suffer from diabetes, weigh more than average for their height, and have never suffered from a mood or anxiety disorder. These are common risk factors for heart disease and, by eliminating these factors, the outcomes of the study would provide little evidence...
that the medication will be effective for people with these risk factors. In other words, the study would have low external validity and, hence, its outcomes to the general population could not be generalized.

This commonly happens in tests of antidepressant medications. Because researchers want to make sure that the antidepressant effects of the medications being tested are not hidden by the effects of extraneous variables, they often have excluded potential participants with one or more of the following characteristics:

- People who are addicted to alcohol or illicit drugs
- People who take various medications
- People who have anxiety disorders (such as phobic disorders)
- People who suffer from depression with psychosis
- People with mild depression (because they would show only a small response to the medication)

If a study excluded people with these characteristic features, then most of the participants suffering from depression would be excluded from the final pool of participants. The outcomes of the study, therefore, would provide little information about how most depressed people will respond to the medication.

**Threats to External Validity**

A threat to external validity is an explanation of how you might be wrong in making a generalization. Usually, generalization is limited when the cause, i.e., independent variable depends on other factors; therefore, all threats to external validity interact with the independent variable.

- **Aptitude-Treatment Interaction:** The sample may have specific characteristic features that may interact with the independent variable, limiting generalization. For example, inferences based on comparative psychotherapy studies often employ specific samples (e.g., volunteers, highly depressed, no comorbidity). If psychotherapy is found effective for these sample patients, will it also be effective for non-volunteers or the mildly depressed or patients with concurrent other disorders?

- **Situation:** All situational features, such as treatment conditions, time, location, lighting, noise, treatment administration, investigator, timing, scope and extent of measurement, and so on, of a study potentially limit generalization.

- **Pre-Test Effects:** If cause and effect relationships can only be found when pre-tests are carried out, then this also limits the generality of the findings.

- **Post-Test Effects:** If cause and effect relationships can only be found when post-tests are carried out, then this also limits the generality of the findings.

- **Reactivity (Placebo, Novelty and Hawthorne Effects):** If cause and effect relationships are found, they might not be generalized to other situations if the effects found only occurred as an effect of studying the situation.

- **Rosenthal Effects:** Inferences about cause-consequence relationships may not be able to generalize to other investigators or researchers.
4.3.5 Focused Groups

Focus group as a method was developed in the 1940s in Columbia University by sociologist Robert Merton and his colleagues as part of a sociological technique. This was used as a method for measuring audience reaction to radio programmes. In fact, the method was uniquely adapted and modified in different branches of social sciences namely anthropology, sociology, psychology, education and advertising. It essentially emerged as an alternative method which was more cost effective and less time consuming, and could generate a large amount of information in a short time span. Another argument given in its favour was that group dynamics play a positive role in generating data that the individual would be hesitant about sharing when he was spoken to individually.

A focus group is a highly versatile and dynamic method of collecting information from a representative group of respondents. The process generally involves a moderator who manoeuvres the discussion on the topic under study. There are a group of carefully-selected respondents who are specifically invited and gathered at a neutral setting. The moderator initiates the discussion, and then the group carries it forward by holding a focused and an interactive discussion. The technique is extensively used and at the same time also criticized. While one school of thought places group dynamics at an important position, another negates its contribution as detrimental. We will examine these as we go along.

**Key elements of a focus group:** There are certain typical requirements for a conducive discussion. These need to be ensured in order to get meaningful and usable outputs from the technique.

**Size:** The size of the group is extremely critical and should not be too large or too small. Edward F. Fern, Professor Emeritus in the Pamplin College of Business at Virginia Tech, stated that every member is assumed to contribute meaningfully to the discussion; however, if the size of the group is too large, then contribution by the members might not be premium. Ideal recommended size, thus, for a group discussion is 8 to 12 members. Less than eight would not generate all the possible perspectives on the topic and the group dynamics required for a meaningful session.

**Nature:** Individuals who are from a similar background—in terms of demographic and psychographic traits—must be included; otherwise the disagreement might emerge as a result of other factors rather than the one under study. For example, a group of homemakers and working women discussing packaged food might not have a similar perspective towards the product because they have different roles to manage and balance; thus, what is perceived as convenience by one is viewed as indifferent and careless attitude towards one’s family by the other. The other requirement is that the respondents must be similar in terms of the subject/policy/product knowledge and experience with the product under study. Moreover, the participants should be carefully screened to meet a certain criteria.

**Acquaintance:** It has been found that knowing each other in a group discussion is disruptive and hampers the free flow of the discussion, and it is believed that
people reveal their perspectives more freely amongst strangers rather than friends. Terry Bristol, University of Arkansas at Little Rock, found that men revealed more about themselves amongst strangers, while females were more comfortable amongst acquaintances. Thus, it is recommended that the group should consist of strangers rather than subjects who know each other. There are exceptions, however, in certain cases; this would be further discussed in a subsequent section.

**Setting:** As far as possible, the external factors which might affect the nature of the discussion are to be minimized. One of these could be the space or setting in which the discussion takes place. Thus, it should be as neutral, informal and comfortable as possible. Even the ones that have one-way mirrors or cameras installed need to ensure that these gadgets are as unobtrusively placed as possible.

**Time period:** The conduction of the discussion should be held in a single setting unless there is a before and after design which requires group perceptions, initially before the study variable is introduced; and later in order to gauge the group’s reactions. The ideal duration of conduction should not exceed one and a half hour. This is usually preceded by a short rapport formation session between the moderator and the group members.

**The recording:** Earlier there were human recorders, either sitting behind one-way mirrors or in the discussion room. Today, these have been replaced by cameras that video-record the entire discussion. This can, then, be replayed for analysis and interpretation. The advantage over human recording is that one is able to observe the non-verbal cues and body language as well. This technology has been further enhanced and one can evaluate the discussion happening at one location, being observed and transmitted at another.

**The moderator:** He is the key conductor of the whole session. The nature, content and validity of the data collected are dependent to a large extent on the skills of the moderator. His role might be that of a participant where he might be a part of the group discussion or he might be a non-participant and has the task of rapport formation, initiating the discussion and steering the discussion forward. Well-known researchers Helen Morgan and Kerry Thomas have stated that any group task has two clear agendas. One is the conscious agenda to complete the overt task and the second, more important, plan is related to the unconscious. This is concerned with the emotional needs of the group and has been described differently as ‘group mind’, ‘group as a whole’ and ‘group as a group’. The moderator is clearly responsible for this as he needs to work with the group as a group in order to maximize the group performance. Thus, he needs to possess some critical moderating skills like:

- Ability to listen attentively and have a positive demeanour that encourages others to discuss. At the same time, he must be detached and give no indication about his personal opinion in order to skew the discussion. He should be dressed in a manner that is informal and similar to the group.
- He needs to make others feel comfortable; thus, the language used should be in the subjects’ lingo, with no use of technical words at all.
He needs to be flexible in approach, so that the discussion flows naturally rather than becoming compartmentalized into a question and answer session. At the same time, he also needs to act as a translator in case someone’s point is not understood or interpreted correctly.

He must also discreetly handle the overbearing and dominating participants and encourage all the members to contribute by drawing out the hesitant ones as well. Thus, sensitivity to the respondents’ feelings must be present at all times.

There is no external signal, so he needs to be sufficiently trained and acquainted with the topic to understand the specific interval when all the possible viewpoints get exhausted and the discussion needs to move on.

In conducting the discussions, he might use the summary and closure approach where he might pick up a similar point made by a participant to another and summarize it and ask for his opinion. Another tactic that can be used is to bring in the extreme opinions on the topic, in case no counter points are coming through; this, then, is able to generate more arguments into the discussion. Sometimes, rather than the moderator introducing another viewpoint, he might ask ‘is that all?’ This might sometimes trigger a fresh stance.

Steps in planning and conducting focus groups: The focus group conduction has to be handled in a structured and stepwise manner as stated below:

(i) Clearly define and enlist the research objectives of the research study that require qualitative research.

(ii) Then these objectives have to be split into information needs to be answered by the group. These may be bulleted as topics of interest or as broad questions to be answered by the group.

(iii) Next, a list of characteristics needs to be prepared, which would be used to select the respondent group. Based on this screening, a questionnaire is prepared to measure the demographic, psychographics, topic-related familiarity and knowledge. In case of a product or policy, one also needs to find out the experience and attitude towards it. Next, a comprehensive moderator’s outline for conducting the whole process needs to be charted out. Here, it is critical to involve the decision-maker (if any), the business researcher as well as the moderator. This is done so that there is complete clarity for the moderator in terms of the intention and potential applicability of the discussion output. This involves extensive discussions among the researcher, client and the moderator. Another advantage of having a structured guideline is that in case of multiple moderators, who might need to conduct focus group discussions at different locales, collection of similar information and reliability of the method can be maintained.

(iv) After this, the actual focus group discussion is carried out. Different sociologists have enlisted various stages that take place in a focus group. The most famous and comprehensive is the linear model of group
development formulated by Research theorist Tuckman. This has been adapted by well-known qualitative researcher Joanna Chrzanowska to explain the stages in the focus group discussions (Table 4.4).

(v) The focus group reveals rich and varied data; thus, the analysis cannot be quantitative or even in frequencies. The summary of the findings are chipped under different heads as indicated in the focus group objectives and reported in a narrative form. This may include expressions like ‘majority of the participants were of the view’ or ‘there was a considerable disagreement on this issue’. A summary report on the focus group discussion held in the organic food study is presented below along with the moderator guide.

**Focus group study: Potential consumers**

Two separate focus group discussions were conducted—one in Noida (UP) and the other in Hi-Tech City, Hyderabad. The group at Noida was predominantly of housewives and the one in Hi-Tech had professionals from different walks of life. Their opinion on a variety of subjects was sought. A summary of the discussions is presented below:

**Adulteration in food**

All the participants were unanimously concerned about adulterated food that they and their families were consuming. The discussion went from pesticides to chemicals and spurious food products. The ladies felt that they experienced a lot of health problems, specifically acidity, because of adulteration in the food. Some stated that they tried to grind all masalas at home as they felt that most of the problem was with masalas. However, some felt that this was meaningless as the whole masala was adulterated and contaminated by chemical residues. Thus, even though it was a matter of concern for them, they felt helpless to verbalize the possible solution.

There was one lady (Noida group), however, who felt that some of the problems were exaggerated and were basically created by the media and were plain hype. Another lady (HT group) felt that the problem of pollution was too deep-rooted and just adulterated food or food grown with chemical fertilizers and pesticides was too elementary and small to comprehend the problem of health hazards of the general population.

**Changes in lifestyle**

The consumers observed major changes in the recent years. The groups were unanimously of the opinion that they were more health conscious and concerned than their mothers and grandmothers. The younger generation (post-teens especially) are extremely conscious about the nutritional content of their food. They actively avoid excess sugar and fats in their diet. As a regime, people said that they exercise in some form or the other. Some said they drink more water and include healthy supplements like sprouts and olive oil in their diets.
Awareness of organic food products

Almost all the consumers, with the exception of one, had read or heard of organic food. One respondent had tried the product and found it very tasty. Three of the group members, as stated earlier, were sceptical about the benefits of organic food.

Willingness to try

The product was formally introduced to the groups and their reactions were noted to the same. Most of them, with the exception of two, were extremely enthusiastic about the products and wanted to know more about them, and had a number of queries about the availability, price, brands and benefits of the products.

Suggestions for marketing the product

- Divided opinion on who should sell the product. Some felt that a government-approved outlet like Mother Dairy/Trinetra should sell the products, whereas others felt that there should be exclusive organic food outlets. There were two or three people who felt that there should be no distinction and the products should be available everywhere. Some were also of the opinion that the products could be sold at high-end grocery stores or departmental stores since this was an expensive product. One consumer suggested the vegetable mandi also as a possible outlet; however, most of the others felt that the products would not be purchased by the masses.

- All the group members were unanimously of the opinion that they would buy a product only if it was certified as organic from an authentic and reputed body.

- The product should be vacuum packed, preferably in a brown paper packet with the label having the certification information and the source of the product clearly displayed.

- All felt that the price difference should not be too steep. At the same time, the Indian consumer who is buying a quality product accepts a price difference, so the product should be slightly expensive than the non-organic option.

- All the respondents felt that television was the best medium for promoting the product. All opined that there was a dire need for creating awareness. They felt that there was absolutely no visibility for the products, and more availability and awareness would mean more sales and more organically converted consumers. Some suggested popular soap operas and others were in favour of educational programmes.

- Some respondents felt that product promotions should be effectively and widely-conducted by tying up with environment-related organizations that would be willing to promote a healthy cause.
In terms of endorsement, they wanted sports personalities, film stars like Hema Malini, Simi Grewal, among others, and politicians like Menaka Gandhi and Sushma Swaraj endorsing the product; some even suggested common people who eat organic products and the farmer who produces.

The groups were generally of the opinion that the campaigns should be targeted at housewives and school children who would be wonderful and effective change agents.

Comparative advertising demonstrating the benefits of organic versus non-organic was another valuable suggestion discussed in the group. Some, however, argued for simply enlisting the benefits and resolving the myths about the products.

Price and availability and the reputation of the organization or brand would be important issues in marketing the product effectively.

Some punch lines suggested for the product were as follows:

- It is the future
- The healthy alternative
- Shudh and swachh
- Shuddhaaar
- Healthorganic
- Go organic

<table>
<thead>
<tr>
<th>Stage</th>
<th>Affective reactions</th>
<th>Behaviour patterns</th>
<th>Moderator role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forming</td>
<td>The group members are uncomfortable, insecure, a little</td>
<td>Silence or general talk, greetings</td>
<td>Time to bring clarity by explaining the purpose of</td>
</tr>
<tr>
<td></td>
<td>left and apprehensive.</td>
<td>and introductions. Mundane activity.</td>
<td>gathering together and the expected behaviour during</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>the discussion.</td>
</tr>
<tr>
<td>Storming</td>
<td>There is chaos, as emotions start flying with members</td>
<td>Arguments directed at each other</td>
<td>Do not take sides. Play poker face and say that all</td>
</tr>
<tr>
<td></td>
<td>questioning others and voicing their own opinion.</td>
<td>or trying to seek support from the</td>
<td>opinions are welcome. Save the direction to the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>moderator. Generally, there is</td>
<td>topic rather than arguments which might go off the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>rigidity in terms of sticking to</td>
<td>tangent. Try to steer the discussion.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>one's position. The leaders and the</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>followers emerge.</td>
<td></td>
</tr>
<tr>
<td>Norming</td>
<td>Cliques and sides start forming, based on the stand that</td>
<td>People have got the hang of the</td>
<td>Takes it easy, is more bothered about sequencing of</td>
</tr>
<tr>
<td></td>
<td>people have taken. More supportive and positive signals,</td>
<td>process and do not really need any</td>
<td>information and managing time now.</td>
</tr>
<tr>
<td></td>
<td>especially non-verbal.</td>
<td>steering by the moderator.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>People need to be more forthcoming</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>about sequencing of information and managing time now.</td>
<td></td>
</tr>
<tr>
<td>Performing</td>
<td>Individuals are subsumed to the group, roles are flexible</td>
<td>Sense of concentration and flow,</td>
<td>Time to introduce difficult issues, stimulate</td>
</tr>
<tr>
<td></td>
<td>and task-oriented.</td>
<td>everything seems easy, high energy,</td>
<td>material, projective techniques.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>group works without being asked.</td>
<td></td>
</tr>
<tr>
<td>Re-adjustment</td>
<td>There might be role reversals. People may have another</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>perspective with which the loosely defined cliques might</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>not agree, so one of the earlier stages might emerge.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mourning</td>
<td>Group task nearing completion, so there might be a sense</td>
<td>1 members do not feel that any</td>
<td>Signal conclusion. If you want to summarize, ask if</td>
</tr>
<tr>
<td></td>
<td>of loss as the energy generated with the discussion might</td>
<td>older stand is emerging, they might</td>
<td>any one has something to add. Thank everyone and</td>
</tr>
<tr>
<td></td>
<td>be tapped.</td>
<td>want to continue and not disrupt the</td>
<td>dispense for refreshment of drinks.</td>
</tr>
</tbody>
</table>

(Source: Chrzanowska, 2002)
Types of focus groups: As stated earlier, there could be several variations to the standard procedure. Some such innovations and alternative approaches are presented below:

Two-way focus group: Here, one respondent group sits and listens to the other and after learning from them or understanding the needs of the group, they carry out a discussion amongst themselves.

For example, in a management school, the faculty group could listen to the opinions and needs of the student group. Subsequently, a focus group of the faculty could be held to study the solutions or changes that they perceive which need to be carried out in the dissemination of the programme.

Dual-moderator group: Here, there are two different moderators: one responsible for the overt task of managing the group discussion and the other for the second objective of managing the “group mind” in order to maximize the group performance.

Fencing-moderator group: The two moderators take opposite sides on the topic being discussed and, thus, in the short time available, ensure that all possible perspectives are thoroughly explored.

Friendship groups: There are situations where the comfort level of the members needs to be high so that they elicit meaningful responses. This is especially the case when a supportive peer group encourages admission about the related organizations or people/issues. Eminent research thinker Stevens used the technique successfully when studying women groups for their experiential consumption of women magazines.

Mini-groups: These groups might be of a smaller size (usually four to six) and are usually expert groups/committees that on account of their composition are able to decisively contribute to the topic under study.

Creativity group: These are usually of longer than one and a half hour duration and might take the workshop mode. Here, the entire group is instructed which then brainstorm into smaller sub-groups and then reassembles to present their sub-groups opinion. They might also stretch across a day or two. A variation of the technique uses projective methods to extract alternative thinking.

Brand-obsessive group: These are special respondent sub-strata who are passionately involved with a brand or product category (say cars). They are selected as they can provide valuable insights that can be successfully incorporated into the brand’s marketing strategy.

Online focus group: This is a recent addition to the methodology and is extensively used today. Thus, it will be elaborated in detail. Like in the case of regular group process, the respondents are selected from an online list of people who have volunteered to participate in the discussion. They are then administered the screening
questionnaire to measure their suitability. Once they qualify, they are given a time, a participating id and password and the venue where they need to be so that they can be connected with the others. The group size here varies from four to six, as otherwise there might be technical problems and lack of clarity in the voices received. To ensure a standardized way of responding, the respondents are mailed details of how to use specific symbols to express emotions while typing the responses. Usual ways of denoting happiness and displeasure are used. These could also be coloured differently; also to show a higher degree of the emotion, additional faces may be used. There are other symbols as well. Besides, a brief about the purpose of the discussion and clarity on specific or technical terms is provided before the conduction. At the designated time, the group assembles in a web-based chat room and enters their id and password to log on. Here the chatting between the moderator and the participant is real time. Once the discussion is initiated, the group is on its own and chats amongst themselves, with the moderator playing the typical role. The session lasts for one to one and a half hour, and the process is much faster than a normal focus group.

The advantage of the method is that geographic locations are not a constraint and persons from varied locations can participate meaningfully in the discussion. Also, since it does not require a commitment to be physically assembled at a particular place and time, people who are busy and otherwise are not able to participate, can also be tapped. Since the addresses of the members are available to the moderators, it is also possible subsequently to probe deeper at a later date or seek clarifications. The interaction is faceless so the person interacting is completely assured of his/her anonymity and is thus, less inhibited. The method also has a cost advantage as compared to a traditional focus group. People are generally less inhibited in their responses and are more likely to fully express their thoughts. A lot of online focus groups go well past their allotted time since so many responses are expressed. Finally, as there is no travel, videotaping or facilities to arrange, the cost is much lower than for traditional focus groups. Firms are able to keep costs between one-fifth and one-half the cost of traditional focus groups.

However, the method can be actively and constructively used only with those who are computer savvy. Another disadvantage is that since anonymity is assured, actual authentication of the respondent being a part of the population under study might be a little difficult to establish. Thus, to verify the details, one may use the traditional telephone method and cross check the information. Since the person is typing his/her response, other sensory cues of tone, body language and facial expressions are not available. Thus, while the apparent emotions or attitudes can be tapped, however, the unconscious or sub-conscious cannot be judged.
These techniques have extensive use for companies that are into e-commerce. Most companies today have started using this technique to get employee reactions to various organizational issues, in what is termed as a 'virtual town hall meeting'. Thus, cyber dialogues can be carried out and meaningful feedback as well as population reaction can be measured with considerable ease and accuracy.

**Evaluating focus group as a method**

Focus groups are extensively criticized and yet have widespread usage in all areas of business research, to the extent that the technique is considered by some as synonymous with qualitative research. Before concluding the discussion on focus groups, let us examine the benefits and drawbacks of using the method.

**Idea generation:** As discussed earlier, the collective group mind creates an atmosphere where ideas and suggestions are churned out which are more holistic and significant than those that would be generated in an individual interview. The other advantage is that the group process works towards vetting each idea as it is presented. The dialogue between the members helps to refine and rephrase the perspective into a usable solution at the end of the discussion.

**Group dynamics:** Once the moderator has initiated the debate and some members have expressed their opinion, the atmosphere becomes charged and the respondents’ involvement with the topic increases with most members presenting reactions and counter reactions. The expressiveness becomes contagious and the contrived discussion slowly becomes a free-flowing discussion. As the comfort level of individuals with the other members increases, they start feeling at ease with the setting and expression becomes more open.

**Process advantage:** The discussion situation permits considerable flexibility in extracting the relevant information as the flow of topics and the extent to which the topic can be debated is dependent upon the group members and the emerging dynamics. Also, the situation permits a simultaneous conduction and collection of information from a number of individuals at a single point of time.

**Reliability and validity:** Since the objectives of the study have been listed out and the structure of the moderator outline is predetermined, the reliability of the information obtained is high. The mechanical recording of the data removes the element of human bias and error in the information collected.

However, the technique is not without shortcomings.

**Group dynamics:** The advantage could also be a disadvantage. On account of the group setting, the members might present a perspective not necessarily their own, but one that is along the lines of the group expression. This is the 'nodding dog syndrome', which is often a result of group conformity.
Scientific process: The group discussion must be treated as indicative and, thus, generalizing must be avoided. The answers obtained are varied and in a narrative form. Thus, coding and analysing this data is quite cumbersome.

Moderator/investigator bias: As discussed in earlier sections, the success or failure of the process depends, to a large extent, on the skills of the moderator. An unbiased and sensitive moderator who is able to generate meaningful and unbiased discussions is quite a rarity.

Check Your Progress
1. What do you mean by research methods?
2. What is survey and its importance?
3. What are the two types of survey?
4. How does case study help in research?
5. What do you mean by census survey?
6. Define variables.
7. What is internal and external validity in experimental research?
8. What are the threats to internal validity?
9. Who developed focus group and when?

4.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The term ‘research methods’ generally refers to the plan or approach that a researcher has used in order to gather data or information. To be more specific, research methods refer to the specific tools or ways which are used to collect and analyse data, such as survey, questionnaire, case study, interview etc.

2. Survey is an important tool in research. No research can be performed without a survey. It is an effective, reliable method of getting information or feedback from the individuals. Some examples of collecting survey data are questionnaires, web-based forms, telephonic questions and answers, interviews etc.

3. The two types of surveys are:
   - Descriptive: These surveys generally collect information on what people think and do.
   - Analytic: These surveys are generally used to either test hypotheses or to answer particular research questions.
4. Case study help to explore and analyse the life of a social entity. The case study method helps to identify the factors and reasons that account for particular behaviour patterns of a sample chapter and its association with other social or environmental factors.

5. Census survey means gathering pertinent information about all the units of population, viz., people, institutions, householders, and so on. Population may consist of persons, institutions, objects, attributes, qualities, families among others.

6. A variable is any feature or aspect of an event, function or process that, with its presence and nature, affects some other event or process which is being studied. According to Professor Fred N. Kerlinger, *Variable is a property that takes on different value.*

7. Internal validity refers to the ability of a research design for providing an adequate test of a hypothesis and the ability to rule out all plausible explanations for the results but the explanation being tested. While external validity refers to the ability of a research design for providing outcomes that can be generalized to other situations, especially to real-life situations.

8. The threats to internal validity are as follows:
   (a) Ambiguous temporal precedence
   (b) Confounding
   (c) Selection bias
   (d) Repeated testing
   (e) Regression towards the mean

9. Focus group as a method was developed in the 1940s in Columbia University by sociologist Robert Merton and his colleagues as part of a sociological technique.

4.5 SUMMARY

- The term ‘research methods’ generally refers to the plan or approach that a researcher use in order to gather data or information.
- There are primary two sources of research. These are primary and secondary sources of research. Primary resources contain first-hand information, meaning that one is reading the author’s own account on a specific topic or event that they have participated in. However, secondary sources describe, summarize, or discuss information or details originally presented in another source.
- Survey is an important tool in research. No research can be performed without a survey. Survey is to inspect, or take a view of; to view with
There are basically two types of surveys, descriptive surveys and analytic surveys. Descriptive surveys are generally collect information on what people think and do. Analytic surveys are generally used to either test hypotheses or to answer particular research questions.

- Some examples of collecting survey data include self-administered posted questionnaires, web-based forms, telephonic question and answer interviews or face-to-face interviews. There are advantages and disadvantages of each approach, primarily to do with sample size and open versus closed questions.

- There are seven stages in survey method namely, planning and designing the survey, collection of data, accessing data, preparing and managing data, analysing data, reporting, and deployment.

- The purpose of case study method is to identify the factors and reasons that account for particular behaviour patterns of a sample chapter and its association with other social or environmental factors.

- By adopting the case study method, a researcher gets to know about happenings in the past, which could be related to the research studies and analyse the problem with better perspectives.

- Census survey means gathering pertinent information about all the units of population, viz., people, institutions, householders, and so on. Census surveys involve employment of huge manpower and monetary resources. This method is also time consuming.

- Getting each respondent to cooperate for data collection is very difficult. Hence, the feasibility of conducting census studies is very limited.

- Experimental research refers to the research activity wherein the manipulation of variables takes place and the resultant effect on other variables is studied. It provides a logical and structured basis for answering questions.

- The experimental researchers manipulate the environment, stimuli or applications and observe the impact of this manipulation on the condition or behaviour of the subject. The manipulation that they undertake is deliberate and systematic.

- Experimentation is the classic method of experimenting in a science laboratory where elements are manipulated and effects observed can be controlled.

- A variable is any feature or aspect of an event, function or process that, with its presence and nature, affects some other event or process which is being studied.
• There are various types of variables namely independent variables, dependent variables, confounding variables.

• Confounding variables are of two types, namely intervening variables and extraneous variables.

• Internal validity is considered as a property of scientific studies which indicates the extent to which an underlying conclusion based on a study is warranted. This type of warrant is constituted by the extent to which a study minimizes a systematic error or a 'bias'.

• Internal validity refers to the ability of a research design for providing an adequate test of a hypothesis and the ability to rule out all plausible explanations for the results but the explanation being tested.

• External Validity is considered as the validity of generalized (causal or fundamental) inferences in scientific studies. It is typically based on experiments as experimental validity. In other words, it is the degree to which the outcomes of a study can be generalized to other situations and people.

• External validity refers to the ability of a research design for providing outcomes that can be generalized to other situations, especially to real-life situations.

• Focus group as a method was developed in the 1940s in Columbia University by sociologist Robert Merton and his colleagues as part of a sociological technique. This was used as a method for measuring audience reaction to radio programmes. In fact, the method was uniquely adapted and modified in different branches of social sciences namely anthropology, sociology, psychology, education and advertising.

• A focus group is a highly versatile and dynamic method of collecting information from a representative group of respondents.

• There are certain typical requirements for a conducive discussions. Some of the key elements of a focus group are size, nature, acquaintance, setting, time period, the recording, the moderator.

• There are five stages in a focus group discussion such as forming, storming, norming, performing, and mourning.

• There could be several variations to the standard procedure. Some such innovations and alternative approaches are as follows: two-way focus group, dual-moderator group, fencing-moderator group, friendship groups, mini-groups, creativity group, brand obsessive group, online focus group.
4.6 KEY WORDS

- **Survey**: It refers to examine and record the area and features so as to construct a map, plan, or description.
- **Descriptive Surveys**: It refers to a type of survey which generally collects information on what people think and do.
- **Analytic Surveys**: It refers to a type of survey which is generally used to either test hypotheses or to answer particular research questions.
- **Case Study**: It refers to a research method which involves an up-close, in-depth and detailed examination of a subject of study.
- **Census**: It refers to a count for official purposes, especially one to count the number of people living in a country and to collect information about them.

4.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Write a brief note on survey and its importance along with examples.
2. What are the assumptions made in a case study method?
3. Briefly discuss about the strengths and limitations of census survey.
4. Differentiate between independent and dependent variables.
5. What is a focus group and its purpose?

**Long Answer Questions**

1. Discuss the different stages of survey method in detail.
2. Explain the advantages and disadvantages of case study method.
3. Assess the characteristics of experimental research.
4. Analyse the various experimental designs.
5. What are the different types of focus groups? Discuss.
6. Describe the benefits and drawbacks of using focus group.

4.8 FURTHER READINGS


5.0 INTRODUCTION

The challenge for a business manager is not only to identify and define the decision problem; the bigger challenge is to convert the decision into a research problem that can lend itself to scientific enquiry. As Powers et al. (1985) have put it, ‘Potential research questions may occur to us on a regular basis, but the process of formulating them in a meaningful way is not at all an easy task.’

One needs to narrow down the decision problem and rephrase it into researchable terms. Well-known authors Bonnie L. Yegidis and Robert W. Weinbach (1991) have also referred to the complexity of phrasing the decision in research terms.

In this unit you will study about the research problem, formulation of the research problem, problem identification process. The unit will also discuss the conversion of management decision problem into research problem and objectives of formulating the problem and techniques involved in formulating the problem.

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Discuss the formulation of the research problem
- Describe the problem identification process
- Define the research problem
- Assess the objectives of formulating the problem
5.2 FORMULATING THE RESEARCH PROBLEM

The first and the most important step of the research process is to identify the path of enquiry in the form of a research problem. It is like the onset of a journey; in this instance, the research journey, and the identification of the problem gives an indication of the expected result being sought. A research problem can be defined as a gap or uncertainty in the decision-makers’ existing body of knowledge which inhibits efficient decision-making. Sometimes, it may so happen that there might be multiple reasons for these gaps, and identifying one of these and pursuing its solution might be the problem. According to professor Fred N. Kerlinger, a well respected positivist, ‘If one wants to solve a problem, one must generally know what the problem is. It can be said that a large part of the problem lies in knowing what one is trying to do.’ The defined research problem might be classified as simple or complex (Hicks, 1991). Simple problems are those that are easy to comprehend, and their components and identified relationships are linear and easy to understand, e.g., the relation between cigarette smoking and lung cancer. Complex problems, on the other hand, talks about interrelationship between antecedents and subsequently with the consequential component. Sometimes the relation might be further impacted by the moderating effect of external variables as well, e.g., the effect of job autonomy and organizational commitment on work exhaustion, at the same time considering the interacting (combined) effect of autonomy and commitment. This might be further different for males and females. These kinds of problems require a model or framework to be developed to define the research approach.

Thus, the significance of a clear and well-defined research problem cannot be overemphasized, as an ambiguous and general issue does not lend itself to scientific enquiry. Even though different researchers have their own methodology and perspective in formulating the research topic, a general framework which might assist in problem formulation is given below.

Problem identification process

The problem recognition process invariably starts with the decision-maker and some difficulty or decision dilemma that he/she might be facing. This is an action oriented problem that addresses the question of what the decision-maker should do. Sometimes, this might be related to actual and immediate difficulties faced by the manager (applied research) or gaps experienced in the existing body of
knowledge (basic research). The broad decision problem has to be narrowed down to information oriented problem which focuses on the data or information required to arrive at any meaningful conclusion. Given in Figure 5.1 is a set of decision problems and the subsequent research problems that might address them.

**Management decision problem**

The entire process explained above begins with the acknowledgement and identification of the difficulty encountered by the business manager/researcher. If the manager is skilled enough and the nature of the problem requires to be resolved by him or her alone, the problem identification process is handled by him or her else he or she outsources it to a researcher or a research agency. This step requires the author to carry out a problem appraisal, which would involve a comprehensive audit of the origin and symptoms of the diagnosed business problem. For illustration, let us take the first problem listed in the Figure 5.1. An organic farmer and trader in Uttarakhand, Nirmal farms, wants to sell his organic food products in the domestic Indian market. However, he is not aware if this is a viable business opportunity and since he does not have the expertise or time to undertake any research to aid in the formulation of the marketing strategy, he decides to outsource the study.

**Discussion with subject experts**

The next step involves getting the problem in the right perspective through discussions with industry and subject experts. These individuals are knowledgeable about the industry as well as the organization. They could be found both within and outside the company. The information on the current and probable scenario required is obtained with the assistance of a semi-structured interview. Thus, the researcher must have a predetermined set of questions related to the doubts experienced in problem formulation. It should be remembered that the purpose of the interview is simply to gain clarity on the problem area and not to arrive at any kind of conclusions or solutions to the problem. For example, for the organic food study, the researcher might decide to go to food experts in the Ministry for Food and Agriculture or agricultural economists, or retailers stocking health food as well as doctors and dieticians. This data, however, is not sufficient in most cases, while in other cases, accessibility to subject experts might be an extremely difficult task as they might not be available. The information should, in practice, be supplemented with secondary data in the form of theoretical as well as organizational facts.
Research is a systematic, objective and scientific study done to collect the research data related to current problems. Research enables a company to exploit the opportunities available in the environment. If the research has not been planned systematically, it is difficult for a firm to achieve the desired objectives. The research process can be described as follows:

- Defining the problem and objectives
- Developing data source
- Data collection
- Data analysis
- Presenting the findings

**Fig. 5.1 Converting Management Decision Problem into Research Problem**

Research is a systematic, objective and scientific study done to collect the research data related to current problems. Research enables a company to exploit the opportunities available in the environment. If the research has not been planned systematically, it is difficult for a firm to achieve the desired objectives. The research process can be described as follows:

- Defining the problem and objectives
- Developing data source
- Data collection
- Data analysis
- Presenting the findings

**Defining the Problem and Objectives**

The defined objectives should be SMART.

- **S** – Specific
- **M** – Measurable
Research Problem

A – Attainable
R – Realistic
T – Time Bound

The first step in research is definition of a problem. Selection of a problem is itself a difficult decision. The success of research depends on right selection of the problem. If the problem has not been identified in right manner, it is very difficult for the researcher to find the right solution of the issue.

The following sources can help a researcher identify the research problems:

- **Brainstorming:** A researcher can learn new dimensions of a problem by discussing ideas, thoughts, facts and data with other people who have knowledge of the subject.
- **Consultations:** By consulting others, the researcher identifies new dimensions of a problem.
- **Daily experience:** Daily experience develops the evaluative thinking in a researcher.
- **Academic experience:** Academic experience helps the researcher develop critical thinking towards the happenings.
- **Field situations:** Research is done because every field today is developing and, hence, changing constantly.

**Objectives of formulating the problem**

‘A well-defined problem is half solved.’ This statement reveals the fact that how important it is to formulate or define a problem. The primary objective of a research is to collect relevant data and analyse this data to get answers to the research problem. This means that the success of research depends upon accuracy of data and information required for investigation. Right formulation solves this purpose.

Proper definition of the problem, its analysis, identifying questions for data collection and the formulation of hypothesis to be tested are key steps which are required for formulation of the problem. Once the exact and accurate data is known to the researcher, he can plan the other steps without wastage of resources. Thus, right formulation of the problem gives the right direction to the entire research and limits the approach towards pertinent facts out of the large variety of facts. It helps us in determining statistical methods to be used for research.

**Criteria for formulation of the problem**

Criteria for formulating one problem out of identified problems can be grouped into internal criteria and external criteria. These are discussed in detail in the following section:
(i) Internal criteria

Internal criteria consist of the following:

- **Interest of the researcher**: The problem should be from the subject of interest of the researcher and can be challenging to him. Without interest in the problem, it becomes very difficult for the researcher to sustain continuity in the research. A researcher’s interest depends on his experience, educational background, sensitivity, and so on.

- **Own resources of the researcher**: Research requires a lot of money. If the researcher does not have enough money and he is unable to manage external finance, the researcher should not go in for research. Moreover, time resource is more important than money. Research requires more time and, hence, it should be utilized properly.

- **Competence of the researcher**: A mere interest in research is not enough. The researcher must be competent enough to plan and carry out a study of the problem. He should have sufficient knowledge of the subject matter, relevant methodology and statistical procedures.

(ii) External criteria

External criteria consist of the following:

- **Potential for research**: Very narrow or extremely vague problems should be avoided. In order to be researched, a problem must be one for which observation or other data collection in real world can provide the answer.

- **Importance and urgency**: Issues that require investigation are unlimited but available research efforts are very limited. Therefore, relative importance and significance of the problem is required. Important and urgent issues should be given priority over an unimportant one.

- **Novelty of the problem**: A problem on which a lot of research work has been done should not be considered for research as there are fewer chances of throwing light on any new factor.

- **Feasibility**: Novelty of the problem is not sufficient if it is not feasible to conduct the study on problem in real world, i.e., it should contain facts which can be analysed. Even if the problem is novel, we should make a small feasibility study first and proceed only after this if study allows.

- **Facilities**: A well-equipped library, proper guidance in data analysis, and so on, are basic facilities which are required to carry on any research.

- **Research personnel**: Availability of adequate research personnel including investigators and research officers is very important for data collection, which is a major issue in many developing countries like India.
Techniques involved in formulating the problem

Defining a research problem properly and clearly is a crucial part of the research study and must, in no case, be done hurriedly. The technique for this purpose involves undertaking the following steps:

- **Statement of the problem in a general way:** The problem should be carefully worded. The problem statement should indicate nature of the problem and intention of researcher.

- **Understanding the nature of the problem:** The best way to understand the nature of the problem is to discuss with those who have prior experience in the same kind of research. This will ensure that the origination of problem and the objectives in view are correct. If the marketer has stated the problem himself, he should consider all the facts that induced him to make a general statement concerning the problem.

- **Developing the ideas through discussions:** Many new ideas are developed by discussing them with others. This discussion provides useful information for research. Discussion is done with those people who have enough experience in the concerned field.

- **Rephrasing the research problem:** After going through the given four steps, the researcher gets a clear idea about the environment in which the problem is to be studied. Now rephrasing the problem into analytical or operational terms is not a difficult task. Through rephrasing, the researcher puts the research problem in as specific terms as possible so that it may become operationally viable and may help in the development of working hypothesis.

Check Your Progress

1. What is the most important step of the research process?
2. Define research problem.
3. What are the steps involved in the research process?
4. Name the sources that can help a researcher to find the right solution of the issue.
5. What is the primary objective of a research?

5.3 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The first and most important step of the research process is to identify the path of enquiry in the form of a research problem.
2. A research problem can be defined as a gap or uncertainty in the decision-makers’ existing body of knowledge which inhibits efficient decision making. Sometimes, it may so happen that there might be multiple reasons for these gaps, and identifying one of these and pursuing its solution might be the problem.

3. The steps involved in the research process can be described in five phases, namely defining the problem and objectives, developing data source, data collection, data analysis, and presentation of the findings.

4. The sources that can help the researcher to find the right solution of the particular issue are brainstorming, consultations, daily experience, academic experience, and field situations.

5. The primary objective of a research is to collect relevant data and analyse this data to get answers to the research problem. This means that the success of research depends upon accuracy of data and information required for investigation.

5.4 SUMMARY

- The first and the most important step of the research process is to identify the path of enquiry in the form of a research problem.
- A research problem can be defined as a gap or uncertainty in the decision-makers’ existing body of knowledge which inhibits efficient decision-making.
- Simple problems are those that are easy to comprehend, and their components and identified relationships are linear and easy to understand.
- Complex problems, on the other hand, talk about interrelationship between antecedents and subsequently with the consequential component.
- The problem recognition process invariably starts with the decision-maker and some difficulty or decision dilemma that he/she might be facing. This is an action-oriented problem that addresses the question of what the decision-maker should do.
- The entire process explained above begins with the acknowledgement and identification of the difficulty encountered by the business manager/researcher.
- Research is a systematic, objective and scientific study done to collect the research data related to current problems. Research enables a company to exploit the opportunities available in the environment. If the research has not been planned systematically, it is difficult for a firm to achieve the desired objectives.
- A few stages are involved in the research process, such as defining the problem and objectives, developing data source, data collection, data analysis, and presenting the finding.
The primary objective of a research is to collect relevant data and analyse this data to get answers to the research problem. This means that the success of research depends upon accuracy of data and information required for investigation.

The right formulation of the problem gives the right direction to the entire research and limits the approach towards pertinent facts out of the large variety of facts. It helps us in determining statistical methods to be used for research.

Criteria for formulating one problem out of identified problems can be grouped into internal criteria and external criteria. Internal criteria consist of interest of the researcher, own resources of the researcher, competence of the researcher.

External criteria consist of potential for research, importance and urgency, novelty of the problem, feasibility, facilities, and research personnel.

Defining a research problem properly and clearly is a crucial part of the research study. The technique involved in formulating the problem involve the steps like statement of the problem in a general way, understanding the nature of the problem, developing the ideas through discussions and, rephrasing the research problem.

5.5 KEY WORDS

- **Research Problem:** It refers to a definite or clear expression about an area of concern, a condition to be improved upon, a difficulty to be eliminated.
- **Applied Research:** It refers to the methodology used to solve a specific, practical problem of an individual or group.
- **Data collection:** It refers to the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research question hypotheses, and evaluate outcomes.

5.6 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. What are simple problems and complex problems?
2. What do you mean by criteria for formulation of the problem?
3. Define internal criteria and external criteria.
4. Write a short note on the techniques involved in formulating the problem.
Long Answer Questions

1. Discuss problem identification process.
2. Explain the process of converting management decision into research problem.
3. Analyse the components of internal criteria and external criteria.

5.7 FURTHER READINGS


UNIT 6   LITERATURE SEARCH

Structure
6.0 Introduction
6.1 Objectives
6.2 Importance of Surveying Related Literature
6.3 Uses of Library and Internet in Research
   6.3.1 Uses of Library in Research
   6.3.2 Uses of Internet in Research
6.4 Research Reviews, Abstracts and Microforms
6.5 Computerized Information Retrieval System
6.6 Answers to Check Your Progress Questions
6.7 Summary
6.8 Key Words
6.9 Self Assessment Questions and Exercises
6.10 Further Readings

6.0 INTRODUCTION

A research review is an article or a piece of information that summarizes the current understanding on a topic or previously published studies. If we talk about a literature review, it is also on the same lines of research review as it includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic.

In this unit, you will study about the features and characteristics of research review, and literature review. You will also study the uses of library and Internet in research and get acquainted with the various library research tools. Besides, you will also get familiar with the concepts like abstracts, index, microforms and would be able to define the computerised information retrieval systems and its types.

6.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the uses of library and Internet in research
- Identify the various library research tools
- Describe the variety of uses of library resources
- Comprehend the concept like research reviews, abstracts and microforms
- Discuss the computerized information retrieval system
6.2 IMPORTANCE OF SURVEYING RELATED LITERATURE

A literature review is a comprehensive compilation of the information obtained from published and unpublished sources of data in the specific area of interest to the researcher. This may include journals, newspapers, magazines, reports, government publications, and also computerized databases. The advantage of the survey is that it provides different perspectives and methodologies to be used to investigate the problem, as well as identify possible variables that may need to be investigated. Second, the survey might also uncover the fact that the research problem being considered has already been investigated and this might be useful in solving the decision dilemma. It also helps in narrowing the scope of the study into a manageable research problem that is relevant, significant and testable.

Once the data has been collected from different sources, the researcher must collate all information together in a cogent and logical manner instead of just listing the previous findings. This documentation must avoid plagiarism and ensure that the list of earlier studies is presented in the researcher’s own words. The logical and theoretical framework developed on the basis of past studies should be able to provide the foundation for the problem statement.

The reporting should cite clearly the author and the year of the study. There are several internationally accepted forms of citing references and quoting from published sources. The *Publication Manual of the American Psychological Association* (2001) and the *Chicago Manual of Style* (1993) are academically accepted as referencing styles in management.

To illustrate the significance of a literature review, given below is a small part of a literature review done on organic purchase.

Research indicates organic is better quality food. The pesticide residue in conventional food is almost three times the amount found in organic food. Baker *et al.* (2002) found that on an average, conventional food is more than five times likely to have chemical residue than organic samples. Pesticides toxicity has been found to have detrimental effects on infants, pregnant women and general public (National Research Council, 1993; Ma *et al.*, 2002; Guillete *et al.*, 1998).

Major factors that promote growth in organic market are consumer awareness of health, environmental issues and food scandals (Yossef and Willer, 2002).

This paragraph helps justify the relevance and importance of organic versus non-organic food products as well as identify variables that might contribute positively to the growth in consumption of organic products.
6.3 USES OF LIBRARY AND INTERNET IN RESEARCH

In this section, you will study the aspects of using library and internet in research.

6.3.1 Uses of Library in Research

It is common for a researcher to be confused and disoriented while using the library for research. Usually, this happens because the researchers feel at a loss as to where from and how to start searching the library resources. Therefore, a systematic and methodical approach towards the vast source of information that libraries usually offer is very essential. This can facilitate researchers in using quality time for conducting his/her search and collecting the essential information. Researchers should, therefore, create a concrete library research plan. Such a plan can enable him/her make an effective use of library materials for research.

Library Research Plan

A library research plan is a predefined activity that gives direction to your research. It is an act that involves evaluation that helps determine the subsequent activities to be followed by the researcher. As such, the research plan is a sequence of steps that the researcher should follow in order to get a comprehensible and reliable outline to adhere to. The various steps contained in a research plan can be stated as follows:

- **Subject evaluation**: This involves analysing the research subject with an informative perspective. The researcher should find out the extent of information that is already available and known to him/her. This can give a clear idea about the unknown information that needs to be searched in the library.

- **Determine the scope of research**: This involves identifying whether the research is a general study of occurrences or is concerned with more specific investigation. For example, whether your research is concerned with studying the eating habits of working women or eating habits of working people. Accordingly, you have to search for the relevant content. You should also check the chronological, geographical, political and other such aspects of your research study. You need to also analyse if your research deals with any specific locality, a particular time span or any current issue.

- **Sort-out keywords**: The research subject should be disintegrated into a set of key terms or key words. A key word can be defined as a term that expresses the most basic words of the research content, which describes your broad topic. The researcher should separate the distinct and unique words and important concepts contained in the research to use as subject
key words. This will greatly help the researcher keep in track the related topics of the research subject and avoid deviation. It is thus the most significant activity of a researcher wherein he/she should determine the basic terms to be adhered to while searching for books and other similar information resources.

- **Select the right library tool**: Depending upon the scope of your research, you can resort to the appropriate library tools required for collecting the information for research. There is a wide range of tools offered by a library in modern times. These tools span from small almanacs and handbooks to the most comprehensive books and anthological volumes, and most obviously the computerized library catalogue, which is the result of information technology revolution.

**Library Research Tools**

Once the researcher has executed the activities involved in the research plan, he/she should start looking for the essential and relevant information. This involves exploring information through traditional and modern library research tools, which contain specific bits of information as well as voluminous records and theories. To be able to make full use of these tools, it is necessary for the researcher to become familiar with the applications of these specific library tools. The most common library research tools available in any library are as follows:

- **Library catalogue**: A library catalogue is an informative list of resources and materials available in the library. It comprises the name of books or journals along with the name of authors, subjects and publishing houses. It thus, informs the researcher what is available in the library. Usually, in the developed countries, such catalogues are stored as computerized databases, which use featured searches with headings like ‘author,’ ‘title,’ ‘subject’ and ‘keywords.’ However, in India, many government-funded educational libraries are still using the paper-oriented catalogue technique except for such private libraries like the British Council Library, American Library, Indian Institute of Technology Library, etc. It is, therefore, advisable for a researcher to refine his/her research subject to specific key words and search for the necessary information using these key words.

- **Almanacs**: An almanac is a chronological tabular publication that is published annually. Traditional almanacs, usually, contained information regarding weather forecasts, astronomical data and several other statistics like the rising and setting of sun, moon, eclipses etc. However, in the current times, almanacs have become all comprehensive and include statistical and explanatory information regarding happenings in the whole world. Topical weather developments, historical events, factual information, etc., are the features of the present almanacs. A researcher can use these for quick grasping of facts of his research topic.
**Dictionaries**: Dictionaries are most often regarded as a superficial source of information as it is supposed to be performing the sole function of defining the meanings of words. However, contemporary editions of dictionaries are much innovative in their own way and explain innumerable terms in context of several usages of the specific term. As such, a researcher should seek factual information in the dictionary and also search for related phrases mentioned in the context of the word being searched to get a broader view of his/her topic of interest.

**Encyclopaedia**: Encyclopaedia, in general, is a bulk informative volume containing a synopsis of the concerned subject, which is published in alphabetical order. Encyclopaedia is, in fact, an extension of the concept of dictionary wherein the words are described precisely. A background context, however, can make the reader get more acquainted with the research term and, therefore, it is always preferable to consult an encyclopaedia to get a better knowledge of the subject matter. This also helps the researcher to understand several jargons and terminologies related to his/her subject of research.

**Bibliographies**: A bibliography usually comprises a list of reference materials mentioned by the researcher. This list gives the names of various sources that the researcher has resorted to, such as books and articles for research. The bibliographies are mentioned at the end of the article or research paper. It gives information regarding any particular topic that has been published together as a book. There are two types of bibliographies depending upon the information they provide. These two types can be stated and explained as follows:

- **Enumerative bibliography**: This is also known as compilative, reference or systematic bibliography. It gives a general idea of the relevant publications in a specific subject matter. The most common format to be used by the researcher while giving citations in such bibliography is as follows:
  - Author
  - Title
  - Publishing company
  - Publication date

- **Analytical bibliography**: This type is further classified into descriptive, historical and textual bibliographies. Usually, these are concerned with the physical attributes and contemporary importance of a book. They take into account the size, format and context in which the book was printed and published, etc. Therefore, such bibliographies are not very closely connected to any form of research.
Bibliographies, however, provide a good research tool to refer to for an effective exploring through the vast data available in a library.

- **Indexes**: Indexes, as is well known, are the alphabetical lists of authors’ name and subjects containing the relevant page numbers where these topics and authors have been discussed or described in the book. In research, they facilitate searching through a number of journals simultaneously and thus provide considerable information at the start of the research process. Indexes cover a large variety of sources, ranging from books, periodicals, conference papers, reports, thesis and articles, etc.

- **Search engines**: Search engines refer to software that browse through the Internet for the queried information and provide sites, which contain the concerned information within a few seconds. They operate automatically and collect words available on a vast number of web pages. A researcher needs to understand and learn the technique of effective utilization of a search engine. He/She should also be aware regarding the evaluation of the results that the search engines provide. The most popular search engines used all over the world today are Google and Alta Vista.

It is evident from the above discussion that with technological inventions, the nature of library research has changed tremendously. It has extended its scope and deals with vast range of information simultaneously. The researcher today, therefore, needs to familiarize himself/herself with various jargons and terminologies to get a good grasp of the research tools. The description that follows includes the various words commonly used in library research.

**Use of Library Resources**

There are innumerable information materials available in a library. In general, the prime sources of information can be classified and explained in the context of research usability as follows:

- **Books**: Undoubtedly, books are the most significant and chief resource of information in any library. However, irrelevance can greatly hamper a book’s usability for the researcher. It is, therefore, very essential for a researcher to examine whether the book he/she is referring is relevant to his/her research study or not. A researcher should also check for the authority of the book, i.e., whether the author is an expert in the field or a well-known publishing house has published the book. Checking for the contemporariness will also facilitate the researcher to remain up-to-date in conducting his/her research study.

- **Journals**: Journals refer to a periodical collection of articles. They also include such sources like reports, bulletins or proceedings, published monthly by any organization or an institution. There are also certain scholarly journals, which are of great help to a researcher working in the field of humanities or
social sciences. Usually, journals contain articles, which can offer thorough knowledge or up-to-date information regarding the subject matter of research.

- **Thesis**: In simple terms, thesis refers to a piece of research work conducted by an individual to qualify for a degree. Usually, libraries often hold copies of thesis, which have been submitted and approved for Ph.D. A researcher can browse through such thesis to get an idea of the layout and presentation options for his/her own thesis. While making use of such thesis, the researcher should remember that they are copyright protected materials and for quoting from a thesis, it is essential to take a prior written consent from the concerned author.

- **Manuscripts and archives**: These are unpublished, sometimes handwritten or typed original and primary sources of information. A study of these manuscripts gives the researcher an idea of the original study conducted in the past and thus guides him/her in conducting the research. Manuscripts can be used for citing examples to prove his/her points by referring to past occurrences and events. It is, however, absolutely important to preserve the integrity of such sources and, therefore, a researcher should repeat the original text exactly if required without any omissions, additions or corrections.

It is absolutely important for the researcher to check the relevance of the concerned materials while referring anything from the mentioned sources in his/her research study. Usually, there is a general tendency to simply use what comes first at hand. Secondly, one is bound to believe anything when it is in print. However, a good research is not the result of such quick and simple adaptations. A researcher should give plenty of time to conduct the research study. He/She should examine every available source of information to its fullest usable degree.

It is also necessary to check the accuracy, timeliness and depth of the content in the concerned source of information. Check out if the topic is being covered exhaustively by the book, or thesis or journal to which you have resorted. It is always advisable for the researcher to consider the target audience he/she is going to address and scrutinize. For example, a research that aims at conducting a study among the students of postgraduate courses should use refined and superior level of vocabulary. In that case, resorting to a book meant for high school students should be restricted till the extent of selecting basic ideas required for the research. However, it is obviously the researcher’s job to adapt his/her target audience and present the ideas in his/her own unique style.

The researcher, while conducting a research, collects a large amount of data and stores them in the computer. However, it depends on the hardware and software capacity of the computer to store the information and data. The hardware and software capacity can be changed or managed as per the requirements of the organization.
6.3.2 Uses of Internet in Research

Before the evolution of the Internet, conducting a research work involved a set of encyclopaedias and a trip to library. However, now we live in an age where the information is easily accessible via computer using Internet. Today, information and data can easily be accessed with the help of the Internet. The Internet is the fastest developing and the largest repository of data. A researcher on the Internet can find information about any topic he/she desires. The Internet acts as a huge database of the content where a researcher can access an unlimited number of informative sources.

Research itself is a very wide term. It means a systematic enquiry of the facts. There are various common applications of Internet research. One such application of the Internet research includes the personal research that is undertaken in order to enquire about a particular subject such as news or health problems. Various other applications of the Internet research also include research undertaken by the students for academic projects and papers, and writers and journalists researching stories.

One of the advantages of conducting research using the Internet is that hundreds or thousands of pages can be found with some relation to the topic, within seconds, which is not possible if the same topic is to be searched from books or encyclopaedias. Moreover, the Internet also includes e-mail, online discussion forums and other communication facilities such as instant messaging and newsgroups that help the researchers have a direct access to the experts and other individuals with relevant knowledge and interests.

There are various tools such as Internet search engine and Internet guide that a researcher can use for collecting the information. A search engine is an online database of Internet resources. When the researcher poses a query about a particular topic, the search engine looks for the likely matches within the database and displays the relevant content accordingly. Unlike, a standard search engine, the information that is contained within an Internet guide is compiled and organized by the humans, not computer programmes. Encyclopaedia Britannica is an example of Internet guide that covers a vast category of different topics.

However, there is one disadvantage for the researchers in conducting a research with the help of the Internet. The disadvantage is that the majority of the content available on the Internet is self-submitted and there are few rules and regulations that a researcher has to adhere with regard to what a researcher can publish and what he/she cannot. Moreover, the content on the Internet may sometimes be inaccurate and opinion based.

However, the Internet must not be disregarded as the major source of conducting research. It is one of the major sources of journals, books, general information and other relevant content. Therefore, we can say that the Internet is
a very important source for the researchers in this modern age for the purpose of collecting information.

<table>
<thead>
<tr>
<th>Check Your Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. What is a literature review?</td>
</tr>
<tr>
<td>2. Name the academically accepted referencing styles in management.</td>
</tr>
<tr>
<td>3. Define library research plan.</td>
</tr>
<tr>
<td>4. What are the most common library research tools available in a library?</td>
</tr>
<tr>
<td>5. What are journals?</td>
</tr>
<tr>
<td>6. What is the importance of thesis in research?</td>
</tr>
<tr>
<td>7. What is the one of the main advantages of conducting research using the Internet?</td>
</tr>
</tbody>
</table>

6.4 RESEARCH REVIEWS, ABSTRACTS AND MICROFORMS

Research reviews, abstracts and microforms are the tools that specify the significant points of published literature and the various research papers, published or unpublished.

Review of any literature or research paper provides the current knowledge including prominent findings, as well as theoretical and methodological contributions to a particular topic. A literature review is a search and evaluation of the available literature in the given subject or chosen topic area. It documents the state of the art with respect to the subject or topic you are writing about.

The purpose of any review paper is to compactly evaluate or review the recent progress for the specific topics. Overall, the review paper summarizes the state of data of the subject. It helps in understanding of the subject for the reader by discussing the findings given in the recent analysed papers.

A literature review may be an intellectual paper, which has this data as well as substantive findings. Literature reviews are secondary sources, and do not report new or original experimental work. Therefore, the literature review helps in explaining the topics of the research as well as in building a rationale or foundation for the research problem that has to be analysed for making decisions and the prerequisite for additional or further research.

A literature analysis or review might accommodate merely an outline or summarized framework of significant sources, however if referred to the social sciences, a literature analysis or review typically has associated structured and organized pattern or arrangement that combines each outline and synthesis, usually at intervals that specify the abstract classes within the framework of precise theoretical or conceptual classifications. An outline may be a recap of
the necessary and significant information about the sources or causes; however, a synthesis may be a re-organization or a re-shuffle of that precise information during an approach that informs or specifies or aims to investigate a research expression or a problem.

The literature analysis or review has the following analytical and methodological features:

1. Providing a new or different interpretation of previous researched material or by mixing the new interpretations with the previously given interpretations.
2. Tracing the intellectual or scholarly progression or advancement of the field under study, which includes major discussion and debates.
3. Depending on the condition the research sources are evaluated and recommend the researcher on the utmost pertinent or appropriate and significant research analysis.
4. Typically, within the conclusion or interpretation of a literature analysis and review for establishing and identifying those research problems wherever gaps exist.

Following are the key objectives of a literature analysis and review in research:

- Organize every context and its related contributions such that it helps to understand and analyse the research problem that is under study.
- Define the relationship or association between every work with the other research work that has been considered for the research process.
- Identify and recognize the new and different techniques for interpreting the previous research analysis.
- Reveal and ascertain any gaps that are existing within the research literature.
- Identify and resolve the disputes/collisions/conflicts between apparently contradictory or inconsistent studies conducted previously.
- Identify/recognize the regions of previous researches and analysis to avoid duplication if any.
- Defining the approach to fulfil the requirement for further research and analysis.
- Discover from the existing research literature to analyse the research framework within the standard or established perspective. This is very significant.

Information in any field has three different layers, as such:

- Researchers conduct and publish the primary studies.
- The reviews of these studies summarize and provide new interpretations specifically designed and occasionally extending on the remote side of the primary studies.
There are perceptions, conclusions, opinion, and interpretations that are shared informally which become a part of the traditional knowledge of field.

Types of Literature Reviews

Argumentative Review

This form examines literature by selection so as to support or refute associate degree argument and the aim is to develop a body of literature that establishes an investor viewpoint. Given the value-laden nature of some science analysis, contentious approaches to analyzing the literature is a legitimate and necessary kind of discourse.

Integrative Review

The body of literature includes all studies that address connected or identical hypotheses and analysis issues. A cooked integrative review meets the identical standards as primary analysis in respect to clarity, rigor, and replication. This is often the foremost common style of review within the social sciences.

Historical Review

Historical literature reviews specialize in examining analysis throughout the amount of your time, typically beginning with the primary time a problem, concept, theory, phenomena emerged within the literature, then tracing its evolution at intervals the scholarship of a discipline. The aim is to put analysis during a historical context to indicate familiarity with progressive developments and to spot the seemingly directions for future research.

Methodological Review

This approach helps highlight moral problems that you ought to remember of and contemplate as you undergo your own study. Reviewing ways of study provides a framework of understanding at completely different levels (i.e. those of theory, substantive fields, analysis approaches, and knowledge assortment and analysis techniques), however researchers draw upon a good style of information starting from the abstract level to sensible documents to be used in munition within the areas of metaphysics and philosophy thought, quantitative and qualitative integration, sampling, interviewing, knowledge assortment, and knowledge analysis.

Theoretical Review

The theoretical literature analysis and review helps in determining that which theories already exist, the relationships between them, and up to what degree or extent the contemporary theories are being investigated, and to develop new specific hypotheses which must be tested. Typically, these types of theories are
essentially employed for assisting and establishing the insufficient or inadequate theories that are applicable to reveal or discover that current or existing theories are not sufficient and satisfactory for explaining new or different increasing analysis issues.

Systematic Review
Systematic reviews formulate analysis queries that are broad or slender in scope, and determine and synthesize studies that directly relate to the systematic review question. They are designed to produce an entire, thoroughgoing outline of current proof relevant to a look question. Systematic reviews of randomised controlled trials are key to the evidence-based medication, and a review of existing studies is usually faster and cheaper than embarking on a replacement study.

Characteristics
1. It aims to produce a whole, thorough outline of current literature relevant to an exploration question. The primary step in conducting a scientific review is to make a structured question to guide the review.
2. The second step is to perform an intensive search of the literature for relevant papers.
3. The methodology section of a scientific review can list all of the databases and citation indexes that were searched like the Internet of Science, EMBASE, and PubMed/MEDLINE and anyone journals that were searched.
4. The titles associate degree abstracts of known articles are checked against pre-determined criteria for eligibility and relevancy to make an inclusion set.
5. This set can relate back to the analysis drawback. Every enclosed study could also be allotted associate degree objective assessment of method quality ideally by exploitation ways orthodox to the well-liked coverage things for Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

Stages
The main stages of systematic review are as follows:
1. Defining a research problem and then agreeing the objective techniques.
2. A search for relevant information from analysis of research that matches bound criteria. As an example, solely choosing analysis that is of sensible quality and then answering the outlined questions. Contacting a trained data skilled or professional person will improve the standard of the systematic review.
3. ‘Extraction’ of relevant information. This could embrace, however, the analysis performed often known as the tactic or ‘intervention’. Further, any
agency participated within the analysis (including what percentage people), however, it was absolutely bought for example funding sources and the outcomes.

4. Assess the standard of the info by judgment it against criteria known at the primary stage.

5. Analyse and mix the information that offers an overall result from all of the data. This mixture of knowledge is visualized employing a blobbogram (also known as a forest plot).

A careful literature review generally comprises of 10 to 25 pages but it could be longer depending on the process as such the process of reviewing the literature requires different kinds of activities and ways of thinking. Shields and Rangarajan (2013) and Granello (2001) link the activities of doing a literature review with Benjamin Bloom’s revised taxonomy of the cognitive domain (ways of thinking: remembering, understanding, applying, analyzing, evaluating, and creating).

A literature review has following four major purposes and uses.

- For surveys of literature in the chosen area of study.
- To synthesize the information in the literature into a summary.
- To critically analyse the information gathered by identifying gaps in current knowledge; by showing limitations of theories and points of view; and by formulating areas for further research and reviewing areas of controversy.
- To present the literature in an organized manner.

A literature review indicates the readers that the writer has an in-depth grasp of the given subject; and that he understands where the current topic of research fits into and adds to an existing body of agreed knowledge.

**Scoping Reviews**

Scoping reviews are distinct from systematic reviews. A scoping review is an endeavour to look for ideas, mapping the language that surrounds those and adjusting the search methodology iteratively. A scoping review could usually be a preliminary stage before a scientific review, that 'scopes' out a region of inquiry and maps the language and key ideas. As it could be a quite review that ought to be consistently conducted (the methodology is repeatable), some research publishers reason them as a sort of 'systematic review', which can cause confusion.

Critical evaluation of each literature review should include the following:

1. Provenance—Author’s credentials are to be monitored, i.e., all the arguments of the author’s should be supported by proper evidences, such as facts, figures, case studies, statistics, recent findings, etc.
Literature Search

NOTES

2. Methodology – Is research problem has been identified properly, the techniques used to gather information and the sample size chosen. Results have been accurately interpreted.

3. Objectivity – Certain data is used or some accurate information being ignored to validate the author’s point of view.

4. Persuasiveness – The highest most convincing or the least point for the author’s work.

5. Value – The work done has come to some significant conclusion.

The following are various stages for the development of Literature review:

1. To formulate the problem.
2. To search the material related to the subject.
3. Evaluation of the data in the right way.
4. To analyze and interpret the findings and conclusions.

The significant consideration is that all the assumptions can be predetermined and essentially be addressed concisely and adequately while performing the literature review and for collecting the problem-specific information. Sometimes when the researcher is operationally defining the variables for literature review, then he/she may have to restructure questions and consecutively the approach will become more authentic and trustworthy.

Thus, the specific framework of literature review specifically created to seek answers to the research question solely depend on the techniques of research analysis methods for obtaining the logically authentic information required to answer the research problem within the specified framework.

In addition, at times the researcher may come across situation when the topic of a research is such that there is no previous information or any literature available on it. In such cases, the researcher can contact the experts to take help who might be able to provide some valuable insights based upon their experience in the field or with the concept. This methodology of obtaining precise and authentic informations from the significant and erudite individuals is termed as the expert information analysis. This methodology might be formal and structured as it is extremely useful when authenticated by a secondary/primary research analysis.

Abstract

An abstract is a brief summary of a literature, research article, thesis, reviews, conference proceeding, or any in-depth analysis of a particular subject and is often used to help the reader quickly ascertain the paper’s purpose. To elaborate further an abstract is a self-contained, short, and powerful statement that describes a larger content work. Components vary according to subject. An abstract of a social science or scientific work may contain the scope, purpose, results, and contents of the work. Whereas an abstract of a humanities work may contain the thesis, background, and conclusion of the larger work. An abstract is not a review,
nor does it evaluate the work being abstracted. While it contains key words found in the larger work, the abstract is an original document rather than an excerpted passage.

As per law, abstract can be defined as a statement that consists of important points of a legal document or several legal papers related to it.

To present the complete synopsis of the paper an abstract need to be written within a tight word limit. A maximum word count for writing an abstract is 200 words. Nowadays, scientific papers are usually placed onto an info, with strict limits on the quantity of words.

While writing an abstract some important points to be kept in mind are:

1. A brief explanation of the topic of the subject
2. Defining the research question
3. The type of methodology used
4. Results
5. Conclusions

An abstract is generally kept at the beginning of a manuscript or typescript, acting as the point-of-thing for any given academic paper or literature. Abstracting and indexing services for various academic books are aimed at compiling the body of literature for that particular subject.

Abstract is also named as synopsis in many a publication just for change but the purpose is same. In some of the management reports, an executive summary usually contains more information (and often more sensitive information) than the abstract does.

Abstract: Purpose and Limitations

Abstract is used in many academic literatures just for clear communication about the complex research. An abstract may act even as a stand-alone entity in place of a full paper. As such, an abstract is used by many organizations as the basis for selecting research that is proposed for presentation in the form of a poster, platform/oral presentation or workshop presentation at an academic conference or seminar.

It will be appropriate to say that most literature database search engines index only the abstracts rather than providing the entire text of the paper. Full texts of scientific papers must often be purchased because of copyright and/or publisher fees and therefore the abstract is a significant selling point for the reprint or electronic form of the full text.

Although the abstract can be used just to convey the main results and conclusions of a scientific article but in any case the full text article must be consulted for details of the methodology, the full experimental results, and a critical discussion of the interpretations and conclusions.
An abstract may provide one to sieve or filter through copious numbers of papers for ones in which the researcher can have more confidence that they will be relevant to their research work. Once papers are chosen based on the abstract, they must be read carefully to be evaluated for relevance. It is generally agreed that one must not base reference citations on the abstract alone, but the content of an entire paper.

As per the results of a study published in some informative literature, the ‘exaggerated and inappropriate coverage of research findings in the news media’ is ultimately related to inaccurately reporting or over-interpreting research results in many abstract conclusions. A study published in an academic literature concluded that ‘inconsistencies in data between abstract and body, reporting of data and other information solely in the abstract are relatively common and that a simple educational intervention directed to the author is ineffective in reducing that frequency’. Other studies comparing the accuracy of information reported in a journal abstract with that reported in the text of the full publication have found claims that are inconsistent with, or missing from, the body of the full article.

Abstracts are protected under the copyright law just as any other form of written manuscript is protected. However, some publishers of the scientific articles invariably make abstracts freely available, even when the article is not free. For example, articles in the biomedical literature are available publicly from MEDLINE which is accessible through PubMed.

An academic abstract typically outlines four elements as mentioned below which may be relevant to the completed work:

- The research focus (i.e., statement of the problem(s)/research issue(s) addressed).
- The research methods used (experimental research, case studies, questionnaires, etc.).
- The results/findings of the research.
- The main conclusions and recommendations.

Abstract may also contain brief references, although some publications have the standard style of excluding the references from the abstract, reserving them for the article body (which, by definition, treats the same topics but in more depth).

‘Length of Abstract’ varies by discipline and publisher requirements. However, typical length ranges from 100 to 500 words, but very rarely more than a page and occasionally just a few words.

An abstract may or may not have the section title as ‘abstract’ prominently listed as an antecedent to the content. Abstracts are generally sectioned logically as an overview of what appears in the paper, with any of the following subheadings: Objectives, Introduction, Methods, Results, Conclusions, Background. Abstracts, where these subheadings are explicitly given, such abstracts are called ‘structured...
Abstracts’ by publishers. Abstracts that comprise one paragraph (no explicit subheadings) are often called ‘unstructured abstracts’ by publishers.

**Informative**

The informative abstract, which is also known as the complete abstract, contain comprehensive summary of a paper’s material including its purpose, background, methodology, results, and conclusion. Usually between 100 to 200 words. The informative abstract in general summarizes the structure of any literature, its major topics and key points. A format for scientific short reports that is similar to an informative abstract has been proposed in recent years. Informative abstracts may be viewed as standalone documents.

**Descriptive**

The descriptive abstract, which has some limitations so is also known as the limited abstract or the indicative abstract. It provides a description of what the paper covers without going into its subject material. A descriptive abstract is limited to a table of contents in paragraph form.

**Graphical Abstracts**

A graphical abstract is a single, concise, pictorial and visual summary of the main findings of the article. A graphical abstract will be a one-image file and will visualize one process or make one-point clear. For ease of browsing, the graphical abstract will have a clear start and end, preferably ‘reading’ from top to bottom or left to right. This could either be the indicative figure from the article or a figure which is typically designed for the purpose, which contain the content of the article for readers at a single glance.

With the influence of storage and retrieval systems of computer such as the Internet browsing, some scientific publications, primarily those published by Elsevier, started including graphical abstracts alongside the text abstracts. The graphic is intended to summarize or be an exemplar for the main thrust of the article. It is not intended to be as exhaustive a summary as the text abstract, rather it is supposed to indicate the type, scope, and technical coverage of the article at a glance. The use of graphical abstracts has been generally well received by the scientific community. Moreover, some journals also include video abstracts and animated abstracts made by the authors to easily explain their papers. Many scientific publishers currently encourage authors to supplement their articles with graphical abstracts, in the hope that such a convenient visual summary will facilitate readers with a clearer outline of papers that are of interest and will result in improved overall visibility of the respective publication. However, the validity of this assumption has not been thoroughly studied, and a recent study statistically comparing publications with or without graphical abstracts with regard to several output parameters reflecting visibility failed to demonstrate an effectiveness of graphical abstracts for attracting attention to scientific publications.
Critical Abstract

A crucial abstract provides, additionally to describing main findings and data, a judgement or comment concerning the study’s validity, liableness, or completeness. The man of science evaluates the paper and infrequently compares it with different works on the identical subject. Crucial abstracts are usually 400-500 words long thanks to the extra interpretative comment. These sorts of abstracts are used sometimes.

Highlight Abstract

A highlight abstract is specifically written to draw in the reader’s attention to the study. No pretence is created in either a balanced or complete image of the paper and, in fact, incomplete and leading remarks could also be accustomed in order to spark the reader’s interest. In this a highlight abstract cannot stand independent of its associated article. Highlight abstract is not a real abstract and, therefore, it is seldom employed in tutorial writing.

Abstract Quality Assessment

Various methods can be used to evaluate abstract quality, e.g. rating by readers, checklists (not necessary in structured abstracts), and readability measures.

When the need does arises of writing the abstracts
- When submitting the articles to journals, especially online journals
- When applying for research grants
- When writing a book proposal
- When completing the Ph.D. dissertation or M.A. thesis
- When writing a proposal for a conference paper
- When writing a proposal for a book chapter

Most likely, the author of the entire work (or prospective work) writes the abstract. However, there are professional abstracting services that hire writers to draft abstracts of other people’s work. In a work with multiple authors, the first author usually writes the abstract. Undergraduates are sometimes asked to draft abstracts of books/articles for classmates who have not read the larger work.

Writing Style

Always use active voice wherever possible; however note that a lot of your abstract might need passive sentence constructions. Regardless, write your abstract in concise, but giving complete sense. Get to the purpose quickly and use past tense each time as a result of reportage on a study that has been completed.

Although it is the primary section of your paper, the abstract, by definition, ought to be written last as it is like the summary of contents of your entire paper. To start composing your abstract, take whole sentences or key
phrases from every section of the paper and put them in a sequence that summarizes the contents of the paper. Then revise or add connecting phrases or words to create the narrative flow clearly. Before handing in your final paper, check to ensure that the data within the abstract fully matches with what you have written in the paper. An abstract is a sequential set of complete sentences which describes the most crucial information using the least required words.

The abstract should not contain the following:
1. Avoid ‘Lengthy’ background data.
2. References to alternative literature [say one thing like, “current analysis shows that...” or “studies have indicated...”].
3. Using elliptical [i.e., ending with “...”] or incomplete sentences.
4. Abbreviations, acronyms, jargon, or terms that will be confusing to the reader.
5. Any form of image, illustration, figure, or table, or references to them.

Abstract Management

Abstract management is a method of accepting and making abstracts ready for presentation at an academic conference. The method consists of either invited or proffered submissions of the abstract or outline of labour. The abstract usually states the hypothesis, tools employed in analysis or investigation, knowledge collected, and an outline or interpretation of the information.

The abstracts typically endure review once that they are accepted or rejected by the conference chair or committee and so allotted to conference sessions. The abstracts could also be bestowed as associate degree oral speak or as an illustrated poster throughout the event. Abstracts are usually printed before or once the event as conference proceedings or in tutorial journals or online. In some cases, submission of a full paper could also be needed before final acceptance is given. In some fields (e.g., laptop science), most thought conferences associate degreed workshops arouse the submission of full papers (rather than simply abstracts) and educational program committees review the complete paper to a typical akin to journal publication before accepting a paper for presentation at the conference and business enterprise it in an altered proceedings series.

The abstract management method is closely tied to the requirement to supply continued education to professionals, particularly continued medical education. Several annual conferences hosted by specialty societies give instructional credit hours in order that attendees might keep current within the field and maintain their skilled certifications.

Software

Historically, abstract management was a long manual method which requires the handling of enormous amounts of paper and created a substantial body work.
The increasing range of organizations currently use Web-based abstract management code to contour and modify the method. The work is typically outsourced to dedicated conference departments at major publishers and skilled conference organisers.

A conference management system is a web-based code that supports the organization of conferences particularly scientific conferences. It helps the program chair(s), the conference organizers, the authors and therefore the reviewers in their various activities.

Abstract submission involves the authors in making their abstracts ready and sending them to the conference organisers through an internet type, and may be a comparatively easy method. The abstracts are either uploaded as documents (typically Microsoft Word, PDF or LaTeX) or, wherever graphics and tables are not needed, they will merely be entered into the shape as plain text. The software package can send an email acknowledgement. Following the committee’s judgements on which abstracts are to be accepted for the conference the submission software package may additionally be used to collect full papers and PowerPoint presentations.

**Fast Abstract**

The conception of the ‘Fast Abstract’ was first created at the 28th Fault-Tolerant Computing Conference (1998) by Ram Chillarege, where the UN agency was the program co-chair of the Conference. They are intended to market communication and discussion of subjects that have nonetheless to be established or completed as analysis comes, and to supply a venue for new barely tested concepts, early experiments, or opinions. Whereas informal discussions, stand-up talks, spherical tables offer an avenue for such sharing, they lack the written account that is essential for progress. Quick abstracts aim to fill this gap. Since their origin, quick abstract sessions are step-by-step process adopted by many conferences, in some cases even being among the foremost attended sessions.

A fast abstract, additionally extended abstract, is a short, gently reviewed technical article that is typically given with a brief speak at a scientific conference. The length of the document is typically restricted to two pages (including all text, figures, references and appendices), though some conferences could enable slightly longer articles. If the conference does not specify a document vogue, the quality double-column IEEE format could be a common follow.

Fast abstracts enable authors to:

1. Report technical work preliminary stages, and solicit early feedback.
2. Present radical opinions on disputed problems or open issues.
3. Introduce new problems to analyse.
4. Present expertise obtained from apply.
Microforms

Microforms—microfilm, microfiche, text/fiche, etc., have been employed in libraries for nearly sixty years. There is currently an awfully massive and growing quantity of research analysis and evaluation of information material, such as the books, journals, newspapers, pamphlets, manuscripts, government documents, dissertations or the microfilms that are available in different formats for the research analysis purpose. Over the years the intellectual moreover because the technical qualities of micro publications have continued to enhance. Since new technologies work side-by-side with, instead of replacement, older ones, microforms may be expected to stay a considerable portion of the holdings of the Library of Congress and alternative analysis establishments, moreover as a major a part of publishers’ offerings.

Microforms: What are Microforms?

The ‘Microform’ is a microfiche, microfilm, microcard and microprint paragraph term. It is usually a type of film, but on paper. These are micro productive materials, about a 25th original document size. These materials are Microform’s advantages include the conservation of rare and fragile materials, increased access and reduced storage space. Nearly any publishing type can be microform, for example books, journals, papers, newspaper, government documents, manuscripts, poetry, and many more.

Depending upon the images being provided the microform images can be positive or negative:

Microforms are in following formats:
1. Microfilm – usually termed as Reels
2. Aperture Cards
3. Microfiche (Flat Sheets) / Micro Cards (Printed on Cardboard)

By the mid of 25th century, these microforms have raised up to maximum level as most of the periodicals, books and other collections were being converted into microform.

Microforms are any forms containing micro-reproductions of documents for transmission, storage, reading, and printing. Microform pictures are unremarkably reduced to concerning one ordinal of the initial document size.

Within the mid-20th century, libraries started using microfilm as a preservation strategy which deteriorated newspaper collections. The books and newspapers that were considered at risk of decay may be preserved on film so as to enhance the access and utilization. The library’s microforms assortment contains newspapers, magazines, and educational journals likewise as educational and government documents.
Most of the library’s microform collections may not be always available in the library because they are held on at the off-site storage facility, which when required should be requested by researcher or analyst for the library use. Below is an example of a record in the Library Catalogue that shows that the item is on microform hold on at the Hampden Centre. To view this microform record within the ‘Catalogue’, use the ‘Request It’ button to possess the item delivered to the library.

Microform in the Twentieth Century

The most important usage of microforms is the storage of documents, images, architectural or technical drawings in a proper way. Along with microforms, digital preservation has become popular since then.

Now, with the increasing usage of digital preservation, microform is existing in today’s date in various libraries and institutions.

First, recorded pictures are the analog type’s images of the initial, reduced in size, and users will access the knowledge with a straightforward instrumentality like a scientific instrument. Digital technologies, however, need rather more complicated devices. Once laptop programs are updated, the initial digital
knowledge might become inaccessible. The life of CDs, DVDs, and alternative digital storage devices remains unsure. For these reasons, digital preservation needs constant and continuous knowledge migration. On the opposite hand, microforms are expected to last concerning five hundred years, if they are properly preserved.

While microform encompasses a variety of benefits, it lacks some functions digital preservation that has: search capability; fast data transfer from one location to a different location; huge storage capacity; simple manipulation of information.

To overcome a number of those disadvantages, knowledge preserved in microform are also digitized. At some establishments, users will favour to access data storage in numerous formats. For instance, the Library of Congress in the U.S. provides copies of their assortment in numerous formats. Users will select and request from the followings:

- 35mm Microfilm
- Film to Paper
- Original Cinematography
- Photocopying
- Cartographical Scanning
- Digital Pictures to Compact Disc Read-Only Memory (CD-ROM)
- All Formats of Photographic Reproduction

**History and Early Developments**

Using the Daguerreotype process, John Benjamin Dancer was one of the first to produce micro-photographs, in 1839. He achieved a reduction ratio of 160:1. Dancer formed his reduction procedures with Frederick Scott Archer’s wet collodion process, developed in 1850-1851. However, he laid off his decades-long work on micro-photographs as a personal hobby, and did not document his procedures. The notion that microphotography could be no more than a novelty was an opinion shared by the 1858 *Dictionary of Photography*, wherein the process is called as ‘somewhat trifling and infantile’.

Microphotography was initially proposed as a document preservation method in 1851 by James Glaisher, associate physicist, and in 1853 by John Herschel. Both men attended the 1851 Great Exhibition held in London, wherever the exhibition on photography greatly influenced Glaisher. He called it ‘the most exceptional discovery of contemporary times’, and argued in his official report for using microphotography to preserve documents.

The developments in microphotography continued through the subsequent decades, however it had been not till the turn of the century that its potential for practical usage was condemned by a wider audience. In 1896, Canadian engineer Reginald A. Fessenden proposed that microforms were a compact solutions to engineers’ unwieldy but frequently consulted materials.
He projected that up to a hundred and fifty lacs words could be made to fit in a square inch, and that a one-foot cube might could contain 1.5 million volumes.

Microfilm first saw military use in the Franco-Prussian War of 1870–1871. During the Seige of Paris, the sole means for the provincial government in Tours to speak with Paris was by Columbiform bird post, and, as the pigeons could not carry paper dispatches, the Tours government turned to microfilm. Employing microphotography unit exhausted from Paris before the seige, clerks in Tours photographed paper dispatches and compressed them to microfilm, which were carried by orientating pigeons into Paris and projected by slide projector whereas clerks copied the dispatches onto paper.

In 1906, Paul Otlet and Robert Goldschmidt projected the ‘livre microphotographique’ as a way to alleviate the price and area limitations obligatory by the codex format. Otlet’s overarching goal was to make a World Centre Library of Juridical, Social and Cultural Documentation, and he saw microfilm as a method to supply a stable and long-lasting format that was cheap, straightforward to use, straightforward to breed, and very compact. In 1925, the team spoke of a large library where every volume existed as master negatives and positives, and where things were written on demand for interested patrons.

Microform Use in Libraries

At the annual meeting of 1936, the American Library Association endorsed microforms. Before this official acceptance, microforms were utilized in connected fields between 1927 and 1935, the Library of Congress (U.S.) microfilmed over 3 million pages of books and manuscripts in the British Library; in 1929 the scientific discipline analysis Council and also the Yankee Council of Learned Societies joined to form a Joint Committee on Materials analysis, that looked closely at microform’s potential to serve little print runs of Educational or Technical Materials. In 1933, Charles C. Peters developed a way to micro format dissertations and in 1934 the U. S. National Agriculture Library enforced the primary microform print-on-demand service which was rapidly followed by an analogous industrial concern, Science Service, and in 1938 University Microfilms was established and also the Harvard Foreign Newspapers Microform Project was enforced.

The ‘MicroCard’ was outdated by ‘Microfilm’. By the eighties, microfilming had become customary policy in libraries as a method of reformatting each books and newspapers.

In 1940, the format most broadly speaking used these days – microfilm – was developed. Formats it bested embody the Photo scope, the Film-O-Graph, the Fiske-O-Scope, film slides.

Early cut sheet microforms and microfilms (to the 1930s) were written on a nitrate film and create high risks to their holding establishments, as nitrate film is explosive and ignitable. From the late 1930s to the 1980s, microfilms were
Literature Search

commonly written on a cellulose acetate base, which is not good for eyes and was prone to tears, vinegar syndrome and chemical reaction blemishes. Vinegar syndrome is the result of chemical decay and produces “buckling and shrinking, embrittlement, and bubbling”. Chemical reaction blemishes are yellow, orange, red spots which are 15-150 micrometres in diameter created by oxidative attacks on the film, and are largely due to the poor storage conditions.

Advantages

The medium has numerous advantages:

- It is compact, with way smaller storage prices than paper documents. Ordinarily ninety-eight document size pages match on one fiche. Compared to filing paper, microforms will cut back house storage needs by up to ninety-five per cent.
- It is cheaper to distribute than paper copy. Most microfilm services get a bulk discount on replica rights, and have lower replica and carriage prices than a comparable quantity of written paper.
- It is a stable depository kind once properly processed. Preservation commonplace microfilms use wet-processed silver salt dyes in onerous gelatin emulsion on a polyester base. With acceptable storage conditions, this film contains a lifespan of five hundred years. Unfortunately, in tropical climates with high wetness, fungus eats the gelatin accustomed bind the silver salt. Thus, diazo-based systems with lower depository lives (20 years) that have polyester or epoxy surfaces are used.
- Since it is ANALOG i.e., an actual image of the initial data, hence it is straightforward to look at. In contrast to digital media, the format needs no package to decipher the info keep thereon—it is instantly intelligible to person’s knowledge within the language; the sole instrumentality that s required may be an easy simple microscope. This reduces the chance of degeneration.

Disadvantages

- The principal disadvantage of microforms is that the image is just too little to scan with the optic. Libraries should use special readers that project life-size pictures on a ground-glass screen.
- Reader machines used to view microform are typically troublesome to use, requiring the user to rigorously wind and rewind until they have arrived at the purpose wherever the information they are searching for is stored.
- Also, photographic illustrations reproduce poorly in microform format, with the image sometimes not being as sharp as the original. The last drawback with the machines is that printers do not seem to be forever out there, limiting the user’s ability to create copies for his or her own functions.
- Colour microfilm is extraordinarily pricy, hence most libraries do not opt for supplying colour films. This ends up in the loss of some info, as colour
materials are usually going to be preserved victimization regular microfilm so as to avoid wasting cash.

- Once keep within the highest-density drawers, it is straightforward to misfile a fiche, that is thenceforth out of stock. Some libraries so keep the microfilm cupboard in a very restricted space, and retrieve fiches on demand. Some fiche services use lower-density drawers with tagged pockets for every card.
- A conventional photocopier cannot reproduce the pictures.
- Like all ANALOG media formats, microfilm lacks options taken as a right by users of digital media. ANALOG copies degrade with every generation, whereas digital copies have abundant higher repeating fidelity. Digital knowledge also can be indexed and searched simply.

Readers and Printers
Desktop readers are boxes with a semi-transparent screen at the front on to that is projected a picture from a microform. They need appropriate fittings for no matter microform is in use. They will supply a selection of magnifications. They typically have powered movement of film. Once secret writing blips are recorded on the film a browser is employed which will read the blips to search out any needed image.

Portable readers are plastic devices that fold for carrying, once open they project an image from microfilm on to a reflective screen.

A microfilm printer contains a photography repetition method, like a photocopier. The image to be printed is projected with synchronous movement on to the drum. These devices supply either tiny image preview for the operator or full size image preview and it is called a reader printer. Microform printers sometimes settle for positive or negative films, to provide positive pictures on paper.

Media
Flat Film - 105 x 148 mm flat film is used for micro images of very large engineering drawings. These may carry a title photographed or written along one edge. Typical reduction is about 20, representing a drawing that is 2.00 x 2.80 metres, that is 79 x 110 inches (2,800 mm). These films are stored as microfiche.

Microfilm
16 mm or 35 mm film to motion picture standard is used, usually unperforated. Roll microfilm is stored on open reels or put into cassettes. The standard length for using roll film is 30.48 m (100 ft). One roll of 35 mm film may carry 600 images of large engineering drawings or 800 images of broadsheet newspaper pages. 16 mm film may carry 2,400 images of letter sized images as a single stream of micro images along the film set so that lines of text are parallel to the sides of the film or 10,000 small documents, perhaps cheques or betting slips, with both sides of the originals set side by side on the film.

**Aperture Cards**

Aperture cards are Hollerith cards into which a hole has been cut. A 35 mm microfilm chip is mounted in the hole inside of a clear plastic sleeve, or secured over the aperture by an adhesive tape. They are used for engineering drawings, for all engineering disciplines. There are libraries of these containing over 3 million cards. Aperture cards may be stored in drawers or in freestanding rotary units.

**Microfiche**

A microfiche is a flat film 105 x 148 mm in size, that is ISO A6. It carries a matrix of micro images. All microfiche are read with text parallel to the long side of the fiche. Frames may be landscape or portrait. Along the top of the fiche a title may be recorded for visual identification. The most commonly used format is a portrait image of about 10 x 14 mm. Office size papers or magazine pages require a reduction of 24 or 25. Microfiche are stored in open top envelopes which are put in drawers or boxes as file cards.

**Image Creation**

To create microform media, a planetary camera is mounted with vertical axis higher than a replica that’s stationary throughout exposure. A flow camera moves the
copy smoothly through the camera to show film that moves with the reduced image. Instead, it is going to be created by computers, i.e., COM (Computer Output Microfilm).

**Film**

Normally microfilming uses high resolution panchromatic high resolution monochrome stock. Positive colour film providing sensible copy and high resolution can even be used. Photographic film is sixteen, thirty-five and one zero five-millimetre-wide in lengths of thirty metres (100 ft) and longer, and is sometimes unperforated. Photographic film is developed, fastened and washed by continuous processors.

Sheet film is equipped in ISO A6 size. This is often either processed by hand or employing a dental X-ray processor. Camera film is equipped with prepared mounted in aperture cards. Aperture cards are developed, fastened and washed straight off when exposure by instrumentation fitted to the camera.

**Cameras**

**Flat Film**

The simplest microfilm camera that is still in use could be a rail mounted structure at the highest of that is a bellows camera for 105 x 148 millimetre film. A frame or copy board holds the initial drawing vertical. The camera contains a horizontal axis that passes through the centre of the copy. The structure could also be moved horizontally on rails. In a room one film could also be inserted into a dark slide or the camera may be fitted with a photographic film holder that when associate exposure advances the film into a box and cuts the frame off the roll for process as a single film.

**Roll Film**

For engineering drawings, a separate open steel structure is usually provided. A camera is also captive vertically on a track. Drawings are placed on an oversized table for photography, with centres below the lens. Mounted lights illuminate the copy. These cameras are typically over three metres (10 feet) high. These cameras settle for film stock of thirty-five or sixteen millimetre.

For workplace documents an identical style is also used. This can be a smaller version of the camera. These are provided either with the selection of sixteen or thirty-five millimetre film or 16 mm film solely. Non-adjustable versions of the workplace camera are provided. These have a rigid frame or associate close box that holds a camera at a set position over a replica board. If this can be to figure at over one reduction quantitative relation there are a selection of lenses.
Some cameras expose a pattern of sunshine, noted as blips, to digitally determine every adjacent frame. This pattern is traced whenever the film is copied for looking.

**Flow Roll Film Cameras**

A camera is made into a box. In some versions this is often for bench prime use, different versions are transportable. The operator maintains a stack of fabric to be recorded in a very receptacle, the camera mechanically takes one document once another for advancement through the machine. The optical lens sees the documents as they pass a slot. Film behind the lens advances specifically with the image.

Special purpose flow cameras film each side of documents, golf shot each pictures aspect by side on sixteen millimetre film. These cameras are accustomed record cheques and sporting slips.

**Microfiche Camera**

All microfilm cameras are planetary with a step and repeat mechanism to advance the film once every exposure. The less complicated versions use a dark slide loaded by the operator during a dark room; once exposure the film is processed one by one, which can be by hand or employing a dental X-ray processor. Cameras for prime output are loaded with a roll of 105 millimetre film. The exposed film is developed as a roll; this can be typically move individual fiche once process or unbroken in roll type for duplication.

**Computer Output Microfilm (COM)**

Equipment is offered that accepts an information stream from a digital computer. This exposes film to supply pictures as if the stream had been sent to a printing machine and therefore the listing had been microfilmed. Due to the supply run one might represent several thousands of pages.

Within the instrumentality character pictures are created by a light-weight supply, this can be the negative of text on paper. COM is usually processed commonly. Alternative applications need that image seems as a standard negative; the film is then reversal processed. This outputs either sixteen millimetre film or fiche pages on a 105 mm roll.
Because listing characters is an easy style, a discount magnitude relation of fifty provides smart quality and puts regarding three hundred pages on a microfilm. A microfilm plotter, generally known as an aperture card plotter, accepts a stream that may be sent to a laptop pen plotter. It produces corresponding frames of microfilm. These manufacture microfilm as thirty-five or sixteen millimetre film or aperture cards.

**Storage and Preservation**

Storage facilities are developed by various manufacturers in such a way that they are exceptionally appropriate for the preservation, safekeeping and protection of the microfilm, microfiche and aperture cards.

The storage facilities has to maintain and preserve a constant 40 per cent - 50 per cent average relative humidity, 68°F/20°C average temperature, irrespective of weather or season. Microfilms are specifically stored in ‘Microfilm Vaults’ which are dehumidified and temperature-controlled as per the specific requirements of the users.

The ANSI/NAPM IT9.11 normal specifies the subsequent mixtures of temperature and ratio for extend-term storage of black and white microfilms of all types: 21°C (70°F) with RH of 20-30%; 15°C (60°F) with RH of 20-40%; and 10°C (50°F) with RH of 20-50%. For colour microfilm, the ANSI/NAPM IT9.11 normal specifies 2°C (36°F) with RH of 20-30%. William Saffady recommends that colour microfilm should be kept into heat-sealed foil luggage for wet protection and to limit exposure to air.

For medium-term storage (at least 10 years), ANSI/NAPM IT9.11 normally specifies that temperature should not exceed 25°C (77°F) and RH should stay stable inside 20-50 per cent. Wetness variations should not exceed tenth per day, and natural action ought to be reduced.

Microfilm ought to be kept in dark enclosures to reduce harm from lightweight. Enclosures ought to suits preservation standards.

Microfilm storage areas ought to be situated in a very non-combustible house that is unbroken, clean and freed from mud particles and alternative contaminants, moreover as gases like sulphur dioxide, sulphide, ammonia, and ozone. All building materials and storage instrumentality ought to be fireproof and noncorrosive. Microfilms ought to be frequently inspected for signs of decay.

**Duplication**

All regular microfilm repeating involves contact exposure struggling. Then the film is processed to produce a permanent image. Hand repeating of single fiche or aperture cards uses exposure over a light-weight box and one by one process of the film. Silver salt film may be a slow version of camera film with a sturdy prime coat. It is appropriate for prints or to be used as associate degree intermediate
from that more prints are also made. The result is a negative copy. Preservation standards need a master negative, a replica negative, and a service copy (positive). Master negatives are unbroken in deep storage, and duplicate negatives are used to produce service copies, that are the copies accessible to researchers. This multi-generational structure ensures the preservation of the master negative.

Diaz-o-sensitised film for dye coupling in ammonia gives blue or black dye positive copies. The black image film is used for more repeating.

Vesicular film is hypersensitive with a caution dye, that once exposure is developed by heat, wherever light-weight has return to the film remains clear, within the areas underneath the dark image the caution compound is destroyed quickly, emotional immeasurable minute bubbles of gas into the film. This produces a picture that diffuses light-weight. It produces a decent black look in a very reader, however, it cannot be used for more repeating.

Finding Microforms for Research in the Library of Congress

The Library of Congress acquires microform materials in the same way it acquires inkprint materials (largely through the operation of the Copyright Law), therefore it currently owns countless items of microforms, as well as many hundred collections, that are housed and maintained by completely different sections and reading rooms. The Manuscript Division, the Law Library, the Serials Division, and therefore the Science and Technology Division, among others, additionally own and repair massive microform collections. The final rule is that a microform item is within the custody of and repaired by the identical section or room that might have custody and service of the item if it were in paper. As an example, an English newspaper would move to the Serials Division not considering the format, whereas a book in Japanese would move to the Asian Division no regard of format, and then forth.

Searching for microforms in LC is generally similar to look for things in print format. Many of the tools and techniques are identical. Items in microform can usually show up in search results, typically unexpectedly. Following is a list of basic sources to assist you find microforms at the Library of Congress.

Basic Sources: Examples

1. LC computerized card catalogues (MUMS and SCORPIO, 1982- ; Premarc (PREM) -1980. (Also known as LOCIS).
2. Card catalogues:
   (a) MicRR Card Catalogue (Pre-1982)
   (b) MicRR Serial File (Pre-1988)
   (c) Main Card Catalogue (Pre-1980, including the Periodicals Catalogue)
3. MRR CD-ROM Network
   Some of the databases available, such as ERIC, Dissertation Abstracts, Congressional Masterfile I and II, and Statistical Master File Index Microforms owned by LC.
4. “Microform Collections and Selected Titles in the Microform Reading Room” (“K-3” and supplement, “K-4”; future editions will be available online and in print) is an annotated list with a subject index which can be consulted in many reading rooms.

5. Guides, indexes, and finding aids to microform collections or collected sets owned by LC:
   (a) produced by micropublisher
   (b) produced by LC staff
   (c) described in bibliographies on which collections are based
   (d) combinations of the above

6. Other publications produced by LC staff
   (a) HSS Research Guides
     (b) Quicksilver (MicRR publications no longer being issued)
   (c) Photoduplication Circulars

7. Reference Librarians
   By providing thousands of leading global, national and local newspapers and hundreds of unique research collections on a broad range of subjects on microfilm, ProQuest enables your library to provide researchers with critical information.

6.5 COMPUTERIZED INFORMATION RETRIEVAL SYSTEM

Definition
Information retrieval (IR) is the action of acquiring data framework assets applicable to a data require from an accumulation of data assets. Pursuits can be found on full-content or other substance-based ordering. Data recovery is the investigation of scanning for data in a record, hunting down reports themselves, and furthermore looking for metadata that portray information, and for databases of writings, pictures or sounds.

According to Salton (1983), ‘An information retrieval system is an information system, that is, a system used to store items of information that need to be processed, searched, retrieved, and disseminated to various user population’.

Computerized data recovery frameworks are utilized to diminish what has been called data over-burden. An IR framework is a product that give access to books, diaries and different records, stores them and deals with the archive. Web indexes are the most obvious IR applications.

Introduction
An information retrieval process starts when a client enters an inquiry into the system. Questions are formal articulations of data needs, for instance search strings
in web search engines. In data recovery a question does not interestingly recognize a solitary item in the gathering. Rather, a few items may match the inquiry, maybe with various degrees of pertinence.

An item is an element that is spoken to by data in a substance collection or database. Client inquiries are matched against the database data. Be that as it may, instead of established SQL questions of a database, in information retrieval the outcomes returned may or may not match the query, so results are ordinarily positioned. This positioning of results is a key contrast of data recovery seeking contrasted with database searching.

Depending upon the application the information items might be, for instance, content records, images, audio, mind maps or recordings. Often the documents themselves are not kept or stored directly in the IR system, but are instead represented in the system by document surrogates or metadata.

Most IR frameworks register a numeric score on how well each item in the database coordinates the inquiry, and rank the articles as indicated by this esteem. The best positioning items are then appeared to the client. The procedure may then be iterated if the client wishes to refine the query.

In today’s scenario, we are being equipped with powerful tools for information handling – for gathering, organization, classification, retrieval and circulation. Since the late 1960s, computers have been used for the storage of large databases for example, library catalogues and bibliographic references. Further to the advancement, development of optical storage media in the form of CD-ROM, External Hard Disk, Pen Drive’s has a great advantage of saving cost by storing large amount of data, graphics, pictures and sound into them. One of the stated feature of this optical memories is that it plays a role of distributed stores for encyclopaedias, databased, books etc.

There are three aspects of computerised information retrieval:

1. Library Catalogues
2. Online Databases
3. Available databased on CD-ROM

We can retrieve information with the help of following tools:

1. Library Catalogue
2. Subject Classification
3. Systematic arrangement of books
4. Bibliographic and table of contents

One of the important point to remember is that these tools can be manual as well as computerized.

With the help of computers, information can be stored in the simplest way so that it can be available at any moment when required. Database are stored in a computer as per predefined standards.
For better understanding, all academic libraries and some of the public libraries are equipped with catalogues which are accessible to public along with retrieval functionalities. To easily identify the location of the books stored, they have come across computer-based shelf list.

How to store data
1. To choose and obtain source data.
2. Creating a vocabulary of terms for their explanation
3. Distributing terms to each and every specific data item
4. To create records, assigned terms should be arranged in some specific formats
5. Creating a way to access for the records by various means.

Steps for retrieving information
The main objective of storing any data or information is its retrieval at any stage of time.
1. For command, we should be using terms from the designed vocabulary.
2. Searching the data and match the stored information with the requests.
3. After retrieving, results to be presented accordingly.

Role of human in Computerised Information System
As we are aware, the computers are being developed to give quick access, storage of large amount of data with simplest form of communication between the human and system.

To retrieve correct path of information stored, systems has to be dependent on the important and crucial role plays by the human skills. It is the way to obtain the information source, to observe and put the data in order, to formulate request for data required.

Type of Information and Retrieval Systems
1. Database Systems – It is the most basic, all others are categorized on the basis of various data available and level of complexity involves in processing. Users which performs this function is Federal & local government agencies, Scientific research institutions, Academic libraries & B2B companies etc.
2. Reference Retrieval Systems – It is a database system which stores documents in the form of references, which means that for retrieval data will have terms descriptive of content. With the help of this, records identify the location of the data and then available for the retrieval to the particular user as requested. Some of the examples can be MEDLINE (Medical Online), OCLC (Online Computer Library Centre), RLIN (Research Libraries Information Network), AGRICOLA (Agricultural Online Access)
3. Documents Retrieval Systems – It has an ability to retrieve desired data in the form of text, images, which can be printed or displayed accordingly.
For example, bounded books and journals of library. In this, documents are being identified by some specific number and location. Storage of the document can be done in the form of printed pages, photographic film, CD-ROMs, Audio tapes and hard disk.

4. Text Retrieval Systems – In this system, text is being used as the source for the retrieval which means that information is saved in sentence or paragraph form. This type of system is mostly used in fields of law. For better explanation, dictionary is stored in Computer memory to derived alternative meanings or words.

5. Image Retrieval Systems – For retrieval the source of content, images are created with the help of word processing in every commercial enterprise, computer-aided designing, cartooning etc. The requirements for processing of image are as follows:
   - Conversion of any equipment to digitized form, it has to be scanned properly.
   - Optical storage should be with large volume of capacity.
   - Can be displayed and printed with high resolution.
   - A better communication device with a manageable bandwidth which can handle the large amount of data in transmitting the image.

Advantages of Computerised Information Retrieval
1. It is time saving.
2. With the help of the computer we can initiate the detailed search as compared to the method done manually. Each and every piece of information stored in computer is searchable at any point of time.
3. The data is more accurate and in current form when stored in a database rather than any relevant printed publication.
4. In single click we can search the information in several subjects during the same search engine.

Check Your Progress
8. What is a research review?
9. What is an argumentative review?
10. What is the role of theoretical review in research?
11. Define abstract?
12. What measures are used to evaluate abstract quality?
13. What do you understand by abstract management?
14. When was the concept of ‘Fast Abstract’ created?
15. What are aperture cards and what were they used for?
16. What are the three aspects of computerised information retrieval system?
NOTES

6.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. A literature review is a comprehensive compilation of the information obtained from published and unpublished sources of data in the specific area of interest to the researcher. This may include journals, newspapers, magazines, reports, government publications, and also computerized databases.


3. A library research plan is a predefined activity that gives direction to your research. It is an act that involves evaluation that helps determine the subsequent activities to be followed by the researcher. The research plan is a sequence of steps that the researcher should follow in order to get a comprehensible and reliable outline to adhere to.

4. The most common library research tools available in a library are as follows:
   - Library Catalogue
   - Almanacs
   - Dictionaries
   - Encyclopaedia
   - Bibliographies
   - Indexes
   - Search Engines

5. Journals refer to a periodical collection of articles. They also include sources like reports, bulletins or proceedings, published monthly by any organization or an institution. There are also certain scholarly journals, which are of great help to a researcher working in the field of humanities or social sciences.

6. Thesis refers to a piece of research work conducted by an individual to qualify for a degree. A researcher can browse through such thesis to get an idea of the layout and presentation options for his/her own thesis.

7. One of the main advantages of conducting research using the Internet is that hundreds or thousands of pages can be found with some relation to the topic, within seconds, which is not possible if the same topic is to be searched from books or encyclopaedias.

8. Review of any literature or research paper provides the current knowledge including prominent findings, as well as theoretical and methodological contributions to a particular topic. A literature review is a search and evaluation of the available literature in the given subject or chosen topic area. It
documents the state of the art with respect to the subject or topic you are
writing about.

9. This form examines literature by selection so as to support or refute associate
degree argument and the aim is to develop a body of literature that establishes
an investor viewpoint. Given the value-laden nature of some science analysis,
contentious approaches to analyzing the literature is a legitimate
and necessary kind of discourse.

10. The theoretical literature analysis or review helps in determining that which
toories already exist, the relationships between them, and up to what
degree or extent the contemporary theories are being investigated, and to
develop new specific hypotheses which must be tested.

11. An abstract is a brief summary of a literature, research article, thesis, reviews,
conference proceeding, or any in-depth analysis of a particular subject and
is often used to help the reader quickly ascertain the paper’s purpose.

12. Various methods are used to evaluate abstract quality e.g., rating by readers,
checklists (not necessary in structural abstracts), and readability measures.

13. Abstract management is a method of accepting and making abstracts ready
for presentation at an academic conference. The method consists of either
invited or proffered submissions of the abstract or outline of labour. The
abstract usually states the hypothesis, tools employed in analysis or
investigation, knowledge collected, and an outline or interpretation of the
information.

14. The conception of the ‘Fast Abstract’ was first created at the 28th Fault-
Tolerant Computing Conference (1998) by Ram Chillarege, where the UN
agency was the program co-chair of the Conference.

15. Aperture cards are Hollerith cards into which a hole has been cut. A 35 mm
microfilm chip is mounted in the hole inside of a clear plastic sleeve, or
secured over the aperture by an adhesive tape. They are used for engineering
drawings, for all engineering disciplines. There are libraries of these containing
over 3 million cards. Aperture cards may be stored in drawers or in
freestanding rotary units.

16. The three aspects of computerised information retrieval system are as follows:

- Library Catalogues
- Online Databases
- Available database on CD-ROM

6.7 SUMMARY

- A library research plan is a predefined activity that gives direction to your
research. It is an act that involves evaluation that helps determine the
subsequent activities to be followed by the researcher. As such, the research plan is a sequence of steps that the researcher should follow in order to get a comprehensible and reliable outline to adhere to.

- Once the researcher has executed the activities involved in the research plan, he/she should start looking for the essential and relevant information. This involves exploring information through traditional and modern library research tools, which contain specific bits of information as well as voluminous records and theories. To be able to make full use of these tools, it is necessary for the researcher to become familiar with the applications of these specific library tools.

- A library catalogue is an informative list of resources and materials available in the library. It comprises the name of books or journals along with the name of authors, subjects and publishing houses.

- An almanac is a chronological tabular publication that is published annually. Traditional almanacs, usually, contained information regarding weather forecasts, astronomical data and several other statistics like the rising and setting of sun, moon, eclipses etc.

- Dictionaries are most often regarded as a superficial source of information as it is supposed to be performing the sole function of defining the meanings of words.

- Encyclopaedia, in general, is a bulk informative volume containing a synopsis of the concerned subject, which is published in alphabetical order.

- A bibliography usually comprises a list of reference materials mentioned by the researcher. This list gives the names of various sources that the researcher has resorted to, such as books and articles for research.

- Enumerative Bibliography is also known as compilative, reference or systematic bibliography.

- Analytical Bibliography is further classified into descriptive, historical and textual bibliographies. Usually, these are concerned with the physical attributes and contemporary importance of a book.

- Indexes, as is well known, are the alphabetical lists of authors’ name and subjects containing the relevant page numbers where these topics and authors have been discussed or described in the book.

- Search engines refer to software that browse through the Internet for the queried information and provide sites, which contain the concerned information within a few seconds.

- Journals refer to a periodical collection of articles. They also include such sources like reports, bulletins or proceedings, published monthly by any organization or an institution.
• Thesis refers to a piece of research work conducted by an individual to qualify for a degree. Usually, libraries often hold copies of thesis, which have been submitted and approved for Ph.D.
• Manuscripts and archives are unpublished, sometimes hand-written or typed, original and primary sources of information.
• One of the advantages of conducting research using the Internet is that hundreds or thousands of pages can be found with some relation to the topic, within seconds, which is not possible if the same topic is to be searched from books or encyclopaedias.
• There are various tools such as Internet search engine and Internet guide that a researcher can use for collecting the information.
• Research reviews, abstracts and microforms are the tools that specify the significant points of published literature and the various research papers, published or unpublished.
• The purpose of any review paper is to compactly evaluate or review the recent progress for the specific topics. Overall, the review paper summarizes the state of data of the subject.
• Literature reviews are a basis for research in nearly every academic field introducing new research, serving to position the current study within the body of the relevant literature and to provide context for the researcher, which generally precedes the methodology and results in possible segments of the work.
• There are different types of literature reviews namely argumentative review, integrative review, historical review, methodological review, theoretical review, systematic review, scoping review.
• An abstract is a brief summary of a literature, research article, thesis, reviews, conference proceeding, or any in-depth analysis of a particular subject and is often used to help the reader quickly ascertain the paper’s purpose.
• Abstract is used in many academic literatures just for clear communication about the complex research. An abstract may act even as a stand-alone entity in place of a full paper. As such, an abstract is used by many organizations as the basis for selecting research that is proposed for presentation in the form of a poster, platform/oral presentation or workshop presentation at an academic conference or seminar.
• The informative abstract, which is also known as the complete abstract, contain comprehensive summary of a paper’s material including its purpose, background, methodology, results, and conclusion.
• The descriptive abstract, which has some limitations so is also known as the limited abstract or the indicative abstract.
• A graphical abstract is a single, concise, pictorial and visual summary of the main findings of the article. A graphical abstract will be a one-image file and will visualize one process or make one-point clear.

• A crucial abstract provides, additionally to describing main findings and data, a judgement or comment concerning the study’s validity, liableness, or completeness.

• A highlight abstract is specifically written to draw in the reader’s attention to the study. No pretence is created in either a balanced or complete image of the paper and, in fact, incomplete and leading remarks could also be accustomed in order to spark the reader’s interest.

• Various methods can be used to evaluate abstract quality, e.g. rating by readers, checklists (not necessary in structured abstracts), and readability measures.

• Abstract management is a method of accepting and making abstracts ready for presentation at an academic conference. The method consists of either invited or proffered submissions of the abstract or outline of labour.

• The conception of the ‘Fast Abstract’ was first created at the 28th Fault-Tolerant Computing Conference (1998) by Ram Chillarege, where the UN agency was the program co-chair of the Conference.

• Microforms—microfilm, microfiche, text/fiche, etc., have been employed in libraries for nearly sixty years. There is currently an awfully massive and growing quantity of research analysis and evaluation of information material, such as the books, journals, newspapers, pamphlets, manuscripts, government documents, dissertations or the microfilms that are available in different formats for the research analysis purpose.

• The ‘Microform’ is a microfiche, microfilm, microcard and microprint paragraph term. It is usually a type of film, but on paper.

• The most important usage of microforms is the storage of documents, images, architectural or technical drawings in a proper way. Along with microforms, digital preservation has become popular since then.

• While microform encompasses a variety of benefits, it lacks some functions digital preservation that has: search capability; fast data transfer from one location to a different location; huge storage capacity; simple manipulation of information.

• Microphotography was initially proposed as a document preservation method in 1851 by James Glaisher, associate physicist, and in 1853 by John Herschel.

• By the eighties, microfilming had become customary policy in libraries as a method of reformatting each books and newspapers.
• Desktop readers are boxes with a semi-transparent screen at the front on to that is projected a picture from a microform. They need appropriate fittings for no matter microform is in use.

• A microfilm printer contains a photography repetition method, like a copier. The image to be printed is projected with synchronous movement on to the drum.

• Aperture cards are Hollerith cards into which a hole has been cut. A 35 mm microfilm chip is mounted in the hole inside of a clear plastic sleeve, or secured over the aperture by an adhesive tape.

• A microfiche is a flat film 105 x 148 mm in size, that is ISO A6. It carries a matrix of micro images. All microfiche are read with text parallel to the long side of the fiche.

• The simplest microfilm camera that is still in use could be a rail mounted structure at the highest of that is a bellows camera for 105 x 148 millimetre film.

• A camera is made into a box. In some versions this is often for bench prime use, different versions are transportable. The operator maintains a stack of fabric to be recorded in a very receptacle, the camera mechanically takes one document once another for advancement through the machine.

• All microfilm cameras are planetary with a step and repeat mechanism to advance the film once every exposure.

• There are different types of information and retrieval systems namely database systems, reference retrieval systems, documents retrieval systems, text retrieval systems, image retrieval systems.

6.8 KEY WORDS

• Enumerative Bibliography: It refers to the most common bibliography type in research writing and it basically the list of sources that most authors put at the end of a piece of writing so that any reader can be able to know where cited information was sourced from.

• Microform: It refers to a process for reproducing printed matter in a much reduced size.

• Literature Review: It refers to a scholarly paper, which includes the current knowledge including substantive findings, as well as theoretical and methodological contributions to a particular topic.

• Abstract Management: It refers to the process of accepting and preparing abstracts for presentation at an academic conference.

• Information Retrieval: It refers to the activity of obtaining information system resources relevant to an information need from a collection.
6.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions
1. What is the purpose of a review paper?
2. Write short notes on historical review and methodological review.
3. Briefly explain the main stages of systematic review.
4. What is an informative abstract?
5. Write a note on graphical and critical abstract.
6. What are microforms?
7. Briefly mention about microfiche.
8. What is a Computer Output Microfilm (COM)?

Long Answer Questions
1. Discuss the various steps of a research plan.
2. Discuss the analytical and methodological features of literature review?
3. What are the characteristics of systematic review? Discuss.
4. Describe the purpose and limitations of an abstract.
5. Discuss the state of microforms in the twentieth century.
6. Examine the advantages and disadvantages of microforms.
7. Assess the different types of information and retrieval systems.
8. Explain the role of human in computerised information retrieval system.

6.10 FURTHER READINGS


BLOCK - III

RESEARCH PROBLEM AND HYPOTHESIS

UNIT 7  HYPOTHESIS

Structure
7.0  Introduction
7.1  Objectives
7.2  Meaning, Importance, Sources and Characteristics
   7.2.1  Criteria for Hypothesis Construction
   7.2.2  Types of Statistical Hypothesis
7.3  Formulation of Hypothesis and Testing
   7.3.1  Procedure for Hypothesis Testing
   7.3.2  Committing Errors: Type I and Type II
   7.3.3  Null and Alternative Hypotheses
7.4  Answers to Check Your Progress Questions
7.5  Summary
7.6  Key Words
7.7  Self Assessment Questions and Exercises
7.8  Further Readings

7.0  INTRODUCTION

A hypothesis is a speculation which is based on inadequate evidences that is advanced to further testing and experimentation. After the testing, a hypothesis can be proven true or false. The unit is all about the importance of hypothesis and its types. The unit also discusses the meaning, sources and characteristics of hypothesis. You will also learn about the rules for constructing good hypothesis, difference between proposition, hypothesis and theory. Besides, it will also focus on procedure for hypothesis testing and will draw a comparison between null hypothesis and alternate hypothesis

7.1  OBJECTIVES

After going through this unit, you will be able to:
- Understand the meaning of hypothesis
- Describe the importance, sources and characteristics of hypothesis
- Define the rules for constructing good hypothesis
- Discuss the types of hypothesis
- Differentiate between proposition, hypothesis and theory
- Understand the formulation of hypothesis and testing
Hypothesis

• Explain the procedure for hypothesis testing
• Draw comparison between null hypothesis and alternate hypothesis

7.2 MEANING, IMPORTANCE, SOURCES AND CHARACTERISTICS

According to Theodorson, ‘a hypothesis is a tentative statement asserting a relationship between certain facts’. Kerlinger describes it as ‘a conjectural statement of the relationship between two or more variables’.

Need of hypothesis

The following points help in understanding the importance of hypothesis:

• A hypothesis is a proposal intended to explain a fact or an observation.
• A hypothesis specifies the sources of data which shall be studied and in what context they shall be studied.
• It determines the data needs.
• Hypothesis suggests the type of research which is likely to be the most appropriate.
• A hypothesis contributes to the development of the theory.

Nature of hypothesis

Hypothesis is more useful when stated in precise and clearly defined terms. A good hypothesis implies that hypothesis which fulfils its intended purposes and is up to the mark. The following are some important points to be kept in mind:

(a) A good hypothesis should be stated in the simplest possible terms. It is also called the principle of the economy or business. It should be clear and precise.
(b) A good hypothesis is in agreement with the observed facts. It should be based on original data derived directly.
(c) It should be so designed that tests will provide an answer to the original problem which forms the primary purpose of the investigation.
(d) Hypothesis should state relationship between variables, if, it happens to be a rational hypothesis.

7.2.1 Criteria for Hypothesis Construction

Once the investigative questions are set up for each of the objectives, the researcher should identify the anticipated or possible answers to the investigative questions. A survey of related theories and earlier studies and discussions with co-scientists will facilitate this process. He, then, should write down those answers as appropriate types of hypotheses—descriptive, relational or causal—as the case may be. He should evaluate these tentative hypotheses in terms of the characteristics of a good hypothesis and refine and record them into logical and testable hypotheses, keeping in mind the rules given below:
Rules for constructing hypotheses

According to research theorist Smith, there are certain rules for constructing good hypotheses. These are as follows:

(i) Search for variable measurements with the most quantitative characteristics available: Precise quantitative measurements are more critical in testing theory than qualitative characteristics. Another theorist Hage gives four techniques to search for and create variables from non-variable concepts. First, the researcher can search for implied dimensions underlying non-variable concepts. Campbell provides a good example of this method in the study of the non-variable concept ‘social group’. He identifies four underlying dimensions of degrees of proximity, similarity, perceived common fate and perceived spatial pattern. Second, one can create new variables by comparing conceptual synonyms or analogies. Price used the synonym technique in his study of organizational measurement concepts like ‘participation in decision making’, ‘organization control’, ‘power’ and ‘influence’ refer to the degree of organizational centralization. Third, one can search the literature for rarely occurring associations between phenomena. The cognitive dissonance theory in sociology got its start in this manner. Fourth, one can generate new variables through ordering many concepts from more or less abstract extension of the research variables and their applicability in the new order of society, say virtual or network organization.

(ii) Make the variable scale properties explicit by stating all of the variable’s mutually exclusive and totally inclusive categories by degrees. For example, a variable like ‘income’ may be categorized into (1) up to ₹ 5000 per month, (2) ₹ 5001 to 10000, (3) ₹ 10001 to 20000, (4) ₹ 20001 and above.

(iii) Describe the means used to sort observations into your variables categories in sufficient detail so that your methods may be evaluated and replicated by others. ‘Personality disintegration’ is a good example of a poorly operationalised variable. It is an unreliable measure and cannot be replicated.

(iv) Always consider alternative operations that might be more appropriate for a given variable.

(v) Analyse variables through their relationships. Non-ratio uni- or multi-variable distribution is arbitrary, since it has no intrinsic lower boundary.

(vi) Link two or more formal propositions through a shared independent or dependent variable where possible. For example, from following concrete observations:

(a) ‘Married persons are less likely than unmarried persons to commit suicide.’

(b) ‘Married persons with children are less likely than married persons without children to commit suicide.’
The following abstract formal hypotheses may be inductively produced:

(i) ‘Suicide rates vary directly with the degree of individualism.’

(ii) ‘Suicide rates vary indirectly with the degree of group cohesion.’

7.2.2 Types of Statistical Hypothesis

In the context of statistical analysis, we generally consider two types of hypothesis:

- Null Hypothesis
- Alternative Hypothesis

When comparing the superiority of both the methods A and B, if we assume that both the methods are equally good, then the assumption is known as ‘null hypothesis’. On the other hand, if we consider method A to be better, it is alternative hypothesis. These may be, symbolically presented as:

Null Hypothesis = Ho
Alternative Hypothesis = Ha

Difference between Proposition, Hypothesis and Theory

(i) A proposition is a logical statement of relationship between two or more variables which has, generally, been confirmed by empirical research. (A proposition should be distinguished from a hypothesis which is a logical statement of an assumed relationship between two or more variables which must be empirically tested, replicated and elaborated before being accepted as confirmed.)

(ii) Proposition is a broad statement drawn from a theory, whereas a hypothesis takes this one step further and formulates a more specific statement that is empirically testable. Proposition states a relationship between two concepts, and a hypothesis operationalizes this relationship and puts it in an empirically testable form.

(iii) The term ‘hypothesis’ is used to refer to an explanation of things that occur. In some cases, it may refer to a simple guess. In other instances, it may be a well-developed set of propositions that are crafted to explain the detailed workings of some occurrence or occurrences. One definition states specifically that it is the antecedent to a conditional proposition.

(iv) The hypothesis is formed and tested within the scientific process. One may develop the hypothesis while observation is occurring, but that may also be considered premature. The act of observation (outside of experimentation) may actually present opportunity to disprove a hypothesis. The hypothesis though is necessarily well defined and inclusive of details. This allows for accurate testing. It also, in many cases, distinguishes it from a theory.

(v) The term ‘theory’ is one of a rather scientific nature, but of a less limited nature. Some uses can refer to explanations of occurrences; some do include usage as referencing a simple guess. There is more though. Theory is used
to refer to a branch of study that is focused on the general and conceptual, as compared to the practical and the applied of the same subject. It is significant that a theory is conjectural in nature.

(vi) A hypothesis is a proposed explanation for something. We call it a theory when that hypothesis has been tested with considerable evidence. As a result, a theory is usually a much larger set of statements than a hypothesis because a theory can grow with every new piece of evidence it explains. In other words, a theory can explain far more than the phenomenon it originally was proposed to explain.

(vii) A hypothesis attempts to answer questions by putting forth a plausible explanation that has yet to be rigorously tested. A theory, on the other hand, has already undergone extensive testing by various scientists and is generally accepted as being an accurate explanation of an observation. This does not mean the theory is correct; only that current testing has not yet been able to disprove it, and the evidence as it is understood, appears to support it. A theory will often start out as a hypothesis—an educated guess to explain observable phenomenon. The more a hypothesis is tested and holds up, the better accepted it becomes as a theory.

7.3 FORMULATION OF HYPOTHESIS AND TESTING

A claim or hypothesis about the values or population parameters is known as the Null Hypothesis and is written as $H_0$. In the case of the above discussed situation, our assumption that a butler is innocent would form the null hypothesis and would be stated as follows:

$H_0 = \text{The butler is innocent}$

This hypothesis is then tested with the available evidence and the decision is made whether to accept this hypothesis or reject it. If this hypothesis is rejected, then we accept the alternate hypothesis which is that the butler is not innocent. This alternate hypothesis is denoted as $H_1$ and is stated as:

$H_1 = \text{The butler is not innocent}$

The process involves testing of the null hypothesis. If the null hypothesis is rejected, then the alternate hypothesis is accepted. It should be noted that the acceptance of the alternate hypothesis does not mean that it is correct. It simply means that there is not enough evidence to be reasonably sure that the null hypothesis is acceptable.

As already explained, there are two types of errors that can be used in making decisions regarding accepting or rejecting the null hypothesis. The first type of error, known as Type I error, is used when the null hypothesis is rejected even if it is true. The second type of error, known as Type II error, is used when a null hypothesis is accepted even if it was not true and should have been rejected.
In statistical hypothesis testing and decision-making about the values of population parameters as defined by the sample statistics, the null hypothesis asserts that there is no true difference between the sample statistics and the corresponding population parameter under consideration and if indeed there is any visible difference, it is considered to be due to natural fluctuations in sampling.

To conclude we say that:

- **Null Hypothesis** $H_0$: An assertion about the population parameter that is being tested by the sample results
- **Alternate Hypothesis** $H_1$: A claim about the population parameter that is accepted when the null hypothesis is rejected
- **Type I Error**: An error made in rejecting the null hypothesis, when in fact it is true
- **Type II Error**: An error made in accepting the null hypothesis, when in fact it is false

Type I error is denoted by $\alpha$ (Alpha) and is expressed as a probability of rejecting a true hypothesis. It is also known as the level of significance. $1 - \alpha$ expresses the level of confidence. For example, $\alpha = 0.05$ means that the confidence level is 95% or 0.95.

Type II error is denoted by $\beta$ (Beta) and is expressed as the probability of accepting a false hypothesis. It is desirable to have the $\beta$ value as low as possible for its value reflects the power of the test being performed and a low $\beta$ value indicates that the test of significance is powerful and reliable. (Type I and Type II errors have been discussed in detail in 7.3.2).

### 7.3.1 Procedure for Hypothesis Testing

The general procedure for hypothesis testing consists of the following steps:

1. **State the Null Hypothesis as well as the Alternate Hypothesis**: This means stating the assumed value of the population parameter which is to be tested. For example, suppose we want to test the hypothesis that the average IQ of our college students is 130. Then this would become our null hypothesis and the alternate hypothesis would be that this average IQ is not 130. These statements are expressed as follows:
   
   \[ H_0 : \mu = 130 \]
   
   \[ H_1 : \mu \neq 130 \]

2. **Establish a Level of Significance Prior to Sampling**: The level of significance signifies the probability of committing Type I error $\alpha$ and is generally taken as equal to 0.05, which really means that after the hypothesis has been tested and a decision is made, we will still be making an error in rejecting the null hypothesis when in fact it is true, i.e., 5% of the time. Sometimes the value $\alpha$ is established as 0.01, but it is at the discretion of the investigator to select its value, depending upon the sensitivity of the study.
3. **Determine a Suitable Test Statistic:** This means the choice of appropriate probability distribution to use with the particular available information under consideration. The normal distribution using the Z-score table or the t-distribution is most often used.

4. **Define the Rejection (Critical) Regions:** The critical region will be established on the basis of the choice of the value of the level of significance \( \alpha \). For example, if we select the value of \( \alpha = 0.05 \), and we use the standard normal distribution as our test statistic for testing the population parameter \( \mu \), then as we have discussed before, the difference between the assumption of null hypothesis, assumed value of this population parameter and the value obtained by the analysis of sample results is not expected to be more than \( \pm 1.96 \sigma \) at \( \alpha = 0.05 \). This relationship is shown in the Figure 7.1.

**Fig. 7.1 Rejection Region**

In the above figure, if the sample \( \overline{X} \) statistic falls within \( 1.96\sigma \) of the assumed value of \( \mu \) under the assumption of null hypothesis \( H_0 \), then we accept the null hypothesis as being correct at 95% confidence level (or 0.05 level of significance). The difference between \( \overline{X} \) and \( \mu \) which may be any value between \( X_1 \) and \( \mu \) or \( X_2 \) and \( \mu \) is considered to be accidental or due to chance element and is not considered significant enough or real enough to reject null hypothesis, so that for all practical purposes the value of \( \overline{X} \) is considered equal to \( \mu \) even though \( \overline{X} \) can have any value between \( X_1 \) and \( X_2 \) as shown above. However, if the value of \( \overline{X} \) falls beyond \( X_2 \) on the upper side or beyond \( X_1 \) on the lower side, then this difference between the values of \( \overline{X} \) and \( \mu \) would be considered significant and it will lead to rejection of null hypothesis. Since 5% of the time, this difference between the values of \( \overline{X} \) and \( \mu \) would be significant with 2.5% of the time \( \overline{X} \) being too far above \( \mu \) (beyond \( X_2 \)) and 2.5% of the time being too far below \( \mu \) (below \( X_1 \)), the area of rejection will be on both sides of the mean extending into the tail sections of the curve. This area of rejection is known as the *critical region*.
5. **Data Collection and Sample Analysis:** This involves the actual collection and computation of the sample data. A sample of the pre-established size $n$ is collected and the estimate of the population parameter is calculated. This estimate is the value of the test statistic. For example, if we are testing a hypothesis about the value of population mean $\mu$, then the test statistic would be the sample mean $\bar{x}$. Then we test this statistic to check whether it falls in the critical region or in the acceptance region. For example, if we want to test for the average IQ of the college students to be 130, then in that case, we have to see that our population mean $\mu$ must be tested. We take a random sample of a given size $n$ and calculate its mean $\bar{x}$, and then test it to see if the value of this $\bar{x}$ falls in the area of acceptance or in the area of rejection at a given level of significance.

6. **Making the Decision:** Before the statistical decision is made, a decision rule must be established. Such decision rule will form the basis on which the null hypothesis will be accepted or rejected. This decision rule is really a formal statement of the obvious purpose of the test. For example, this rule could be stated as follows:

   *Accept the null hypothesis if the value of sample statistic $\bar{x}$ falls within the area of acceptance, otherwise reject the null hypothesis.*

Based upon this established decision rule, a decision can be made whether to accept or reject the null hypothesis.

7.3.2 **Committing Errors: Type I and Type II**

**Types of Errors:** There are two types of errors in statistical hypothesis, which are as follows:

- **Type I Error:** In this type of error, you may reject a null hypothesis when it is true. It means rejection of a hypothesis, which should have been accepted. It is denoted by $\alpha$ (alpha), and is also known alpha error.

- **Type II Error:** In this type of error, you are supposed to accept a null hypothesis when it is not true. It means accepting a hypothesis, which should have been rejected. It is denoted by $\beta$ (beta), and is also known as beta error.

Type I error can be controlled by fixing it at a lower level. For example, if you fix it at 2%, then the maximum probability to commit Type I error is 0.02. However, reducing Type I error has a disadvantage when the sample size is fixed as it increases the chances of Type II error. In other words, it can be said that both types of errors cannot be reduced simultaneously. The only solution of this problem is to set an appropriate level by considering the costs and penalties attached to them or to strike a proper balance between both types of errors.

In a hypothesis test, a Type I error occurs when the null hypothesis is rejected when it is in fact true; that is, $H_0$ is wrongly rejected. For example, in a clinical trial
of a new drug, the null hypothesis might be that the new drug is no better, on
average, than the current drug; that is $H_0$: there is no difference between the two
drugs on average. A Type I error would occur if we concluded that the two drugs
produced different effects when in fact there was no difference between them.

In a hypothesis test, a Type II error occurs when the null hypothesis $H_0$ is
not rejected when it is in fact false. For example, in a clinical trial of a new drug,
the null hypothesis might be that the new drug is no better, on average, than the
current drug; that is $H_0$: there is no difference between the two drugs on average.
A Type II error would occur if we concluded that the two drugs produced the
same effect, that is, there is no difference between the two drugs on average,
when in fact, they produced different ones.

In how many ways can we commit errors?

We reject a hypothesis when it may be true. This is Type I Error.
We accept a hypothesis when it may be false. This is Type II Error.

<table>
<thead>
<tr>
<th>Accept $H_0$</th>
<th>Reject $H_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$: True</td>
<td>Accept True $H_0$ Desirable</td>
</tr>
<tr>
<td>$H_1$: False</td>
<td>Accept False $H_0$ Type II Error</td>
</tr>
</tbody>
</table>

The level of significance implies the probability of Type I error. A 5% level implies that
the probability of committing a Type I error is 0.05. A 1% level implies 0.01 probability of
committing Type I error.

Lowering the significance level and, hence, the probability of Type I error is
good, but unfortunately, it would lead to the undesirable situation of committing
Type II error.

To Sum Up:

- **Type I Error**: Rejecting $H_0$ when $H_0$ is True.
- **Type II Error**: Accepting $H_0$ when $H_0$ is False.

*Note:* The probability of making a Type I error is the level of significance of a statistical test.
It is denoted by $\alpha$.

Where, $\alpha = \text{Prob. (Rejecting } H_0 / H_0 \text{ True)}$

$1 - \alpha = \text{Prob. (Accepting } H_0 / H_0 \text{ True)}$

The probability of making a Type II error is denoted by $\beta$.

Where, $\beta = \text{Prob. (Accepting } H_0 / H_1 \text{ False)}$

$1 - \beta = \text{Prob. (Rejecting } H_0 / H_1 \text{ False)} = \text{Prob. (The test correctly rejects } H_0 \text{ when } H_1 \text{ is false).}$
1 – \( \beta \) is called the power of the test. It depends on the level of significance \( \alpha \), sample size \( n \) and the parameter value.

### 7.3.3 Null and Alternative Hypotheses

Hypothesis is usually considered as the principal instrument in research. The basic concepts regarding the testability of a hypothesis are as follows:

**Null Hypothesis and Alternative Hypothesis**

In the context of statistical analysis, while comparing any two methods, the following concepts or assumptions are taken into consideration:

- **Null Hypothesis**: While comparing two different methods in terms of their superiority, wherein the assumption is that both the methods are equally good, it is called null hypothesis. It is also known as statistical hypothesis and is symbolized as \( H_0 \).

- **Alternate Hypothesis**: While comparing two different methods, regarding their superiority, wherein stating a particular method to be good or bad as compared to the other one, it is called alternate hypothesis. It is symbolized as \( H_1 \).

**Comparison of Null Hypothesis with Alternate Hypothesis**

Following are the points of comparison between null hypothesis and alternate hypothesis:

- Null hypothesis is always specific, while alternate hypothesis gives an approximate value.
- The rejection of Null hypothesis involves great risk, which is not in the case of alternate hypothesis.
- Null hypothesis is more frequently used in statistics than alternate hypothesis because it is specific and is not based on probabilities.

The hypothesis to be tested is called the null hypothesis and is denoted by \( H_0 \). This is to be tested against other possible states of nature called alternative hypothesis. The alternative is usually denoted by \( H_1 \).

The null hypothesis implies that there is no difference between the statistic and the population parameter. To test whether there is no difference between the sample mean \( \bar{X} \) and the population \( \mu \), we write the null hypothesis as:

\[
H_0: \bar{X} = \mu
\]

The alternative hypothesis would be:

\[
H_1: \bar{X} \neq \mu
\]

This means \( \bar{X} > \mu \) or \( \bar{X} < \mu \). This is called a two-tailed hypothesis.

The alternative hypothesis \( H_1: \bar{X} > \mu \) is right tailed.

The alternative hypothesis \( H_1: \bar{X} < \mu \) is left tailed.
These are one-sided or one-tailed alternatives.

**Note 1:** The alternative hypothesis $H_1$ implies all such values of the parameter, which are not specified by the null hypothesis $H_0$.

**Note 2:** Testing a statistical hypothesis is a rule, which leads to a decision to accept or reject a hypothesis.

A one-tailed test requires rejection of the null hypothesis when the sample statistic is greater than the population value or less than the population value at a certain level of significance.

1. We may want to test if the sample mean exceeds the population mean $\mu$.
   Then the null hypothesis is:
   $$ H_0: \mu < \mu $$

2. In the other case, the null hypothesis could be:
   $$ H_0: \mu > \mu $$

Each of these two situations leads to a one-tailed test and has to be dealt with in the same manner as the two-tailed test. Here the critical rejection is on one side only, right for $\mu > \mu$ and left for $\mu < \mu$. Both the Figures 7.2 and 7.3 here show a 5% level of test of significance.

For example, a minister in a certain government has an average life of 11 months without being involved in a scam. A new party claims to provide ministers with an average life of more than 11 months without scam. We would like to test if, on the average, the new ministers last longer than 11 months. We may write the Null Hypothesis $H_0: \mu = 11$ and Alternative Hypothesis $H_1: \mu > 11$. 

![Fig. 7.2](image)

**Fig. 7.2** $H_0: \mu < \mu$

![Fig. 7.3](image)

**Fig. 7.3** $H_0: \mu > \mu$
Hypothesis

Check Your Progress

1. Define hypothesis as described by Theodorson and Kerlinger.
2. What do you mean by a good hypothesis?
3. Name the two types of hypothesis.
4. Define theory.
5. What is null hypothesis?
6. Define alternate hypothesis.

7.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. According to Theodorson, ‘a hypothesis is a tentative statement asserting a relationship between certain facts’. Kerlinger describes it as ‘a conjectural statement of the relationship between two or more variables’.
2. A good hypothesis is that which fulfils its intended purposes and is up to the mark. Also, it should be stated in the simplest possible terms and should be in agreement with the observed facts.
3. The two types of hypothesis are: Null Hypothesis and Alternative Hypothesis.
4. Theory is used to refer to a branch of study that is focused on the general and conceptual, as compared to the practical and the applied of the same subject. It is significant that a theory is conjectural in nature.
5. A claim or hypothesis is about the values or population parameters is known as the null hypothesis and is written as \( H_0 \).
6. While comparing two different methods, regarding their superiority, wherein, stating a particular method to be good or bad as compared to the other one, it is called alternate hypothesis. It is symbolized as \( H_1 \).

7.5 SUMMARY

- According to Theodorson, ‘a hypothesis is a tentative statement asserting a relationship between certain facts’. Kerlinger describes it as ‘a conjectural statement of the relationship between two or more variables’.
- A hypothesis specifies the sources of data which shall be studied and in what context they shall be studied.
- Hypothesis is more useful when stated in precise and clearly defined terms. A good hypothesis implies that hypothesis which fulfils its intended purposes and is up to the mark.
- A good hypothesis should be stated in the simplest possible terms. It is also called the principle of the economy or business. It should be clear and precise.
Precise quantitative measurements are more critical in testing theory than qualitative characteristics.

There are generally two types of hypothesis, namely null hypothesis and alternative hypothesis.

When comparing the superiority of both the methods A and B, if we assume that both the methods are equally good, then the assumption is known as ‘null hypothesis’. On the other hand, if we consider method A to be better, it is alternative hypothesis. These may be, symbolically presented as:

Null Hypothesis = Ho
Alternative Hypothesis = Ha

A proposition is a logical statement of relationship between two or more variables which has, generally, been confirmed by empirical research.

Proposition states a relationship between two concepts, and a hypothesis operationalizes this relationship and puts it in an empirically testable form.

The hypothesis is formed and tested within the scientific process. One may develop the hypothesis while observation is occurring, but that may also be considered premature.

The term ‘theory’ is one of a rather scientific nature, but of a less limited nature. Theory is used to refer to a branch of study that is focused on the general and conceptual, as compared to the practical and the applied of the same subject. It is significant that a theory is conjectural in nature.

A theory is usually a much larger set of statements than a hypothesis because a theory can grow with every new piece of evidence it explains. In other words, a theory can explain far more than the phenomenon it originally was proposed to explain.

There are two types of errors that can be used in making decisions regarding accepting or rejecting the null hypothesis. The first type of error, known as Type I error, is used when the null hypothesis is rejected even if it is true. The second type of error, known as Type II error, is used when a null hypothesis is accepted even if it was not true and should have been rejected.

A null hypothesis (H0) is an assertion about the population parameter that is being tested by the sample results.

An alternate hypothesis (H1) is a claim about the population parameter that is accepted when the null hypothesis is rejected.

Type I error is denoted by $\alpha$ (Alpha) and is expressed as a probability of rejecting a true hypothesis. It is also known as the level of significance. Type II error is denoted by $\beta$ (Beta) and is expressed as the probability of accepting a false hypothesis.

Data Collection and sample analysis involves the actual collection and computation of the sample data.

While comparing two different methods in terms of their superiority, wherein the assumption is that both the methods are equally good, it is called null hypothesis. It is also known as statistical hypothesis and is symbolized as $H_0$. 
While comparing two different methods, regarding their superiority, wherein stating a particular method to be good or bad as compared to the other one, it is called alternate hypothesis. It is symbolized as $H_1$.

### 7.6 KEY WORDS

- **Hypothesis:** It refers to the supposition or proposed explanation made on the basis of limited evidence as a starting point for further investigation.
- **Null Hypothesis:** It refers to a general statement or default position that there is no relationship between two measured phenomena, or no association among groups.
- **Alternative Hypothesis:** It refers to the one that states there is a statistically significant relationship between two variables.
- **Hypothesis Testing:** It refers to an act in statistics whereby an analyst tests an assumption regarding a population parameter.

### 7.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Write a short note on the nature of hypothesis.
2. What do you understand by Type I Error and Type II Error?
3. Briefly explain Type I error in a hypothesis test with an example.
4. Differentiate between null hypothesis and alternate hypothesis.

**Long Answer Questions**

1. Discuss some rules for constructing good hypotheses.
2. Examine the steps involved in hypothesis testing.
3. Describe the difference between proposition, hypothesis and theory.

### 7.8 FURTHER READINGS

UNIT 8  PLANNING OF RESEARCH AND RESEARCH DESIGN

Structure
8.0  Introduction
8.1  Objectives
8.2  Planning of Research and its Process
8.3  Research Design: Essentials of a Good Design and its Importance
  8.3.1  Preparation of the Research Design
8.4  Writing the Research Proposal
  8.4.1  Contents of a Research Proposal
  8.4.2  Types of Research Proposals
8.5  Answers to Check Your Progress Questions
8.6  Summary
8.7  Key Words
8.8  Self Assessment Questions and Exercises
8.9  Further Readings

8.0  INTRODUCTION

A research plan is a vital part of the research. With a research plan, a researcher gets an opportunity to discuss the proposed research, its significance and the process in which it will be carried out. A research plan must be designed in a way that it satisfy the questions like the intention of the research, importance of the research, previous findings, and steps need to be taken to complete the research.

In this unit, you will study about the effective planning of research and also the complete process of research. In addition to this, you will also study about the essentials of research design and its importance, preparation and classification of research design. The unit goes on discussing the contents of a research proposal and the types of research proposals.

8.1  OBJECTIVES

After going through this unit, you will be able to:
- Describe the process and planning of research
- Discuss the concept of research design
- Understand the basics of a good research design and its importance
- Explain the preparation and classification of research design
- Explain the contents of a research proposal
- Identify the types of research proposals
8.2 PLANNING OF RESEARCH AND ITS PROCESS

We have placed planning and collecting data for research as simultaneous to the sampling plan. This is because these two—based on the research design—need to be developed concurrently. The reason for this is that the sampling plan helps in identifying the population under study and the data collection plan helps in working out ways of obtaining information from the specified population. There is a huge variety and number of data collection instruments available to the researcher. Broadly, these may be classified into secondary and primary data methods. Each has multiple sub-divisions available. Primary, as the name suggests, is original and collected first hand for the problem under study. There are a variety of primary data methods available to the researcher ranging from subjective, non-quantifiable interviews, focus group discussions, personal/telephonic interviews/mail survey to the well-structured and quantifiable questionnaires. Secondary data is information that has been collected and compiled earlier. For example, company records, magazine articles, expert opinion surveys, sales records, customer feedback, government data and previous researches done on the topic of interest. For example, a study that measures the acceptability of orange-flavoured drink versus natural orange juice by consumers requires empirical and primary information. On the other hand, a descriptive financial investment behaviour study of consumers might be able to make use of secondary data. There are sub-steps involved at this stage—primary data instrument design and pilot testing. For example, if we want to measure the work-family conflict experienced by women in the healthcare sector and the steps that women professionals take to balance this, the study requires empirical data collection and instrument design. Once the instrument has been designed, it has to be tested and refined (pilot testing) before actual data collection can take place. In case a pre-constructed instrument is available and has been developed to measure the specific construct, the two steps of instrument design and testing can be done away with.

This step in the research process requires careful and rigorous quality checks to ensure the reliability and validity of the data collected. There are measurement options available to establish these criteria for the data collection instrument, which have been discussed in the subsequent chapter. Once the instrument is ready, the field work begins and the data is collected from the respondent population based on the devised sampling plan.

8.3 RESEARCH DESIGN: ESSENTIALS OF A GOOD DESIGN AND ITS IMPORTANCE

Once you have established the what of the study, i.e., the research problem, the next step is the how of the study, which specifies the method of achieving the stated research objectives in the best possible manner.
As stated earlier, different paradigms will guide the selection of the gamut of techniques available. These differences in approach have led to varying definitions of what constitutes a research design.

Green et al. (2008) defines research designs as ‘the specification of methods and procedures for acquiring the information needed. It is the overall operational pattern or framework of the project that stipulates what information is to be collected from which sources by what procedures. If it is a good design, it will insure that the information obtained is relevant to the research questions and that it was collected by objective and economical procedures.’

Thyer (1993) states that, ‘A traditional research design is a blueprint or detailed plan for how a research study is to be completed—operationalizing variables so they can be measured, selecting a sample of interest to study, collecting data to be used as a basis for testing hypotheses, and analysing the results.’ The essential requirement of the design is thus to provide a framework and direction to the investigation in the most efficient manner. Sellitz et al. (1962) states that ‘A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure.’

One of the most comprehensive and holistic definition has been given by Kerlinger (1995). He refers to a research design as, ‘… a plan, structure and strategy of investigation so conceived as to obtain answers to research questions or problems. The plan is the complete scheme or programme of the research. It includes an outline of what the investigator will do from writing the hypotheses and their operational implications to the final analysis of data.’

Thus, the formulated design must ensure three basic tenets:

(a) Convert the research question and the stated assumptions/hypotheses into operational variables that can be measured.
(b) Specify the process that would be followed to complete the above task, as efficiently and economically as possible.
(c) Specify the ‘control mechanism(s)’ that would be used to ensure that the effect of other variables that could impact the outcome of the study have been controlled.

The important consideration is that none of these assumptions can be foregone; all of them must be addressed succinctly and adequately in the design for it to be able to lead on to the methods to be used for collecting the problem-specific information. Thus, it follows the problem definition stage and precedes the data collection stage. However, this is not an irreversible step. Sometimes when the researcher is operationally defining the variables for study, it might emerge that the research question needs to be restructured and consecutively the approach for data collection also might oscillate from the quantitative to the qualitative or vice versa.
At this juncture, one needs to understand the distinction between research design and research method. While the design is the specific framework that has been created to seek answers to the research question, the research method is the technique to collect the information required to answer the research problem, given the created framework.

Thus, research designs have a critical and directive role to play in the research process. The execution details of the research question to be investigated are referred to as the research design.

8.3.1 Preparation of the Research Design

Once the researcher has identified the research scope and objectives, he has also established his/her epistemological position. This could be positivistic—in which case the method of enquiry would necessarily be scientific and empirical. Subsequently, this would require a statistical method of analysis (Ackroyd, 1996). The constructivists on the other hand argue for methods that are richer and more applicable to the social sciences, unlike the more pedantic experimental approach. Qualitative is a more definitive choice here than the quantitative (Atkinson and Hammersley, 1994). Yet another approach is the principle of triangulation (Jick, 1979), which advocates the simultaneous or a sequential use of the qualitative and quantitative methods of investigation. The proponents state that when the findings from diverse methods are collated, then the results are richer, more wholistic and this, in turn, improves the sanctity of the analysis.

The formulated research questions are then, through a comprehensive theoretical review, put into a practical perspective. The conceptual design thus developed requires and entails specifications of the variables under study as well as approach to the analysis. This might in turn lead to a refining or rephrasing of the defined research questions. Thus, the formulation of the research design is not a stagnant stage in the research process; rather it is an ongoing backward and forward integrated process by itself.

- An illustration: Let us take the example of the organic food study. The formulated research problem was:

  To investigate the consumer decision-making process for organic food products and to segment the market according to the basket size.

  On conducting an extensive review of the literature, it was found that organic consumption is not always a self-driven choice; rather it could be the seller who might influence the product choice. Thus, a research design was formulated to study the organic consumer’s decision stages. However, once the design is selected and a proposed sampling plan is developed, the next step required is that the constructs and the variables to be studied must be operationalized. On defining the organic consumer, we realized the significance of the psychographics of the individual—the attitude, interest and opinion—which were extremely critical. Thus, to get a wholistic view,
one needs to look at the psychographic profile of the existing consumer, as well as of the potential consumer with a similar mindset. This led to a revision of the research problem:

To investigate the consumer decision-making process for organic food products and to segment the market—existing and potential—according to their psychographic profile.

Classification of Research Designs

The researcher has a number of designs available to him/her for investigating the research objectives. There are various typologies that can be adopted for classifying them. The classification that is universally followed and is simple to comprehend is the one based upon the objective or the purpose of the study. A simple classification that is based upon the research needs ranging from simple and loosely structured to the specific and more formally structured is given in Figure 8.1. This depiction shows the two types of researches—exploratory and conclusive as separate design options, with subcategories in each.

The demarcation between the designs in practice is not this compartmentalized. Thus, a more appropriate approach would be to view the designs on a continuum as in Figure 8.2. Hence, in case the research objective is diffused and requires a fine-tuning and refinement, one uses the exploratory design, this might lead to the slightly more concrete descriptive design—here one describes all the aspects of the constructs and concepts under study. This leads to a more structured and controlled causal research design.

Fig 8.1 Classification of Research Designs
Check Your Progress

1. List some primary data methods available to the researcher.
2. Define research design as explained by Thyer.
3. What is the difference between research design and research method?
4. What role does research designs play in the research process?
5. On what grounds the classification of research designs is done?
6. Name the two types of research designs.

8.4 WRITING THE RESEARCH PROPOSAL

We have learnt that research always begins with a purpose. Either this is the researcher’s own pursuit, or it is carried out to address and answer a specific managerial question and arrive at a solution. This clear statement of purpose guides the research process and must be converted into a plan for the study. This framework or plan is termed as the research proposal. A research proposal is a formal document that presents the research objectives, design of achieving these objectives and the expected outcomes/deliverables of the study.

This step is essential both for academic and corporate research, as it clearly establishes the research process to be followed to address the research questions. In a business or corporate setting, this step is often preceded by a PR (Proposal...
Another advantage of a formal proposal is that sometimes the manager may not be able to clearly tell his problem or the researcher might not be able to understand and convert the decision into a workable research problem. The researcher lists the objectives of the study and then together with the manager, is able to review whether or not the listed objectives and direction of the study will be able to deliver output for arriving at a workable solution.

For the researcher, the document provides an opportunity to identify any shortfalls in the logic or the assumption of the study. It also helps to monitor the methodical work being carried out to accomplish the project.

8.4.1 Contents of a Research Proposal

There is a broad framework that most proposals follow. In this section we will briefly discuss these steps.

Executive summary

This is a broad overview that gives the purpose and objective of the study. In a short paragraph, the author gives a summary about the management problem/academic concern.

Background of the problem

This is the detailed background of the management problem. It requires a sequential and systematic build-up to the research questions and also why the study should be done. The researcher has to be able to demonstrate that there could be a number of ways in which the management dilemma could be answered. For example, a pharmaceutical company develops a new hair growing solution and packages it in two different types of bottles. They want to know which one people will buy. The product testing could be done internally in the company, or the two sample bottles could be formulated and tested for their acceptability amongst likely consumers or retailers keeping the product; or the two types would be developed and test launched and tested for their sales potential. The researcher thus has to spell out all probabilities and then systematically and logically argue for the research study. This section has to be objective and written in simple language, avoiding any metaphors or idioms to dramatize the plan. The logical arguments should speak for themselves and be able to convince the reader of the need for the study in order to find probable solutions to the management dilemma.

Problem statement and research objectives

The clear definition of the problem broken down into specific objectives is the next step. This section is crisp and to the point. It begins by stating the main thrust area of the study. For example, in the above case, the problem statement could be:
To test the acceptability of a spray or capped bottle dispenser for a new hair growing formulation.

The basic objectives of this research would be to:

- Determine the comparative preference of the two prototypes amongst customers of hair growing solutions.
- To conduct a sample usage test of both the bottles with the identified population.
- To assess the ease of use for the bottles amongst the respondents.
- To prepare a comparative analysis of the advantages and problems associated with each bottle, on the basis of the sample usage test.
- To prepare a detailed report on the basis of the findings.

If the study is addressed towards testing some assumptions in the form of hypotheses, they have to be clearly stated in this section.

Research design

This is the working section of the proposal as it needs to indicate the logical and systematic approach intended to be followed in order to achieve the listed objectives. This would include specifying the population to be studied, the sampling process and plan, sample size and selection. It also details the information areas of the study and the probable sources of data, i.e., the data collection methods. In case the process has to include an instrument design, then the intended approach needs to be detailed here. A note of caution has to be given here: this is not a simple statement of the sampling and data collection plan; it requires a clear and logical justification of using the techniques over the methods available for research.

Scheduling the research

The time-bound dissemination of the study with the major phases of the research has to be presented. This can be done using the CPM/GANTT/PERT charts. This gives a clear way for monitoring and managing the research task. It also has the additional benefit of providing the researcher with a means of spelling out the payment points linked to the delivered phase outputs.

Results and outcomes of the research

Here the clear terms of contract or expected outcomes of the study have to be spelt out. This is essential even if it is an academic research. The expected deliverables need to clearly demonstrate how the researcher intends to link the findings of the proposed study design to the stated research objectives. For example, in the pharmaceutical study, the expected deliverables are:

- To identify the usage problems with each bottle type.
- To recommend on the basis of the sample study on which bottle to use for packaging the liquid.
Costing and budgeting the research

In all instances of business research, both internal and external, an estimated cost of the study is required.

In addition to these sections, academic research proposals require a section on review of related literature; this generally follows the ‘problem background’ section. If the proposal is meant to establish the credentials of the research supplier, then detailed qualifications of the research team, including the research experience in the required or related area, help to aid in the selection of the research proposal.

Sometimes, the research study requires an understanding of some technical terms or explanations of the constructs under study; in such cases the researcher needs to attach a glossary of terms in the appendix of the research proposal.

The last section of the proposal is to state the complete details of the references used in the formulation of the research proposal. Thus the data source and address have to be attached with the formulated document.

8.4.2 Types of Research Proposals

Basically, the proposals formulated could be of three types:

- Academic research proposals
- Internal organizational proposal
- External organizational proposals

Academic research proposal

The academic research proposal might be generated by students or academicians pursuing the study for fundamental academic research. These kind of studies need extensive search of past studies and data on the topic of study. An example is an academician wanting to explore the viability of different eco-friendly packaging options available to a manufacturer.

Internal organizational proposal

The internal organizational proposals are conducted within an organization and are submitted to the management for approval and funding. They are of a highly focused nature and are oriented towards solving immediate problems. For example, a pharmaceutical company, which has developed a new hair growing formulation wants to test whether to package the liquid in a spray type or capped dispenser. The solutions are time-driven and applicability is only for this product. These studies do not require extensive literature review but do require clearly stated research objectives, for the management to assess the nature of work required.
External organizational proposals

External organizational proposals have the base or origin within the company, but the scope and nature of the study requires a more structured and objective research. For example, if the above stated pharmaceutical company wishes to explore the herbal cosmetic market and wants market analysis and feasibility study conducted; the PR might be spelt out to solicit proposals to address the research question, and execute an outsourced research.

Check Your Progress

7. What is a research proposal?
8. List the contents of a research proposal.

8.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. There are a variety of primary data methods available to the researcher ranging from subjective, non-quantifiable interviews, focus group discussions, personal/telephonic interviews/mail survey to the well-structured and quantifiable questionnaires.

2. According to Thyer (1993), 'A traditional research design is a blueprint or detailed plan for how a research study is to be completed—operationalizing variables so they can be measured, selecting a sample of interest to study, collecting data to be used as a basis for testing hypotheses, and analysing the results.' The essential requirement of the design is thus to provide a framework and direction to the investigation in the most efficient manner.

3. The research design is the specific framework that has been created to seek answers to the research question, the research method is the technique to collect the information required to answer the research problem, given the created framework.

4. Research designs play a critical and directive role in the research process. The execution details of the research question to be investigated are referred to as the research design.

5. The classification of research designs is based upon the research needs ranging from simple and loosely structured to the specific and more formally structured.

6. The two types of research designs are exploratory research design and conclusive research design.
7. A research proposal is a formal document that presents the research objectives, design of achieving these objectives and the expected outcomes/deliverables of the study.

8. The contents of a research proposal are as follows:
   - Executive Summary
   - Background of the problem
   - Problem statement and research objectives
   - Research design
   - Scheduling the research
   - Results and outcomes of the research
   - Costing and budgeting the research

8.6 SUMMARY

- There are a huge variety and number of data collection instruments available to the researcher. Broadly, these may be classified into secondary and primary data methods. Primary, as the name suggests, is original and collected first hand for the problem under study. There are a variety of primary data methods available to the researcher ranging from subjective, non-quantifiable interviews, focus group discussions, personal/telephonic interviews/mail survey to the well-structured and quantifiable questionnaires. Secondary data is information that has been collected and compiled earlier. For example, company records, magazine articles, expert opinion surveys, sales records, customer feedback, government data and previous researches done on the topic of interest.
- This step in the research process requires careful and rigorous quality checks to ensure the reliability and validity of the data collected. There are measurement options available to establish these criteria for the data collection instrument.
- Green et al. (2008) defines research designs as ‘the specification of methods and procedures for acquiring the information needed. It is the overall operational pattern or framework of the project that stipulates what information is to be collected from which sources by what procedures. If it is a good design, it will insure that the information obtained is relevant to the research questions and that it was collected by objective and economical procedures.’
- One of the most comprehensive and holistic definition has been given by Kerlinger (1995). He refers to a research design as, ‘…… a plan, structure and strategy of investigation so conceived as to obtain answers to research
questions or problems. The plan is the complete scheme or programme of the research. It includes an outline of what the investigator will do from writing the hypotheses and their operational implications to the final analysis of data.'

- One needs to understand the distinction between research design and research method. While the design is the specific framework that has been created to seek answers to the research question, the research method is the technique to collect the information required to answer the research problem, given the created framework.

- Once the researcher has identified the research scope and objectives, he has also established his/her epistemological position. This could be positivistic—in which case the method of enquiry would necessarily be scientific and empirical.

- Once the design is selected and a proposed sampling plan is developed, the next step required is that the constructs and the variables to be studied must be operationalized.

- There are two types of researches—exploratory and conclusive as separate design options, with subcategories in each.

- A research always begins with a purpose; either it is researcher’s own pursuit or it is carried out to address and answer a specific managerial question and arrive a solution. The clear purpose guides the research process and must be converted into a plan for the study. This framework or plan is termed as the research proposal.

- For the researcher, the research proposal document provides an opportunity to identify any shortfalls in the logic or the assumption of the study. It also helps to monitor the methodical work being carried out to accomplish the project.

- The contents of a research proposal include executive summary, background of the problem, problem statement and research objectives, research design, scheduling the research, results and outcomes of the research, costing and budgeting the research.

- Basically, the research proposals formulated could be of three types:
  - Academic research proposals
  - Internal organizational proposals
  - External organizational proposals

- The academic research proposal might be generated by students or academicians pursuing the study for fundamental academic research.

- The internal organizational proposals are conducted within an organization and are submitted to the management for approval and funding. They are of a highly focused nature and are oriented towards solving immediate problems.
• External organizational proposals have the base or origin within the company, but the scope and nature of the study requires a more structured and objective research.

8.7 KEY WORDS

• **Research Design:** It refers to the set of methods and procedures used in collecting and analysing measures of the variables specified in the problem research.

• **Exploratory Research Designs:** It refers to a research conducted for a problem that has not been studied more clearly, intended to establish priorities, develop operational definitions and improve the final research design.

• **Conclusive Research Designs:** It is meant to provide information that is useful in reaching conclusions or decision-making.

• **Research Proposal:** It refers to a document proposing a research project and generally constitutes a request for sponsorship of that research.

8.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Write a note on research process.
2. Mention the three basic tenets for a formulated research design.
3. How does a research proposal help a researcher?
4. Briefly discuss the basic objectives of the research.

**Long Answer Questions**

1. What is a research design process? Explain.
2. Describe the entire process or steps involved in writing a research proposal.
3. Discuss the three types of research proposals in detail.

8.9 FURTHER READINGS


NOTES


UNIT 9  TYPES OF RESEARCH DESIGNS

Structure
9.0 Introduction
9.1 Objectives
9.2 Exploratory Research Design
   9.2.1 Secondary Resource Analysis
   9.2.2 Case Study Method
   9.2.3 Expert Opinion Survey
   9.2.4 Focus Group Discussions
9.3 Descriptive Research Design
   9.3.1 Cross-sectional Studies
   9.3.2 Longitudinal Studies
9.4 Experimental Designs
9.5 Diagnostic Research Design
9.6 Answers to Check Your Progress Questions
9.7 Summary
9.8 Key Words
9.9 Self Assessment Questions and Exercises
9.10 Further Readings

9.0 INTRODUCTION

As we know, a research design is a set of methods and procedures used in collecting and analysing measures of the variables specified. In this unit, you will study about the different types of research designs like exploratory research design, descriptive research design, experimental designs and diagnostic research design. Besides the definitions of different types of research designs, you will also study about their categories and sub-categories.

9.1 OBJECTIVES

After going through this unit, you will be able to:
- Describe the different types of research designs
- Discuss the purpose and characteristics of exploratory research design
- Understand the objective and sub-categories of descriptive research design
Types of Research Designs

- Explain the conditions and types of experimental designs
- Define diagnostic research design

NOTES

9.2 EXPLORATORY RESEARCH DESIGN

Exploratory designs, as stated earlier, are the simplest and most loosely structured designs. As the name suggests, the basic objective of the study is to explore and obtain clarity about the problem situation. It is flexible in its approach and mostly involves a qualitative investigation. The sample size is not strictly representative and at times it might only involve unstructured interviews with a couple of subject experts. The essential purpose of the study is to:

- Define and understand the research problem to be investigated.
- Explore and evaluate the diverse and multiple research opportunities.
- Assist in the development and formulation of the research hypotheses.
- Define the variables and constructs under study.
- Identify the possible nature of relationships that might exist between the variables under study.
- Explore the external factors and variables that might impact the research.

For example, a university professor might decide to do an exploratory analysis of the new channels of distribution that are being used by the marketers to promote and sell products and services. To do this, a structured and defined methodology might not be essential as the basic objective is to understand how to teach this to students of marketing. The researcher can make use of different methods and techniques in an exploratory research - like secondary data sources, unstructured or structured observations, expert interviews and focus group discussions with the concerned respondent group. Here, we will discuss them in brief in the light of their use in exploratory research.

9.2.1 Secondary Resource Analysis

Secondary sources of data, as the name suggests, are data in terms of the details of previously collected findings in facts and figures—which have been authenticated and published. It is a fast and inexpensive way of collecting information. The past details can sometimes point out to the researcher that his proposed research is redundant and has already been established earlier. Secondly, the researcher might find that a small but significant aspect of the concept has not been addressed and should be studied. For example, a marketer might have extensively studied the potential of the different channels of communication for promoting a ‘home maintenance service’ in Greater Mumbai. However, there is no impact of any mix that he has tested. An anthropologist research associate, on going through the findings, postulated the need for studying the potential of WOM (word of mouth) in a close-knit and predominantly Parsi colony where this might be the most effective
culture-dependent technique that would work. Thus, such insights might provide leads for carrying out an experimental and conclusive research subsequently.

Another valuable secondary resource is the compiled and readily available databases of the entire industry, business, or construct. These might be available on free and public domains or through a structured acquisition process and cost. These are both government and non-government publications. Based on the resources and the level of accuracy required, the researcher might decide to make use of them.

9.2.2 Case Study Method

Another way of conducting an exploratory research is the case study method. This requires an in-depth study and is focused on a single unit of analysis. This unit could be an employee or a customer; an organization or a complete country analysis. They are by their nature, generally, post-hoc studies and report those incidences which might have occurred earlier. The scenario is reproduced based upon the secondary information and a primary interview/discussion with those involved in the occurrence. Thus, there might be an element of bias as the data, in most cases, becomes a judgemental analysis rather than a simple recounting of events.

For example, BCA Corporation wants to implement a performance appraisal system in the organization and is debating between the merits of a traditional appraisal system and a 360° appraisal system. For a historical understanding of the two techniques, the HR director makes use of books on the subject. However, for better understanding, he should do an in-depth case accounting of Allied Association which had implemented traditional appraisal formats, and Surakhsha International which uses 360° appraisal systems. Thus, the two exploratory researches carried out were sufficient to arrive at a decision in terms of what would be best for the organization.

9.2.3 Expert Opinion Survey

At times, there might be a situation when the topic of a research is such that there is no previous information available on it. In these cases, it is advisable to seek help from experts who might be able to provide some valuable insights based upon their experience in the field or with the concept. This approach of collecting particulars from significant and knowledgeable people is referred to as the expert opinion survey. This methodology might be formal and structured and is useful when authenticated or supported by a secondary/primary research or it might be fluid and unstructured and might require an in-depth interviewing of the expert. For example, the evaluation of the merit of marketing organic food products in the domestic Indian market cannot be done with the help of secondary data as no such structured data sources exist. In this case the following can be contacted:

- Doctors and dieticians as experts would be able to provide information whether consumers would eat organic food products as a healthier alternative.
- Chefs who are experimental and would like to look at providing better value to their clients.
- Retailers who like to sell contemporary new products.
Types of Research Designs

These could be useful in measuring the viability of the proposed plan. Discussions with knowledgeable people may reveal some information regarding who might be considered as potential consumers. Secondly, the question whether a healthy proposition or a lifestyle proposition would work better to capture the targeted consumers needs to be examined. Thus, this method can play a directional role in shaping the research study.

9.2.4 Focus Group Discussions

Another way to conduct an exploratory analysis is carry out discussions with individuals associated with the problem under study. This technique, though originally from sociology, is actively used in business research. In a typical focus group, there is a carefully selected small set of individuals representative of the larger respondent population under study. It is called a focus group as the selected members discuss the concerned topic for the duration of 90 minutes to, sometimes, two hours. Usually the group is made up of six to ten individuals. The number thus stated is because less than six would not be able to throw enough perspectives for the discussion and there might emerge a one-sided discussion on the topic. On the other hand, more than ten might lead to more confusion rather than any fruitful discussion and that would be unwieldy to manage. Generally, these discussions are carried out in neutral settings by a trained observer, also referred to as the moderator. The moderator, in most cases, does not participate in the discussion. His prime objective is to manage a relatively non-structured and informal discussion. He initiates the process and then maneuvers it to steer it only to the desired information needs. Sometimes, there is more than one observer to record the verbal and non-verbal content of the discussion. The conduction and recording of the dialogue requires considerable skill and behavioural understanding and the management of group dynamics. In the organic food product study, the focus group discussions were carried out with the typical consumers/buyers of grocery products. The objective was to establish the level of awareness about health hazards, environmental concerns and awareness of organic food products. A series of such focus group discussions carried out across four metros—Delhi, Mumbai, Bengaluru and Hyderabad—revealed that even though the new age consumer was concerned about health, the awareness about organic products varied from extremely low to non-existent. (This study was carried out in the year 2004–05 by one of the authors for an NGO located in Delhi).

Check Your Progress

1. What is the basic objective of exploratory research design?
2. What do you understand by secondary sources of data?
3. List the characteristics of a case study method.
4. What is the motive of expert opinion survey?
9.3 DESCRIPTIVE RESEARCH DESIGN

As the name implies, the objective of descriptive research studies is to provide a comprehensive and detailed explanation of the phenomena under study. The intended objective might be to give a detailed sketch or profile of the respondent population being studied. For example, to design an advertising and sales promotion campaign for high-end watches, a marketer would require a holistic profile of the population that buys such luxury products. Thus a descriptive study, (which generates data on who, what, when, where, why and how of luxury accessory brand purchase) would be the design necessary to fulfill the research objectives.

Descriptive research thus are conclusive studies. However, they lack the precision and accuracy of experimental designs, yet it lends itself to a wide range of situations and is more frequently used in business research. Based on the time period of the collection of the research information, descriptive research is further subdivided into two categories: cross-sectional studies and longitudinal studies.

9.3.1 Cross-sectional Studies

As the name suggests, cross-sectional studies involve a slice of the population. Just as in scientific experiments one takes a cross-section of the leaf or the cheek cells to study the cell structure under the microscope, similarly one takes a current subdivision of the population and studies the nature of the relevant variables being investigated.

There are two essential characteristics of cross-sectional studies:

- The cross-sectional study is carried out at a single moment in time and thus the applicability is most relevant for a specific period. For example, one cross-sectional study was conducted in 2002 to study the attitude of Americans towards Asian-Americans, after the 9/11 terrorist attack. This revealed the mistrust towards Asians. Another cross-sectional study conducted in 2012 to study the attitude of Americans towards Asian-Americans revealed more acceptance and less mistrust. Thus the cross-sectional studies cannot be used interchangeably.

- Secondly, these studies are carried out on a section of respondents from the population units under study (e.g., organizational employees, voters, consumers, industry sectors). This sample is under consideration and under investigation only for the time coordinate of the study.

There are also situations in which the population being studied is not of a homogeneous nature but composed of different groups. Thus it becomes essential to study the sub-segments independently. This variation of the design is termed as multiple cross-sectional studies. Usually this multi-sample analysis is carried out at the same moment in time. However, there might be instances when the data is obtained from different samples at different time intervals and then they are compared. Cohort analysis is the name given to such cross-sectional surveys conducted on different sample groups at different time intervals. Cohorts are...
9.3.2 Longitudinal Studies

A single sample of the identified population that is studied over a longer period of time is termed as a longitudinal study design. A panel of consumers specifically chosen to study their grocery purchase pattern is an example of a longitudinal design. There are certain distinguishing features of the same:

- The study involves the selection of a representative panel, or a group of individuals that typically represent the population under study.
- The second feature involves the repeated measurement of the group over fixed intervals of time. This measurement is specifically made for the variables under study.
- A distinguishing and mandatory feature of the design is that once the sample is selected, it needs to stay constant over the period of the study. That means the number of panel members has to be the same. Thus, in case a panel member due to some reason leaves the panel, it is critical to replace him/her with a representative member from the population under study.

Longitudinal study using the same section of respondents thus provides more accurate data than one using a series of different samples. These kinds of panels are defined as true panels and the ones using a different group every time are called omnibus panels. The advantages of a true panel are that it has a more committed sample group that is likely to tolerate extended or long data collecting sessions. Secondly, the profile information is a one-time task and need not be collected every time. Thus, a useful respondent time can be spent on collecting some research-specific information.

However, the problem is getting a committed group of people for the entire study period. Secondly, there is an element of mortality and attrition where the members of the panel might leave midway and the replaced new recruits might be vastly different and could skew the results in an absolutely different direction. A third disadvantage is the highly structured study situation which might be responsible for a consistent and structured behaviour, which might not be the case in the real or field conditions.
5. What types of questions are answered in descriptive research study?

6. Define cohort analysis.

9.4 EXPERIMENTAL DESIGNS

Experimental designs are conducted to infer causality. In an experiment, a researcher actively manipulates one or more causal variables and measures their effects on the dependent variables of interest. Since any changes in the dependent variable may be caused by a number of other variables, the relationship between cause and effect often tends to be probabilistic in nature. It is virtually impossible to prove a causality. One can only infer a cause-and-effect relationship.

The necessary conditions for making causal inferences are: (i) concomitant variation, (ii) time order of occurrence of variables and (iii) absence of other possible causal factors. The first condition implies that cause and effect variables should have a high correlation. The second condition means that causal variable must occur prior to or simultaneously with the effect variable. The third condition means that all other variable except the one whose influence we are trying to study should be absent or kept constant.

There are two conditions that should be satisfied while conducting an experiment. These are:

(i) Internal validity: Internal validity tries to examine whether the observed effect on a dependent variable is actually caused by the treatments (independent variables) in question. For an experiment to be possessing internal validity, all the other causal factors except the one whose influence is being examined should be absent. Control of extraneous variables is a necessary condition for inferring causality. Without internal validity, the experiment gets confounded.

(ii) External validity: External validity refers to the generalization of the results of an experiment. The concern is whether the result of an experiment can be generalized beyond the experimental situations. If it is possible to generalize the results, then to what population, settings, times, independent variables and the dependent variables can the results be projected. It is desired to have an experiment that is valid both internally and externally. However, in reality, a researcher might have to make a trade-off between one type of validity for another. To remove the influence of an extraneous variable, a researcher may set up an experiment with artificial setting, thereby increasing its internal validity. However, in the process the external validity will be reduced.

There are four types of experimental designs. These are explained below:

1. Pre-experimental designs: There are three designs under this – one short case study where observation is taken after the application of treatment, one
Types of Research Designs

1. **Self-instructional Material**
   - Group pre-test-post test design: Where one observation is taken prior to the application of treatment and the other one after the application of treatment, and static group comparison, where there are two groups – experimental group and control group. The experiment group is subjected to treatment and a post test measurement is taken. In the control group measurement is taken at the time when it was done for experimental group. These do not make use of any randomization procedures to control the extraneous variables. Therefore, the internal validity of such designs is questionable.

2. **Quasi-experimental designs**: In these designs the researcher can control when measurements are taken and on whom they are taken. However, this design lacks complete control of scheduling of treatment and also lacks the ability to randomize test units’ exposure to treatments. As the experimental control is lacking, the possibility of getting confounded results is very high. Therefore, the researchers should be aware of what variables are not controlled and the effects of such variables should be incorporated into the findings.

3. **True experimental designs**: In these designs, researchers can randomly assign test units and treatments to an experimental group. Here, the researcher is able to eliminate the effect of extraneous variables from both the experimental and control group. Randomization procedure allows the researcher the use of statistical techniques for analysing the experimental results.

4. **Statistical designs**: These designs allow for statistical control and analysis of external variables. The main advantages of statistical design are the following:
   - The effect of more than one level of independent variable on the dependent variable can be manipulated.
   - The effect of more than one independent variable can be examined.
   - The effect of specific extraneous variable can be controlled.

Statistical design includes the following designs:

   (i) **Completely randomized design**: This design is used when a researcher is investigating the effect of one independent variable on the dependent variable. The independent variable is required to be measured in nominal scale i.e. it should have a number of categories. Each of the categories of the independent variable is considered as the treatment. The basic assumption of this design is that there are no differences in the test units. All the test units are treated alike and randomly assigned to the test groups. This means that there are no extraneous variables that could influence the outcome.

   Suppose we know that the sales of a product is influenced by the price level. In this case, sales are a dependent variable and the price is the independent variable. Let there be three levels of price, namely, low, medium and high. We wish to determine the most effective price level i.e. at which price level the sale is highest. Here, the test units are the stores which are randomly assigned to the three treatment level. The average sales for each price level is computed and examined to
Types of Research Designs

see whether there is any significant difference in the sale at various price levels. The statistical technique to test for such a difference is called analysis of variance (ANOVA).

The main limitation of completely randomized designs is that it does not take into account the effect of extraneous variables on the dependent variable. The possible extraneous variables in the present example could be the size of the store, the competitor’s price and price of the substitute product in question. This design assumes that all the extraneous factors have the same influence on all the test units which may not be true in reality. This design is very simple and inexpensive to conduct.

(ii) Randomized block design: As discussed, the main limitation of the completely randomized design is that all extraneous variables were assumed to be constant over all the treatment groups. This may not be true. There may be extraneous variables influencing the dependent variable. In the randomized block design it is possible to separate the influence of one extraneous variable on a particular dependent variable, thereby providing a clear picture of the impact of treatment on test units.

In the example considered in the completely randomized design, the price level (low, medium and high) was considered as an independent variable and all the test units (stores) were assumed to be more or less equal. However, all stores may not be of the same size and, therefore, can be classified as small, medium and large size stores. In this design, the extraneous variable, like the size of the store could be treated as different blocks. Now the treatments are randomly assigned to the blocks in such a way so that each treatment appears in each block at least once. The purpose of forming these blocks is that it is hoped that the scores of the test units within each block would be more or less homogeneous when the treatment is absent. What is assumed here is that block (size of the store) is correlated with the dependent variable (sales). It may be noted that blocking is done prior to the application of the treatment.

In this experiment one might randomly assign 12 small-sized stores to three price levels in such a way that there are four stores for each of the three price levels. Similarly, 12 medium-sized stores and 12 large-sized stores may be randomly assigned to three price levels. Now the technique of analysis of variance could be employed to analyse the effect of treatment on the dependent variable and to separate out the influence of extraneous variable (size of store) from the experiment.

(iii) Factorial design: A factorial design may be employed to measure the effect of two or more independent variables at various levels. The factorial designs allow for interaction between the variables. An interaction is said to take place when the simultaneous effect of two
or more variables is different from the sum of their individual effects. An individual may have a high preference for mangoes and may also like ice-cream, which does not mean that he would like mango ice cream, leading to an interaction.

The sales of a product may be influenced by two factors, namely, price level and store size. There may be three levels of price—low ($A_1$), medium ($A_2$) and high ($A_3$). The store size could be categorized into small ($B_1$) and big ($B_2$). This could be conceptualized as a two-factor design with information reported in the form of a table. In the table, each level of one factor may be presented as a row and each level of another variable would be presented as a column. This example could be summarized in the form of a table having three rows and two columns. This would require $3 \times 2 = 6$ cells. Therefore, six different levels of treatment combinations would be produced each with a specific level of price and store size. The respondents would be randomly selected and randomly assigned to the six cells.

The tabular presentation of $3 \times 2$ factorial design is given in Table 9.1.

<table>
<thead>
<tr>
<th>Price</th>
<th>Store</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Level ($A_1$)</td>
<td>$B_1$</td>
</tr>
<tr>
<td>Medium Level ($A_2$)</td>
<td>$B_1$</td>
</tr>
<tr>
<td>High Level ($A_3$)</td>
<td>$B_1$</td>
</tr>
</tbody>
</table>

Respondents in each cell receive a specified treatment combination. For example, respondents in the upper left hand corner cell would face small level of price and small store. Similarly, the respondents in the lower right hand corner cell will be subjected to both high price level and big store.

The main advantages of factorial design are:

- It is possible to measure the main effects and interaction effect of two or more independent variables at various levels.
- It allows a saving of time and effort because all observations are employed to study the effects of each factor.
- The conclusion reached using factorial design has broader applications as each factor is studied with different combinations of other factors.

The limitation of this design is that the number of combinations (number of cells) increases with increased number of factors and levels. However, a fractional
factorial design could be used if interest is in studying only a few of the interactions or main effects.

**Errors Affecting Research Design**

We have discussed three types of research designs, namely, exploratory, descriptive and experimental. All of these have some scope of error. There could be various sources of errors in research design.

Exploratory research is conducted using focus group discussion, secondary data, analysis of case study and expert opinion survey. It is quite likely that members of the focus group have not been selected properly. Secondary data may not be free from errors (in fact, one needs to evaluate the methodology used in collecting such data). Also, the experts chosen for the survey may not be experts in the field. As a matter of fact, getting an expert is a very difficult task. All these factors could lead to errors in the exploratory design.

In the descriptive design, the purpose is to describe a phenomenon. For this one could use a structured questionnaire. It could always happen that the respondents do not give correct responses to some of the questions, thereby resulting in wrong information.

In the true experimental design and statistical design, the respondents are selected at random which may not be the case in real life. Many times, in actual business situation, the value judgements play a very important role in selecting the respondents. Further, there can always be errors in observations.

### 9.5 DIAGNOSTIC RESEARCH DESIGN

In this type of research design, the researcher wants to know about the foundational causes of the problem. The researcher describes the factors responsible for the problematic situation. Therefore it can be said that diagnostic research design is a problem-solving research design that comprises mainly:

- Emergence of the problem
- Diagnosis of the problem
- Solution for the problem and
- Suggestion for the problem solution

**Check Your Progress**

- Name the type of experimental design in which researchers can randomly assign test units and treatments to an experimental group.
- State the limitation of the factorial design.
Types of Research Designs

9.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

NOTES

1. The basic objective of the exploratory research design is to explore and obtain clarity about the problem situation. It is flexible in its approach and mostly involves a qualitative investigation.

2. The secondary sources of data are the details of previously collected findings in facts and figures—which have been authenticated and published. It is a fast and inexpensive way of collecting information.

3. The main characteristics of case study method are that it is narrowly focused, provides high level of detail, in-depth understanding. Generally, they are post-hoc studies and report by nature which are based upon the secondary information.

4. An expert opinion survey is aimed to collect particulars from significant and knowledgeable people. This methodology might be formal and structured and is useful when authenticated or supported by a secondary/primary research or it might fluid or unstructured or require an in-depth interviewing of the expert.

5. Descriptive studies are conclusive studies and it collects data that are used to answer a wide range of who, what, when, why, where and how questions pertaining to a particular population or group.

6. Cohort analysis is the name given to such cross-sectional surveys conducted on different sample groups at different time intervals. Cohorts are essentially groups of people who share a time zone or have experienced an event that took place at a particular time period.

7. True experimental designs are the type of experimental design in which researchers can randomly assign test units and treatments to an experimental group.

8. The limitation of the factorial design is that the number of combinations increases with the increased number of factors and levels.

9.7 SUMMARY

- Exploratory designs, as stated earlier, are the simplest and most loosely structured designs. As the name suggests, the basic objective of the study is to explore and obtain clarity about the problem situation. It is flexible in its approach and mostly involves a qualitative investigation.

- Secondary sources of data, as the name suggests, are data in terms of the details of previously collected findings in facts and figures—which have been authenticated and published.
• Another way of conducting an exploratory research is the case study method. This requires an in-depth study and is focused on a single unit of analysis.

• At times, there might be a situation when the topic of a research is such that there is no previous information available on it. In these cases, it is advisable to seek help from experts who might be able to provide some valuable insights based upon their experience in the field or with the concept.

• Another way to conduct an exploratory analysis is carry out discussions with individuals associated with the problem under study. This technique, though originally from sociology, is actively used in business research. In a typical focus group, there is a carefully selected small set of individual’s representative of the larger respondent population under study. It is called a focus group as the selected members discuss the concerned topic for the duration of 90 minutes to, sometimes, two hours.

• As the name implies, the objective of descriptive research studies is to provide a comprehensive and detailed explanation of the phenomena under study. The intended objective might be to give a detailed sketch or profile of the respondent population being studied.

• The conduction and recording of the dialogue requires considerable skill and behavioural understanding and the management of group dynamics. In the organic food product study, the focus group discussions were carried out with the typical consumers/buyers of grocery products.

• As the name implies, the objective of descriptive research studies is to provide a comprehensive and detailed explanation of the phenomena under study. The intended objective might be to give a detailed sketch or profile of the respondent population being studied.

• A descriptive study, (which generates data on who, what, when, where, why and how of luxury accessory brand purchase) would be the design necessary to fulfill the research objectives.

• Descriptive research thus are conclusive studies. However, they lack the precision and accuracy of experimental designs, yet it lends itself to a wide range of situations and is more frequently used in business research.

• The cross-sectional study is carried out at a single moment in time and thus the applicability is most relevant for a specific period.

• Cross-sectional studies are carried out on a section of respondents from the population units under study (e.g., organizational employees, voters, consumers, industry sectors). This sample is under consideration and under investigation only for the time coordinate of the study.

• There are also situations in which the population being studied is not of a homogeneous nature but composed of different groups. Thus, it becomes essential to study the sub-segments independently. This variation of the design is termed as multiple cross-sectional studies.
Types of Research Designs

- **Cohort analysis** is the name given to such cross-sectional surveys conducted on different sample groups at different time intervals. Cohorts are essentially groups of people who share a time zone or have experienced an event that took place at a particular time period.

- A single sample of the identified population that is studied over a longer period of time is termed as a longitudinal study design. A panel of consumers specifically chosen to study their grocery purchase pattern is an example of a longitudinal design.

- Longitudinal study using the same section of respondents thus provides more accurate data than one using a series of different samples. These kinds of panels are defined as true panels and the ones using a different group every time are called omnibus panels.

- Experimental designs are conducted to infer causality. In an experiment, a researcher actively manipulates one or more causal variables and measures their effects on the dependent variables of interest. Since any changes in the dependent variable may be caused by a number of other variables, the relationship between cause and effect often tends to be probabilistic in nature. It is virtually impossible to prove a causality. One can only infer a cause-and-effect relationship.

- The necessary conditions for making causal inferences are: (i) concomitant variation, (ii) time order of occurrence of variables and (iii) absence of other possible causal factors.

- The first condition implies that cause and effect variables should have a high correlation. The second condition means that causal variable must occur prior to or simultaneously with the effect variable.

- Internal validity tries to examine whether the observed effect on a dependent variable is actually caused by the treatments (independent variables) in question.

- External validity refers to the generalization of the results of an experiment. The concern is whether the result of an experiment can be generalized beyond the experimental situations. If it is possible to generalize the results, then to what population, settings, times, independent variables and the dependent variables can the results be projected.

- There are three designs under pre-experimental designs – one short case study where observation is taken after the application of treatment, one group pre test-post test design where one observation is taken prior to the application of treatment and the other one after the application of treatment, and static group comparison, where there are two groups – experimental group and control group.

- In Quasi-experimental designs, the researcher can control when measurements are taken and on whom they are taken. However, this design
lacks complete control of scheduling of treatment and also lacks the ability to randomize test units’ exposure to treatments.

- In true experimental designs, researchers can randomly assign test units and treatments to an experimental group. Here, the researcher is able to eliminate the effect of extraneous variables from both the experimental and control group. Randomization procedure allows the researcher the use of statistical techniques for analysing the experimental results.

- Statistical designs allow for statistical control and analysis of external variables. Statistical designs includes different designs, namely completely randomized design, randomized block design, and factorial design.

- Completely randomized design is used when a researcher is investigating the effect of one independent variable on the dependent variable.

- In the randomized block design it is possible to separate the influence of one extraneous variable on a particular dependent variable, thereby providing a clear picture of the impact of treatment on test units.

- A factorial design may be employed to measure the effect of two or more independent variables at various levels. The factorial designs allow for interaction between the variables. An interaction is said to take place when the simultaneous effect of two or more variables is different from the sum of their individual effects.

- Exploratory research is conducted using focus group discussion, secondary data, analysis of case study and expert opinion survey. It is quite likely that members of the focus group have not been selected properly.

- In the descriptive design, the purpose is to describe a phenomenon. For this one could use a structured questionnaire. It could always happen that the respondents do not give correct responses to some of the questions, thereby resulting in wrong information.

- In the true experimental design and statistical design, the respondents are selected at random which may not be the case in real life.

- Many a times, in actual business situation, the value judgements play very important role in selecting the respondents. Further, there can always be errors in observations.

### 9.8 KEY WORDS

- **Exploratory Research**: It refers to a research used to investigate a problem which is not clearly defined. It is conducted to have better understanding of the existing problem.

- **Descriptive Research Design**: It is used to describe characteristics of a population or phenomenon being studied.
Types of Research Designs

- **Longitudinal Study**: It refers to a research design that involves repeated observations of the same variables over short or long periods of time.
- **Quasi-experimental Designs**: It refers to an empirical interventional study used to estimate the causal impact of an intervention on target population without random assignment.

9.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Write a short note on focus group discussions.
2. What are the two essential characteristics of cross-sectional studies?
3. Briefly mention about the longitudinal study design.
4. Distinguish between internal and external validity of the experiments.
5. Write short notes on:
   - Completely randomized design
   - Randomized block design
   - Factorial Design
6. What is diagnostic research design?

**Long Answer Questions**

1. What are exploratory designs? What are the methods that can be used in an exploratory design?
2. Discuss the two categories of descriptive research design in detail.
3. Explain the four types of experimental designs.

9.10 FURTHER READINGS


UNIT 10 DATA COLLECTION

Structure
10.0 Introduction
10.1 Objectives
10.2 Primary and Secondary Data
10.3 Methods of Data Collection
   10.3.1 Observation
   10.3.2 Interview
   10.3.3 Questionnaire
10.4 Sampling: Introduction
10.5 Scaling and Checklist
10.6 Library Records and Reports
10.7 Answers to Check Your Progress Questions
10.8 Summary
10.9 Key Words
10.10 Self Assessment Questions and Exercises
10.11 Further Readings

10.0 INTRODUCTION

Data collection is a vital aspect of a research study as inaccurate or improper data collection can inversely effect the results of a study and lead to invalid results. Several methods of data collection are employed in collecting different types of data. In the same line, the unit first gives a slight insight into the types of data collection like primary and secondary data and then goes on discussing various methods of data collection such as observation, interviews, and questionnaires. Apart from their definition, the unit will also describe their characteristics, advantages and disadvantages, and uses of respective methods in research. In addition to this, you will also learn about sampling design, checklists, and the meaning and importance of library records and reports.

10.1 OBJECTIVES

After going through this unit, you will be able to:
- Understand the concept of data collection
- Discuss the primary and secondary data collection
- Analyse the various methods of data collection like observation, interview, and questionnaire
- Comprehend the design, procedure, implications, and advantages of sampling
- Describe the concept of checklists
- Assess the meaning, importance of library records and reports
10.2 PRIMARY AND SECONDARY DATA

Primary data, as the name suggests, is original, problem- or project-specific and collected for the specific objectives and needs spelt out by the researcher. The accuracy and relevance is reasonably high. The time and money required for this are quite high and sometimes a researcher might not have the resources or the time or both to go ahead with this method. In this case, the researcher can look at alternative sources of data which are economical and reliable enough to take the study forward. These include the second category of data sources—namely the secondary data.

Secondary data as the name implies is that information which is not topical or research-specific and has been collected and compiled by some other researcher or investigative body. This type of data is recorded and published in a structured format, and thus, is quicker to access and manage. Secondly, in most instances, unless it is a data product, it is not too expensive to collect. The information required is readily available as a data product or as the audit information which the researcher or the organization can get and use it for arriving at quick decisions. In comparison to the original research-centric data, secondary data can be economically and quickly collected by the decision maker in a short span of time. However, one must remember that this is a little low on accuracy as what is primary and original for one researcher would essentially become secondary and historical for someone else.

Table 10.1 gives a snapshot of the major differences between the two methods.

Table 10.1 Primary vs Secondary Data

<table>
<thead>
<tr>
<th></th>
<th>Primary Data</th>
<th>Secondary Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collection purpose</td>
<td>For the problem at hand</td>
<td>For other problems</td>
</tr>
<tr>
<td>Collection process</td>
<td>Very involved</td>
<td>Rapid &amp; easy</td>
</tr>
<tr>
<td>Collection cost</td>
<td>High</td>
<td>Relatively low</td>
</tr>
<tr>
<td>Collection time</td>
<td>Long</td>
<td>Short</td>
</tr>
<tr>
<td>Accuracy</td>
<td>As Desired</td>
<td>Not Known</td>
</tr>
<tr>
<td>Control</td>
<td>Complete</td>
<td>No Control</td>
</tr>
<tr>
<td>Period</td>
<td>Up to date</td>
<td>Dated</td>
</tr>
</tbody>
</table>

10.3 METHODS OF DATA COLLECTION

This section discusses the various methods of data collection in detail.

10.3.1 Observation

Observation can be defined as viewing or seeing. Observation means specific viewing with the purpose of gathering the data for a specific research study.
Data Collection

Observation is a classical method of scientific study. It is very important in any research study as it is an effective method for data collection.

Characteristics of Observation Method

The following are the characteristics of observation method of data collection:

- **Physical and mental activity**: Eyes observe so many things in our surroundings but our focus or attention is only on data which is relevant to research study.

- **Observation is selective**: It is very difficult for a researcher to observe everything in his surroundings. He only observes the data which is purposive for his research study and meets with the scope of his study. The researcher ignores all the data which is not relevant to the study.

- **Observation is purposive and not casual**: Observation is purposive as it is relevant to a particular study. The purpose of observation is to collect data for the research study. It focuses on human behaviour which occurs in a social phenomenon. It analyses the relationship of different variables in a specific context.

- **Accuracy and standardization**: Observation of pertinent data should be accurate and standardized for its applications.

Types of Observation

Different concepts define the classification of observations.

With respect to an investigator’s role, observation may be:

- Participant observation
- Non-participant observation

With respect to the method of observation, it can be classified into the following:

- Direct observation
- Indirect observation

With reference to the control on the system to be observed, observation can be classified into the following:

- Controlled observation
- Uncontrolled observation

(i) Participant observation

In participant type of observation, the observer is an active participant of the group or process. He participates as well as observes as a part of phenomenon; for example, to study the behaviour of management students towards studying and understanding marketing management, the observer or researcher has to participate in the discussion with students without telling them about the observation.
or purpose. When respondents are unaware of observations, then only their natural interest can be studied.

**Advantages**

The following are the main advantages of participant observation:

- In-depth understanding of the respondent group.
- The context which is meaningful to observed behaviour can be recorded or documented by the researcher.

**Disadvantages**

The following are the disadvantages of participant observation:

- If a participant is at lower level in hierarchy of group, his participation may be less.
- Emotions of the observer may result in loss of objectivity.

(ii) Non-participant observation

In non-participant observation, the observer does not participate in the group process. He acknowledges the behaviour of the group without telling the respondents. It requires a lot of skills to record observations in an unnoticeable manner.

(iii) Direct observation

In direct observation, the observer and researcher personally observe all the happenings of a process or an event when the event is happening. In this method, the observer records all the relevant aspects of an event which are necessary for study. He is free to change the locations and focus of the observation. One major limitation of the method is that the observer may not be able to cover all relevant events when they are happening.

(iv) Indirect observation

Physical presence of an observer is not required and recording is done with the help of mechanical, photographic or electronic devices; for example, close circuit TV (CCTV) cameras are used in many showrooms to observe the behaviour of customers. It provides a permanent record for an analysis of different aspects of the event.

(v) Controlled observation

All observations are done under pre-specified conditions over extrinsic and intrinsic variables by adopting experimental design and systematically recording observations. Controlled observations are carried out either in the laboratory or the field.
(vi) Uncontrolled observation

There is no control over extrinsic and intrinsic variables. It is mainly used for descriptive research. Participant observation is a typical uncontrolled one.

Prerequisites of Observation

The following are the prerequisites of observation:

- The conditions of observation must provide accurate results. An observer should be in a position to observe the object clearly.
- The right number of respondents should be selected as the sample size for the observation to produce the desired results.
- Accurate and complete recording of an event.
- If it is possible, two separate observers and sets of instruments can be used in all or some observations. Then the result can be compared to measure accuracy and completeness.

Advantages and Limitations of Observation

The following are the advantages of observations:

- It ensures the study of behaviour in accordance with the occurrence of events. The observer does not ask anything from the representatives; he just watches the doing and saying of the sample.
- The data collected by observation defines the observed phenomenon as they occur in their natural settings.
- When an object is not able to define the meaning of its behaviour, observation is best method for analysis; for example, animals, birds and children.
- Observation covers the entire happenings of an event.
- Observation is less biased as compared to questioning.
- It is easier to conduct disguised observation studies as opposed to disguised questioning.
- The use of mechanical devices can generate accurate results.

The following are the limitations of observation:

- Past studies and events are of no use to observation. For these events and study, one has to go through narrations, people and documents.
- It is difficult to understand attitudes with the help of observation.
- Observations cannot be performed by the choice of the observer. He has to wait for an event to occur.
- It is difficult to predict when and where the event will occur. Thus, it may not be possible for an observer to reach in every event.
- Observation requires more time and money.
Data Collection

Uses of observation in business research

Observation is very useful in the following business research purposes:

- Buying behaviour of customer, lifestyles, customs, interpersonal relations, group dynamics, leadership styles, managerial style and actions
- Physical characteristics of inanimate things such as houses, factories, stores, and so on
- Movements in a production plant
- Flow of traffic, crowd and parking on road

10.3.2 Interview

Interviewing is a very effective method of data collection. It is a systematic and objective conversation between an investigator and respondent for collecting relevant data for a specific research study. Along with conversation, learning about the gestures, facial expressions and environmental conditions of a respondent are also very important. Generally, interview collects a wide range of data from factual demographic data to highly personal and intimate information relating to a person’s opinions, attitudes, values and beliefs, past experience and future intentions. The interview method is very important in the collection of data from the respondent who is less educated or illiterate. Personal interview is more feasible when the area covered for survey is compact. Probing is a very important part of an interview.

Types of Interview

The following are the various types of interviews:

(i) Structured or directive interview

In this type of interview, the investigator goes to the respondent with a detailed schedule. Some questions in same sequence are asked from all respondents.

(ii) Unstructured or non-directive interview

In this type of interview, the respondent is encouraged to give his honest opinion on a given topic without or with minimum help of others.

(iii) Focused interview

This is a semi-structured interview where the respondent shares the effect of the experience to the given conditions with the researcher or investigator. It is conducted with those respondents only who have prior experience of conditions given by the investigator. Analysis of the attitude and emotional feelings for the situations under study is main purpose behind conducting these interviews. A set of fix questions may not be required in this interview but a relevant topic is required which is known to the respondent.
(iv) Clinical interview

While a focused interview is concerned with effects of specific experience, clinical interviews are concerned with broad underlying feelings or motivations or the course of the individual's life experiences with reference to the research study. It encourages the interviewee to share his experience freely.

(v) Depth interview

To analyse or study the respondent's emotions, opinions, and so on, depth interviews are conducted. This kind of interview aims to collect intensive data about individuals, especially their opinions. It is a lengthy process to get unbiased data from the respondent. Interviewers should avoid advising or showing this agreement. Instead, the investigator has to motivate the respondent to answer the questions.

Features of Interviews

The following are some of the features of interviews:

- The interviewer and the respondent are the participants in any interview. They both are unknown to each other and so it is important for an interviewer to introduce himself first to the respondent.
- An interview has a beginning and a termination point in the relationship between the participants.
- Interview is not a mere casual conversational exchange. It has a specific purpose of collecting data which is relevant to the study.
- Interview is a mode of obtaining a verbal response to questions to put verbally. It is not always face to face.
- Success of interview depends on the interviewer and respondent, and how they perceive each other.
- It is not a standardized process.

Essentials for an Effective Interview

The following are the requirements for a successful interview:

- **Data availability**: The respondent should have complete knowledge of the information required for specific study.
- **Role perception**: The interviewer and the respondent should be aware of their roles in the interview process. The respondent should be clear about the topic or questions which have to be answered by him. Similarly, it is the responsibility of the interviewer to make the respondent comfortable by introducing himself first. The investigator should not affect the interview situation through subjective attitude and argumentation.
- **Respondent's motivation**: The respondent can hesitate to answer the questions. In this case, the approach and skills of the interviewer are very important as he has to motivate the respondent to answer or express himself.
Advantages and Disadvantages of Interviews

The following are the advantages of the interview method:

- In-depth and detailed information is collected.
- The interviewer tries to improve the responses and quality of data received. He can control the conditions in favour of the research study.
- Interviews help in gathering supplementary information which can be helpful to the study.
- Interviews use special scoring devices, visuals and materials to improve the quality of data or information collected.
- Interviews use observation and probing by the interviewer to see the accuracy and dependability of given data by the respondent.
- Interviews are flexible in nature.

The following are the disadvantages of interviews:

- Interviews consume more time and cost.
- The respondent’s responses can be affected by the way the interviewer asks the questions.
- The respondent may refuse to answer some personal questions which are relevant to the study.
- Recording and coding of data during the interview process may sometimes be difficult for the interviewer.
- The interviewer may not have good communication or interactive skills.

Interview Process

The following are the stages in an interview process:

(i) Preparation

The interviewer needs to make certain preparations to make an interview successful. The interviewer should keep all the copies of the schedule or guide ready. He should prepare the lists of respondents with their addresses, contact number and meeting time. He should prepare himself with all the approaches and skills required to conduct an interview. He should prepare himself to face all adverse situations during the interview. If the interviewer is not doing such planning, he can fail to collect the right information from the respondent.

(ii) Introduction

The interviewer is not known to the respondent. Therefore, the interviewer must introduce himself first to every respondent. In the introduction, the interviewer should tell about himself, his organization details and the purpose of his visit. If the interviewer knows someone who the respondent is familiar with, then he can use that person’s reference to make the respondent more comfortable. The following are some steps which help in motivating the respondent:
• The interviewer should introduce himself with a smiling face and always greet the respondent.
• He should identify and call the respondent by name.
• He must describe how the respondent is selected.
• He should explain the purpose and usefulness of the study.
• He should focus on the value of the respondent’s cooperation.

(iii) Developing rapport
It is important for an interviewer to develop a rapport with the respondent before starting the interview. By doing this, a cordial relationship is established between them. It helps the interviewer understand the inherent nature of the respondent which helps in building a rapport and the discussion can be started with some general topic or with the help of a person who is commonly known to both of them.

(iv) Carrying the interview forward
After establishing a rapport, the skills of the interviewer are required to carry the interview forward. The following are some guidelines which should be followed:
• Start the interview in an informal and natural manner.
• Ask all the questions in the same sequence as in the schedule.
• Do not take an answer for granted. It is not necessary that an interviewee will know all answers or will give all answers. The interviewer has to create interest for answering questions.
• The objective of the question should be known to the interviewer to ensure that the correct information is collected for research study.
• Explain the question if it has not been understood properly by the respondent.
• Listen to the respondent carefully with patience.
• Never argue with the respondent.
• Show your concern and interest in the information given by the respondent.
• Do not express your own opinion for answers of any question in the schedule.
• Continue to motivate the respondent.
• If the respondent is unable to frame the right answer, the interviewer should help him by providing alternate questions.
• Ensure that the conversation does not go off the track.
• If the respondent is unable to answer a particular question due to some reasons, drop the question at that moment. This question can be asked indirectly later on.
(v) Recording the interview

Responses should be recorded in the same sequence as they are given by the respondent. The response should be recorded at the same time as it is generated. It may be very difficult to remember all the responses later for recording them. Recording can be done in writing but there may be some problems if the writing skills of an interviewer are not good. Hence, use of electronic devices like tape recorders can help in this purpose. The interviewer should also record all his probes and other comments on the schedule, but they should be in brackets to ensure that they are set off from response.

(vi) Closing the interview

After the interview is over, the interviewer must thank the respondent for his cooperation. He must collect all the papers before leaving the respondent. If the respondent wants to know the result of the survey, the interviewer must ensure that the results are mailed to him when they are ready.

(vii) Editing

At the end, the interviewer must edit the schedule to check that all the questions have been asked and recorded. Also, abbreviations in recording should be replaced by full words.

Problems Faced in an Interview

The following are some of the main problems faced in an interview:

(i) Inadequate response

Kahn and Cannel laid down five principal symptoms of inadequate response. They are given as follows:

- Partial response in which the respondent gives a relevant but incomplete answer.
- Non-response in which the respondent remains silent or refuses to answer the questions.
- Irrelevant response in which the respondent’s answer is not relevant to the question asked.
- Inaccurate response in which the reply is biased.
- Verbalized response problem which arises because of the respondent’s failure to understand the question.

(ii) Interviewer’s biasness, refusal, incapability to understand questions, and so on.

An interviewer can affect the performance of an interview with his own responses and suggestions. Such biasing factors can never be overcome fully, but their effect can be reduced by training and development techniques.
(iii) Non-response
Some respondents out of the total respondents fail to respond to the schedule. The reasons for this non-response can be non-availability, refusal, incapability to understand questions, and so on.

(iv) Non-availability
Some respondents are not available at their places at the time of call. This could be because of odd timings or working hours.

(v) Refusal
Some respondents refuse to answer the questions. There can be many reasons for this, such as language, odd hours, sickness, no interest in such studies, and so on.

(vi) Inaccessibility
Some respondents can be inaccessible because of various reasons such as migration, touring job, and so on.

Methods and aims of controlling non-response
Researcher Kish suggests the following methods to reduce either the percentage of non-response or its effects:

1. Improved procedure for collecting data is the most obvious remedy for non-response. The improvements advocated are as follows:
   - Guarantee of anonymity
   - Motivation of the respondent to cooperate
   - Arousing the respondent’s interest by clever opening remarks and questions
   - Advance notice to the respondent
2. Call backs are the most effective way of reducing not-at-home responses in personal interviews, as are repeated mailings in no-returns in mail survey.
3. Substitution for non-response is often suggested as a remedy. Usually, this is a mistake because the substitutes resemble the responses rather than the non-responses. Nevertheless, beneficial substitution methods can sometimes be designed with references to important characteristics of population.

Attempts to reduce the percentage or effect of non-response is aimed at reducing the bias caused by vast differences between non-respondents and respondents. The response bias should not be confused with the reduction of sample size due to non-response. The latter effect can be easily overcome either by anticipating the size of non-response in designing the sample size or by compensating for it with a supplement. These adjustments increase the size of the response and the sampling precision, but they do not reduce the non-response percentage or bias.
Data Collection

Telephonic Interview

Telephonic interview is a non-personal method of data collection. It may be used as a major method or supplementary method of data collection. It is useful in the following conditions:

- When the population is composed of those people who are listed in telephone directories
- When less number of questions have to be answered by the respondents
- When the time available for the survey is less
- When the subject is of the interest to the respondent
- When the respondents are widely scattered

Advantages

The following are the advantages of telephonic interviews:

- Less time and low cost
- Good quality of response
- Less demanding on interviewer
- No field work is required
- Easy to contact those respondents who cannot be reached

Disadvantages

Telephonic interviews have the following limitations:

- Restricted to persons who are listed in telephone or other relevant directories
- Not feasible to conduct long interviews
- Limitation of information collected
- No answer to personal questions by respondents
- Respondent’s emotions, facial expressions and other environmental factors cannot be recorded
- Difficult to develop rapport

Group Interview

Group interview is the method of collecting primary data from a number of individuals with common interests. In group interviews, the interviewer performs the role of a discussion leader. Free discussion is encouraged on the same aspects of the subject under the study. Information is collected either through a self-administered questionnaire or through an interview. Samples for the group can be selected from schools, colleges, clubs and other associations.
Advantages

The following are the advantages of this technique:

- Respondent gets freedom to express his views.
- Flexible method.
- Use of visual aids.
- Less time consuming as group can be interviewed in the time required for one respondent’s interview.
- Respondents are more confident in groups.
- Eliminates the limitation of individual interviews.

Disadvantages

The following are the main disadvantages of group interviews:

- Difficulty in selecting the desired sample group.
- Dominance of one individual in a group.
- Respondents can be biased or they can try to please the interviewer or others.

10.3.3 Questionnaire

Primary data can be collected with the help of mails and surveys. The respondents receive the questionnaires from the researcher, and are asked to fill them completely and return them to the researcher. It can be performed only when the respondents are educated. The mail questionnaire should be simple and easy to understand, so that the respondents can answer all questions easily. In mail questionnaires, all the answers have to be given and recorded by the respondents and not by the researcher or investigator, as in the case of personal interview method. There is no face-to-face interaction between the investigator and respondent, and so the respondent is free to give answers of his own choice.

Importance of questionnaires

A questionnaire is a very effective method as well as research tool in any research study. It ensures the collection of a diversified and wide range of scientific data to complete the research objectives. The questionnaire provides all the inputs in the form of relevant data to all statistical methods used in a research study.

Types of Questionnaire

The following are the various categories of questionnaires:

(i) Structured or standard questionnaire

Structured or standard questionnaires contain predefined questions in order to collect the required data for research study. These questions are the same for all
the respondents. Questions are in the same language and in the same order for all the respondents.

(ii) Unstructured questionnaire

In unstructured questionnaires, the respondent has the freedom to answer all the questions in his own frame of reference and in his own terms.

Process of Data Collection

The researcher prepares the mailing list by collecting the addresses of all the respondents with the help of primary and secondary sources of data. A covering letter must accompany every questionnaire, indicating the purpose and importance of the research, and importance of cooperation of the respondent for the success of the research study.

Alternate modes of sending questionnaire

The following are the alternate modes of distributing questionnaire to respondents:

(i) Personal delivery

The researcher or investigator himself delivers the questionnaire to the respondents and requests them to fill it within a specific duration, i.e., one day or two days, as per the convenience of the researcher. After the given duration, they collect the questionnaire from the respondents. This adds the advantage of personal interview and mail survey. Alternatively, the questionnaire can be delivered personally to the respondents and the respondents return the questionnaire by mail to the researcher.

(ii) Attaching questionnaire to a product

When a firm is launching a new product or wants to collect the feedback on old products, the firm attaches a questionnaire with its product and requests the customers to fill the questionnaire. The company can give some discount or gift to the respondent of every return questionnaire.

(iii) Advertising the questionnaire

The questionnaire is advertised in magazines and newspapers with instructions to complete it. After filling the questionnaire from the magazine or newspaper, the respondents mail it to the advertiser.

(iv) News-stand inserts

In this method of sending questionnaires to the respondents, the questionnaire, along with covering letter and a self-addressed reply-paid envelope, is inserted into a random sample of news-stand copies of a newspaper or magazine.

Improving the Response in a Mail Survey

Generally, the response rate in mail surveys in countries like India is very low. The following techniques can be adopted to increase the rate of response:
• **Covering letter:** The covering letter should be in a language which generates the interest of the respondent. It should address the respondent by name.

• **Quality printing:** Sometimes the quality of the printed questionnaire is so bad that the respondent faces a lot of problems in reading it. This results in loss of interest, and so the quality of printing should be excellent and attractive.

• **Prior information:** Prior information can be given to the concerned respondent by telephone, e-mail, newsletters, and so on. Such steps bring more success than follow-ups.

• **Incentives:** Monetary and non-monetary incentives can be given to respondents who are filling the questionnaire. This generates a higher response.

• **Follow-ups:** The respondent can be approached with the help of an investigator to collect the questionnaire or to solve the problems faced by the respondent in filling the questionnaire.

• **Larger sample size:** We should always select a sample size which is larger than what is actually required. This will help the researcher in getting answers from the effective sample size.

### Advantages and Disadvantages of Questionnaires

The following are the advantages of questionnaires:

- Low cost
- Wide reach and extensive coverage
- Easy to contact the person who is busy
- Respondent’s convenience in completion of the questionnaire
- More impersonal; provides more anonymity
- No interviewer’s bias
- Accuracy

The following are the disadvantages of questionnaires:

- Low response by the respondent
- Low scope in many societies where literary level is low
- More time requirement

### Preparation of an Effective Questionnaire

While preparing a questionnaire, the researcher must focus on some key parameters to prepare it. These key parameters are as follows:

- Proper use of open and close probe
- Proper sequence of questions
- Use of simple language
Data Collection

NOTES

Collecting Data through Schedule

This method is very similar to the collection of data through questionnaires. The only difference is that in schedule, enumerators are appointed. These enumerators go to the respondents, ask the stated questions in the same sequence as the schedule and record the reply of the respondents. Schedules may be given to the respondents and the enumerators should help them solve the problems faced while answering the questions in the given schedule. Thus, enumerator selection is very important in data collection through schedules.

Distinction between schedule and questionnaire

Both questionnaire and schedule are popular methods of data collection. The following are the main differences between a questionnaire and a schedule:

- A questionnaire is generally sent to the respondents through mail, but in case of schedule, it is sent through enumerators.
- Questionnaires are relatively cheaper mediums of data collection as compared to schedules. In the case of questionnaires, the cost is incurred in preparing it and mailing it to the respondent, while in schedule, more money is required for hiring enumerators, training them and incurring their field expenses.
- The response rate in questionnaires is low as many people return it without filling. On the other hand, the response rate in schedules is high because they are filled by the enumerators.
- In collecting data through questionnaires, the identity of the respondent may not be known, but this is not the case when it comes to schedules.
- Data collection through questionnaires requires a lot of time, which is comparatively very less in case of schedules.
- Generally, there is no personal contact in case of questionnaires, but in schedules, personal contact is always there.
- The literacy level of the respondent is very important while filling questionnaires, but in schedules, the literacy level of the respondent is not a major concern as the responses have to be recorded by the enumerators.
- Wider distribution of questionnaires is possible but this is difficult with schedules.
- There is less accuracy and completeness of responses in questionnaires as compared to schedules.
- The success of questionnaires depends on the quality of questions but the success of a schedule depends on the enumerators.
• The physical appearance of questionnaire matters a lot, which is less important in case of schedules.
• Observation method cannot be used along with questionnaires but it can be used along with schedules.

Check Your Progress
1. Define primary and secondary data.
2. List the advantages of participant observation.
3. Mention the uses of observation in business research.
4. What is a focused interview?
5. Name the different stages of an interview.
6. State the advantages of questionnaires.

10.4 SAMPLING: INTRODUCTION

A part of the population is called sample. Selecting a part of the ‘universe’ with a view to draw conclusions about the ‘universe’ or ‘population’ for a study is known as sampling. A researcher uses sampling for saving time and costs as a selected sample is a replica of the population.

Sampling design: Census and sample survey

All items in any field of inquiry constitute a ‘universe’ or ‘population’. A complete enumeration of all the items in the ‘population’ is known as a census inquiry. It can be presumed that in such an inquiry, when all the items are covered, no element of chance is left and highest accuracy is obtained. In practice, this may not be true. Even the slightest element of bias in such an enquiry will get larger and larger as the number of observations increase. Moreover, there is no way of checking the element of bias or its extent, except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Therefore, when the field of inquiry is large, this method becomes difficult to adopt because of the resources involved. At times, this method is practically beyond the reach of ordinary researchers. Perhaps, the government is the only institution which can get the complete enumeration carried out. Even the government adopts this method in very rare cases, such as population census conducted once in a decade. Further, many a times, it is not possible to examine every item in the population and sometimes it is possible to obtain sufficiently accurate results by studying only a part of the total population. In such cases, there is no utility of census surveys.

However, it needs to be emphasized that when the universe is a small one, it is no use resorting to a simple survey. When field studies are undertaken in practical life, consideration of time and cost invariably leads to a selection of respondents, i.e., selection of only a few items. The respondents selected should
Data Collection

ACT as representatives of the total population in order to produce a miniature cross-section. The selected respondents constitute what is technically called a ‘sample’ and the selection process is called the ‘sampling technique’. The survey so conducted is known as a ‘sample survey’.

Algebraically, let the population size be N and if a part of size n (which is < N) of this population is selected according to some rule for studying some characteristics of the population, the group consisting of these n units is known as a ‘sample’. The researcher must prepare a sample design for his study, i.e., he must decide how a sample should be selected and of what size such a sample would be.

Implications of a sample design

A sample design is a definite plan for obtaining a sample from a given population. It refers to the technique or the procedure the researcher would adopt in selecting items for the sample. Sample design may as well lay down the number of items to be included in the sample, i.e. the size of the sample. Sample design is determined before data are collected. There are many sample designs from which a researcher can choose. Some designs are relatively more precise and easier to apply than others. The researcher must select/prepare a sample design which should be reliable and appropriate for his research study.

Advantages of sampling

The following are the advantages of sampling:

- **Size of population**: It is very difficult to study a large population for a research study; hence, a sample from the population is selected for the study and it represents all characteristics of population.

- **Funds requirement for the study**: When the funds availability is lesser than the anticipated cost of census survey, sampling is an effective method.

- **Facilities**: When facilities like technology and staff members are limited, sampling is preferable.

- **Time**: As the time required for the sampling procedure is less, a researcher prefers this method.

Sampling procedure

Sampling is a complicated process. A researcher has to identify all the factors which can affect the sample. The various criteria related to choice of sampling procedure are as follows:

- **Purpose of survey**: Defining the purpose of a survey helps the researcher in the selection of a particular method of sampling. A particular method of sampling choice depends on the geographical area of the survey, and the size and nature of the study.
• **Measurability**: The application of statistical inference theory requires computation of the sampling error from the sample itself. Probability samples only allow such computation. Hence, where the research objectives require statistical inference, the sample should be drawn by applying simple random sampling method or stratified random sampling method, depending whether the population is homogeneous or heterogeneous.

• **Degree of precision**: A desired level of precision of the result of the survey decides the method adopted for sampling.

• **Information about population**: Details of information available about the population to be studied help in deciding the method of sampling. If no data is available about population, it is difficult to apply probability random sampling. In this condition, the non-probability sampling method can be used for getting an idea of the population.

• **Nature of population**: Whether the population is homogeneous or heterogeneous decides the variables to be studied. Simple random sampling can be used for a homogeneous population. If the population is heterogeneous, stratified random sampling is a better option.

• **Geographical area of study and size of population**: Multi-stage, cluster sampling is used for the study of wide geographical area and large size of population.

• **Financial resources**: Availability of finance decides the need of sampling method.

• **Time limitation**: The time limit to complete a study decides the method of sampling.

**Characteristics of a good sample**

The following are the characteristics of a good sample:

• Representative
• Accuracy
• Precision
• Size

10.5 **SCALING AND CHECKLIST**

There are four types of measurement scales—nominal, ordinal, interval and ratio scales. We will discuss each one of them in detail. The choice of the measurement scale has implications for the statistical technique to be used for data analysis.

**Nominal scale**: This is the lowest level of measurement. Here, numbers are assigned for the purpose of identification of the objects. Any object which is assigned a higher number is in no way superior to the one which is assigned a lower number. In the nominal scale there is a strict one-to-one correspondence between the
numbers and the objects. Each number is assigned to only one object and each object has only one number assigned to it. It may be noted that the objects are divided into mutually exclusive and collectively exhaustive categories.

Examples of nominal scale:

- What is your religion?
  (a) Hinduism
  (b) Sikhism
  (c) Christianity
  (d) Islam
  (e) Any other, (please specify)

A Hindu may be assigned a number 1, a Sikh may be assigned a number 2, a Christian may be assigned a number 3 and so on. Any religion which is assigned a higher number is in no way superior to the one which is assigned a lower number. The assignment of numbers is only for the purpose of identification. We also note that all respondents have been divided into mutually exclusive and collectively exhaustive categories. For example:

- Are you married?
  (a) Yes
  (b) No
  If a person is married, he or she may be assigned a number 101 and an unmarried person may be assigned a number 102.

- In which of the following departments do you work?
  (a) Marketing
  (b) HR
  (c) Information Technology
  (d) Operations
  (e) Finance and Accounting
  (f) Any other, (please specify)

Here also, a person working for the marketing department may be assigned a number 1, the one working for HR may be assigned a number 2 and so on.

Nominal scale measurements are used for identifying food habits (vegetarian or non-vegetarian), gender (male/female), caste, respondents, brands, attributes, stores, the players of a hockey team and so on.

The assigned numbers cannot be added, subtracted, multiplied or divided. The only arithmetic operations that can be carried out are the count of each category. Therefore, a frequency distribution table can be prepared for the nominal scale variables and mode of the distribution can be worked out. One can also use chi-square test and compute contingency coefficient using nominal scale variables.
Ordinal scale: This is the next higher level of measurement than the nominal scale measurement. One of the limitations of the nominal scale measurements is that we cannot say whether the assigned number to an object is higher or lower than the one assigned to another option. The ordinal scale measurement takes care of this limitation. An ordinal scale measurement tells whether an object has more or less of characteristics than some other objects. However, it cannot answer how much more or how much less. An ordinal scale tells us the relative positions of the objects and not the difference between the magnitudes of the objects. Suppose Shashi scores the highest marks in marketing and is ranked no. 1; Mohan scores the second highest marks and is ranked no. 2; and Krishna scores third highest marks and is ranked no. 3. However, from this statement we cannot say whether the difference in the marks scored by Shashi and Mohan is the same as between Mohan and Krishna. The only statement which can be made under ordinal scale is that Shashi has scored higher than Mohan and Mohan has scored higher than Krishna. The difference between the ranks does not have any meaningful interpretation in the sense that it cannot tell the difference in absolute marks between the three candidates. Another example of the ordinal scale could be the CAT score given in percentile form. Suppose a candidate’s score is 95 percentile in the CAT exam. What it means is that 95 per cent of the candidates that appeared in the CAT examination have a score below this candidate, whereas only 5 per cent have scored more than him. The actual score is how much less or more cannot be known from this statement. Examples of the ordinal scale include quality ranking, rankings of the teams in a tournament, ranking of preference for colours, soft drinks, socio-economic class and occupational status, to mention a few. Some of the examples of ordinal scales are listed below:

- Rank the following attributes while choosing a restaurant for dinner. The most important attribute may be ranked one, the next important may be assigned a rank of 2 and so on.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food quality</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td></td>
</tr>
<tr>
<td>Menu variety</td>
<td></td>
</tr>
<tr>
<td>Ambience</td>
<td></td>
</tr>
<tr>
<td>Service</td>
<td></td>
</tr>
</tbody>
</table>

- Rank the following by placing a 1 beside the attribute you think is the most important, a 2 beside the attribute you think is the second most important and so on while purchasing a two-wheeler.
In the ordinal scale, the assigned ranks cannot be added, multiplied, subtracted or divided. One can compute median, percentiles and quartiles of the distribution. The other major statistical analysis which can be carried out is the rank order correlation coefficient, sign test. As the ordinal scale measurement is higher than the nominal scale measurement, all the statistical techniques which are applicable in the case of nominal scale measurement can also be used for the ordinal scale measurement. However, the reverse is not true. This is because ordinal scale data can be converted into nominal scale data but not the other way round.

**Interval scale:** The interval scale measurement is the next higher level of measurement. It takes care of the limitation of the ordinal scale measurement where the difference between the score on the ordinal scale does not have any meaningful interpretation. In the interval scale the difference of the score on the scale has meaningful interpretation. It is assumed that the respondent is able to answer the questions on a continuum scale. The mathematical form of the data on the interval scale may be written as

\[ Y = a + bX \]

where \( a \neq 0 \)

The interval scale data has an arbitrary origin (non-zero origin). The most common example of the interval scale data is the relationship between Celsius and Fahrenheit temperature. It is known that:

\[ C^\circ = \frac{5}{9} (F^\circ - 32) \]

\[ C^\circ = -160 + \frac{5}{9} F^\circ \]

This is of the form \( Y = a + bX \), where \( a = -\frac{160}{9} \) and \( b = \frac{5}{9} \) and hence it represents the interval scale measurement. In the interval scale, the difference in score has a meaningful interpretation while the ratio of the score on this scale does not have a meaningful interpretation. This can be seen from the following interval scale question:
• How likely are you to buy a new designer carpet in the next six months?

<table>
<thead>
<tr>
<th>Scale</th>
<th>Very unlikely</th>
<th>Unlikely</th>
<th>Neutral</th>
<th>Likely</th>
<th>Very likely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale A</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Scale B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scale C</td>
<td>-7</td>
<td>-4</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Suppose a respondent ticks the response category ‘likely’ and another respondent ticks the category ‘unlikely’. If we use any of the scales A, B or C, we note that the difference between the scores in each case is 2. Whereas, when the ratio of the scores is taken, it is 2, 3 and –1 for the scales A, B and C respectively. Therefore, the ratio of the scores on the scale does not have a meaningful interpretation. The following are some examples of interval scale data.

• How important is price to you while buying a car?

<table>
<thead>
<tr>
<th>Least important</th>
<th>Unimportant</th>
<th>Neutral</th>
<th>Important</th>
<th>Most important</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

• How do you rate the work environment of your organization?

<table>
<thead>
<tr>
<th>Very good</th>
<th>Good</th>
<th>Neither good nor bad</th>
<th>Bad</th>
<th>Very bad</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

• The counter-clerks at ICICI Bank, (Vasant Kunj Branch) are very friendly.

<table>
<thead>
<tr>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

• Rate the life of the battery of your inverter.

• Indicate the degree of satisfaction with the overall performance of Wagon R.

<table>
<thead>
<tr>
<th>Very unsatisfied</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Very satisfied</th>
</tr>
</thead>
</table>

• How expensive is the restaurant ‘Punjabi By Nature’?

<table>
<thead>
<tr>
<th>Extremely expensive</th>
<th>Definitely expensive</th>
<th>Somewhat expensive</th>
<th>Somewhat inexpensive</th>
<th>Definitely inexpensive</th>
<th>Extremely inexpensive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

• How likely are you to buy a new car within the next six months?

<table>
<thead>
<tr>
<th>Definitely will buy</th>
<th>Probably will buy</th>
<th>Neutral</th>
<th>Probably will not buy</th>
<th>Definitely will not buy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
The numbers on this scale can be added, subtracted, multiplied or divided. One can compute arithmetic mean, standard deviation, correlation coefficient and conduct a t-test, Z-test, regression analysis and factor analysis. As the interval scale data can be converted into the ordinal and the nominal scale data, therefore all the techniques applicable for the ordinal and the nominal scale data can also be used for interval scale data.

**Ratio scale:** This is the highest level of measurement and takes care of the limitations of the interval scale measurement, where the ratio of the measurements on the scale does not have a meaningful interpretation. The ratio scale measurement can be converted into interval, ordinal and nominal scale. But the other way round is not possible. The mathematical form of the ratio scale data is given by \( Y = bX \).

In this case, there is a natural zero (origin), whereas in the interval scale we had an arbitrary zero. Examples of the ratio scale data are weight, distance travelled, income and sales of a company, to mention a few.

**Checklist**

A checklist is a type of a data collection aid which is employed to decrease failure by compensating for potential limits of human memory and attention. It helps to make sure the stability and completeness in carrying out a task. A basic example of a check list would be a 'to do list'. A more advanced checklist would be a schedule, which lays out tasks to be done according to the time of day or other factors. A primary task in checklist is documentation of the task and auditing against the documentation.

Checklists often come in a format with small checkboxes down the left hand side of the page. Individuals are required to draw a small tick or checkmark after the item has been completed. This is the standard format of a checklist, however, other formats are also used sometimes.

<table>
<thead>
<tr>
<th>Check Your Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. What is sampling and why a researcher uses sampling?</td>
</tr>
<tr>
<td>8. Define census inquiry.</td>
</tr>
<tr>
<td>9. What are the advantages of sampling?</td>
</tr>
<tr>
<td>10. List a few characteristics of a good sample.</td>
</tr>
<tr>
<td>11. Name the four types of measurement scales.</td>
</tr>
<tr>
<td>12. What is the major difference between ratio and internal scale?</td>
</tr>
</tbody>
</table>

**10.6 LIBRARY RECORDS AND REPORTS**

We can understand Library definition as an accumulation of various sources of data and information, made open to individuals, schools, colleges, business
professionals etc., for reference or obtaining information. Libraries range in size from a few racks of books to a few million things.

Library Record is an independent record label founded in Melbourne, Australia in 1998. A library’s collection includes data and information in various forms such as books, periodicals, newspapers, manuscripts, films, maps, microform, CDs, videotapes, DVDs, e-books, audiobooks, databases and many other.

A Library runs through OPAC

It can be defined as Online Public Access Catalogue, a multisource finder for all the available resources kept in library. With the help of this system, we can come to know the lists for the number of data and information available along with their numbers and current status of their availability in Library or out for issuance to someone. Records are being created and pen down by some person’s or associations during their course of actions.

Most records are confirmations of authoritative, official, administration or business related exchanges. Conventional records positions incorporate letters, minutes, memoranda and reports, and so forth. The electronic record positions incorporate spreadsheets, databases, messages and copied. Here and there, what innovation has done is to re-characterize customary records designs. For instance, email is fundamentally the same as a letter. The reality it is transmitted electronically that separates email from letters.

ISO 15489: International Standard on Records Management defines a record as: ‘recorded information in any form, including data in a computer system, created, received and maintained as evidence and information by an organisation or person, in the transaction of business and kept as evidence of such activity’.

Records give data as portrayal, insights, or schedules did by an individual or a gathering working in an establishment or association. While all records pass on data, not all wellsprings of data are essentially records. For instance, a distributed book or a remotely given database (on the web or disconnected) won’t be a record, in spite of the fact that data chose from it and reused in another setting may itself turn into a record.

A library needs records for a variety of reasons such as follows:

- Records are expected to convey benefits in a reliable and impartial way as records help to bring and keep up consistency in the schedules.
- Records give benchmark measures to screen and assess the advancement in execution, effectiveness and accomplishments.
- Records are expected to archive moves and choices made in a library and to direct library business in an organized, effective and responsible way.
- Records are expected to give consistency, progression and profitability in the library and organization.
• Records also encourage the development of arrangements, projects, frameworks and long haul key designs.
• Records are expected to help and report strategy plan and administrative basic leadership and to advance educated basic leadership.
• Records fill in as source materials to build library history and keep up library memory.
• Records, as a list, fill in as an apparatus for proficient access to the library assets.

Methods to Store Records in Library
In a fully computerised library, records are computer generated by the relevant software modules. In non-automated libraries, the dominant forms of records are: ledgers, registers, files, cards and statistical sheets.

Category for records keeping may be divided into the following:
• Information presented in which format
• Content Shared
• Their life span

Further to elaborate more, please understand the below concept carefully

Records as per their life span is:
• **Permanent Records**: These records are pertaining to library origin, its plan of growth, approvals for sanctioning of staff, accession registers, user registration forms, audit reports, furniture inventory and fittings, building plan, electrical and sewage system structure etc.

• **Semi-permanent Records**: These records incorporate among others finances use enrol, concurrences with firms about terms and states of supply of books and access to databases and library dissemination records, in-house utilization records and online use records, and so forth. Budgetary records are held till the review is finished. Also records relating to the terms and states of supply of a firm need not be held once managing the firm is finished.

• **Temporary Records**: Records which are produced for a particular work are to be endless supply of the work. For instance, book determination cards/slips are typically devastated once the books have been inventoried. These are not decimated in the event that they are utilized as increase record. Sequential enlistment card is additionally a brief record. When it is full, it is supplansted by another card.

Content Records include following:
• **Records of Resources**: Variety of books and non-books like brochures, posters, slides, microfilms, maps, globes, computer software, audio books, audio visual and photos etc.
• **Financial Records**: This pertains to books of accounts such as budget register, allocation of funds register, bills and all those related to financial parameters.

• **Library Service Records**: These records include among others inter-library loan record, circulation record and reference service record, etc.

• **Admin Records**: There are numerous of staff working under library administration, so all they maintain is a service book which depicts their overall service starting from his application for that particular service.

Records by Presentation:

• **Descriptive**: All those reports and records linked with library including annual reports, stated policies, minutes of meeting are maintained under this section.

• **Statistical**: These kind of records are related to the number of books purchased, issued, and number of queries attended by supplying photocopies in a year. At the end of the session, the data collected are being analysed.

**Call number in the library**

This is one of the important feature of Library Record keeping which refer to an alphanumeric code that uniquely identifies and indicates the location of an item in a library’s collection. Some libraries use Library of congress classification (LC Number) in which the call number begins with letters, other libraries use Dewey Decimal method in which the call number begins with numerals.

**Card Catalogue**

The main motive of Card Catalogue is to help Library users to find a particular book through multiple options. It can be stated that card catalogue is brief index for the Library collection. By the help of this method one can easily locate the book he/she is looking for as all the cards are alphabetically arranged and being labelled outside the shelves/ trays.

There are three types of card catalogue.

• **Author Catalogue**: In this all the author last name are being written in alphabetical manner. Author who are with common surnames are being differentiated by their date of birth or death anniversary or by their middle initial. If a particular book is being written by various authors, so each author card is being supported by co-authors, so that just be knowing one author name they can still find the book.

• **Title Catalogue**: This includes main title and subtitle of the book along with their publishing date and total number of pages.

• **Subject Catalogue**: In this specifications are being done on the topics starting from the broadest one and become more specific by subdividing subjects.
Check Your Progress

13. What is a library record and what does it contain?

14. What is library card catalogue and name the three types of card in the card catalogue?

10.7 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Primary data is original, problem-or project-specific and collected for the specific objectives and needs spelt out by the researcher. The accuracy and relevance is reasonably high. The time and money required for this are quite high. However, a secondary data is that information which is not topical or research-specific and has been collected and compiled by some other researcher or investigative body. This type of data is collected and published in a structured format, and thus, is quicker to access and manage.

2. The following are the advantages of participant observation:
   - In-depth understanding of the respondent group.
   - The context which is meaningful to observed behaviour can be recorded or documented by the researcher.

3. Observation is very useful in the following business research purposes:
   (a) Buying behaviour of customer, lifestyles, customs, interpersonal relations, group dynamics, leadership styles, managerial style and actions
   (b) Physical characteristics of inanimate things such as houses, factories, stores, and so on
   (c) Movements in a production plant
   (d) Flow of traffic, crowd and parking on road

4. This is a semi-structured interview where the respondent shares the effect of the experience to the given conditions with the researcher or investigator. It is conducted with those respondents only who have prior experience of conditions given by the investigator. Analysis of the attitude and emotional feelings for the situations under study is main purpose behind conducting these interviews.

5. The different stages of an interview are as follows:
   - Preparation
   - Introduction
   - Developing rapport
   - Carrying the interview forward
• Recording the interview
• Closing the interview
• Editing

6. The following are the advantages of questionnaires:
   (a) Low cost
   (b) Wide reach and extensive coverage
   (c) Easy to contact the person who is busy
   (d) Respondent’s convenience in completion of questionnaire
   (e) More impersonal, provides more anonymity
   (f) No interviewer’s biasness
   (g) Accuracy

7. A part of population is called sample. Selecting a part of the ‘universe’ with a view to draw conclusions about the ‘universe’ or ‘population’ for a study is known as sampling. A researcher uses sampling for saving time and costs as a selected sample is a replica of the population.

8. A complete enumeration of all the items in the ‘population’ is known as a census inquiry. It can be presumed that in such an inquiry, when all the items are covered, no element of chance is left and highest accuracy is obtained.

9 The following are the advantages of the sampling:
   (a) **Size of population**: It is very difficult to study a large population for a research study; hence, a sample from the population is selected for the study and it represents all characteristics of population.
   (b) **Funds requirement for the study**: When the funds availability is lesser than the anticipated cost of census survey, sampling is an effective method.
   (c) **Facilities**: When facilities like technology and staff members are limited, sampling is preferable.
   (d) **Time**: As the time required for the sampling procedure is less, a researcher prefers this method.

10. The following are the characteristics of a good sample:
   • Representative
   • Accuracy
   • Precision
   • Size

11. The four types of measurement scales are – nominal, ordinal, interval and ratio scales.

12. The major difference between ratio and interval scale is that an interval scale has an arbitrary origin i.e., have an arbitrary zero (non-zero origin) whereas in the case of ratio scale, there is a natural zero (origin).
13. A Library Record is an independent record label founded in Melbourne, Australia in 1998. A library’s collection includes data and information in various forms such as books, periodicals, newspapers, manuscripts, films, maps, microform, CDs, videotapes, DVDs, e-books, audiobooks, databases and many other.

14. Card Catalogue is a set of cards in alphabetical order that helps library users to find a particular book through multiple options. It can be stated that card catalogue is brief index for the library collection. By the help of this method one can easily locate the book he/she is looking for as all the cards are alphabetically arranged and being labelled outside the shelves/trays. The three types of card catalogues are author catalogue, title catalogue and subject catalogue.

10.8 SUMMARY

- Observation can be defined as viewing or seeing. Observation means specific viewing with the purpose of gathering the data for a specific research study. Observation is a classical method of scientific study. It is very important in any research study as it is an effective method for data collection.
- Some of the major characteristics of observation method is it is selective, it is purposive and not casual, it should be accurate and standardized for its applications.
- With respect to an investigator’s role, observation may be classified into two types namely participant observation and non-participant observation.
- With respect to the method of observation, it can be classified into the following: direct observation and indirect observation.
- With reference to the control on the system to be observed, observation can be classified into the following: controlled observation and uncontrolled observation.
- In participant type of observation, the observer is an active participant of the group or process.
- In non-participant observation, the observer does not participate in the group process.
- In direct observation, the observer and researcher personally observe all the happenings of a process or an event when the event is happening. In this method, the observer records all the relevant aspects of an event which are necessary for study.
- In indirect observation, physical presence of an observer is not required and recording is done with the help of mechanical, photographic or electronic devices; for example, close circuit TV (CCTV) cameras are used in many showrooms to observe the behaviour of customers.
• All observations are done under pre-specified conditions over extrinsic and intrinsic variables by adopting experimental design and systematically recording observations.

• Uncontrolled observation is mainly used for descriptive research. Participant observation is a typical uncontrolled one.

• Interviewing is a very effective method of data collection. It is a systematic and objective conversation between an investigator and respondent for collecting relevant data for a specific research study.

• The interview method is very important in the collection of data from the respondent who is less educated or illiterate. Personal interview is more feasible when the area covered for survey is compact. Probing is a very important part of an interview.

• There are various types of interview: structured or directive interview, unstructured or non-directive interview, focused interview, clinical interview, depth interview, telephonic interview, group interview.

• Some of the main problems faced in an interview are inadequate response, non-response, interviewer’s biasness, refusal and inaccessibility.

• Primary data can be collected with the help of mails and surveys. The respondents receive the questionnaires from the researcher, and are asked to fill them completely and return them to the researcher.

• The mail questionnaire should be simple and easy to understand, so that the respondents can answer all questions easily. In mail questionnaires, all the answers have to be given and recorded by the respondents and not by the researcher or investigator, as in the case of personal interview method.

• A questionnaire is a very effective method as well as research tool in any research study. It ensures the collection of a diversified and wide range of scientific data to complete the research objectives.

• The two types of questionnaires are structured or standard questionnaire and unstructured questionnaire.

• Group interview is the method of collecting primary data from a number of individuals with common interests. In group interviews, the interviewer performs the role of a discussion leader. Free discussion is encouraged on the same aspects of the subject under the study.

• Both questionnaire and schedule are popular methods of data collection.

• A questionnaire is generally sent to the respondents through mail, but in case of schedule, it is sent through enumerators.

• A part of the population is called sample. Selecting a part of the ‘universe’ with a view to draw conclusions about the ‘universe’ or ‘population’ for a study is known as sampling. A researcher uses sampling for saving time and costs as a selected sample is a replica of the population.

• All items in any field of inquiry constitute a ‘universe’ or ‘population’. A complete enumeration of all the items in the ‘population’ is known as a census inquiry.
• A sample design is a definite plan for obtaining a sample from a given population.
• It refers to the technique or the procedure the researcher would adopt in selecting items for the sample.
• The few advantages of sampling are it represents all characteristics of population, sampling is an effective method when the fund availability is less, it is an effective method when facilities are limited, and a sampling procedure takes less time.
• Representation, accuracy, precision, and size are the few characteristics of a good sample.
• There are four types of measurement scales—nominal, ordinal, interval and ratio scales. We will discuss each one of them in detail. The choice of the measurement scale has implications for the statistical technique to be used for data analysis.
• Nominal scale is the lowest level of measurement. Here, numbers are assigned for the purpose of identification of the objects. Any object which is assigned a higher number is in no way superior to the one which is assigned a lower number.
• Ordinal scale is the next higher level of measurement than the nominal scale measurement. One of the limitations of the nominal scale measurements is that we cannot say whether the assigned number to an object is higher or lower than the one assigned to another option. The ordinal scale measurement takes care of this limitation.
• The interval scale measurement is the next higher level of measurement. It takes care of the limitation of the ordinal scale measurement where the difference between the score on the ordinal scale does not have any meaningful interpretation. In the interval scale the difference of the score on the scale has meaningful interpretation.
• Ratio scale is the highest level of measurement and takes care of the limitations of the interval scale measurement, where the ratio of the measurements on the scale does not have a meaningful interpretation. The ratio scale measurement can be converted into interval, ordinal and nominal scale. But the other way round is not possible.
• A checklist is a type of a data collection aid which is employed to decrease failure by compensating for potential limits of human memory and attention. It helps to make sure the stability and completeness in carrying out a task.
• Library Record is an independent record label founded in Melbourne, Australia in 1998. A library’s collection includes data and information in various forms such as books, periodicals, newspapers, manuscripts, films, maps, microform, CDs, videotapes, DVDs, e-books, audiobooks, databases and many other.
Online Public Access Catalogue, a multisource finder for all the available resources kept in library. With the help of this system, we can come to know the lists for the number of data and information available along with their numbers and current status of their availability in library or out for issuance to someone. Records are being created and pen down by some person’s or associations during their course of actions.

Records give data as portrayal, insights, or schedules did by an individual or a gathering working in an establishment or association. While all records pass on data, not all wellsprings of data are essentially records.

Permanent records are pertaining to library origin, its plan of growth, approvals for sanctioning of staff, accession registers, user registration forms, audit reports, furniture inventory and fittings, building plan, electrical and sewage system structure etc.

Semi-permanent records incorporate among others finances use enrol, concurrences with firms about terms and states of supply of books and access to databases and library dissemination records, in-house utilization records and online use records, and so forth.

Temporary records are those which are produced for a particular work are to be endless supply of the work. For instance, book determination cards/slips are typically devastated once the books have been inventoried.

Card Catalogue helps is to help Library users to find a particular book through multiple options. It can be stated that card catalogue is brief index for the Library collection.

The three types of card catalogues are author catalogue, title catalogue, and subject catalogue.

10.9 KEY WORDS

- **Descriptive surveys**: It refers to the surveys that generally collect information on what people think and do.
- **Analytic surveys**: It refers to the surveys that are generally used to either test hypotheses or to answer particular research questions.
- **Group interview**: It refers to a method of collecting primary data from a number of individuals with common interests.
- **Observation**: It means specific viewing with the purpose of gathering the data for a specific research study.
- **Sampling**: It refers to selecting a part of the ‘universe’ with a view to draw conclusions about the ‘universe’ or ‘population’ for a study.
- **Sample design**: It refers to a definite plan for obtaining a sample from a given population.
10.10 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions
1. Write a short note on the types of observation.
2. What are the features of an interview?
3. Briefly mention the advantages and disadvantages of interviews.
4. Briefly describe the various stages of an interview process.
5. What is the significance of questionnaires?
6. What are the alternate modes of distributing questionnaires to respondents?
7. Tabulate the advantages and disadvantages of questionnaires.
8. Differentiate between schedule and questionnaires.
9. List the merits of sampling.
10. What do you mean by probability sampling?
11. Write a brief note on library records.

Long Answer Questions
1. Discuss the different methods of data collection.
2. Describe the advantages and limitations of observation.
3. What are the different types of interviews? Explain
4. Identify the problems faced in an interview. Also, distinguish between telephonic interview and group interview.
5. Discuss various methods of sampling.
6. Elaborate the four types of measurement scales.

10.11 FURTHER READINGS


UNIT 11 QUESTIONNAIRE DESIGN

Structure
11.0 Introduction
11.1 Objectives
11.2 Questionnaire Construction, Design and Types
  11.2.1 Types of Questionnaire
11.3 Process of Questionnaire Designing
11.4 Secondary Data Sources and Precautions
  11.4.1 Uses of Secondary Data
  11.4.2 Advantages and Disadvantages of Secondary Data
  11.4.3 Types and Sources of Secondary Data
11.5 Answers to Check Your Progress Questions
11.6 Summary
11.7 Key Words
11.8 Self Assessment Questions and Exercises
11.9 Further Readings

11.0 INTRODUCTION

A discussion on data collection would be incomplete if one did not talk about the questionnaire method. This is the most cost effective and widely used method, apart from being extremely user friendly. The questionnaire method is flexible enough to reveal data that is in the respondents own words and language. It can be made extremely scientific by framing questions which enable a very advanced level of quantitative measurement and analysis. The pattern of questioning is always designed in accordance with the respondent’s comfort and ease of answering. Today, with the wide use of technology it is very easy to use the questionnaire method even without being present physically in front of the respondent.

In this unit, you will study about the significance of questionnaire method in research process. The construction and design and the types of questionnaire is also being discussed in the unit. In addition to this, the unit will also discuss the advantages and disadvantages of the questionnaire method and will also explain the importance of secondary data sources along with its advantages, uses and types.

11.1 OBJECTIVES

After going through this unit, you will be able to:

- Recognize the relevance of questionnaire method in research
- Describe the construction and design process of a questionnaire
- Discuss the different types of questionnaire
- Determine the types of questions asked in questionnaire
11.2 QUESTIONNAIRE CONSTRUCTION, DESIGN AND TYPES

The questionnaire is a research technique that consists of a series of questions asked to respondents, in order to obtain statistically useful information about a given topic. It is one of the most cost-effective methods of collecting primary data, which can be used with considerable ease by most individual and business researchers. It has the advantage of flexibility of approach and can be successfully adapted for most research studies. The instrument has been defined differently by various researchers. Some take the traditional view of a written document requiring the subject to record his/her own responses (Kervin, 1999). Others have taken a broader perspective to include structured interview also as a questionnaire (Bell, 1999). It is essentially a data-collection instrument that has a predesigned set of questions, following a particular structure (De Vaus, 2002). Since it includes a standard set of questions, it can be successfully used to collect information from a large sample in a reasonably short period of time.

However, the use of questionnaire is not always the best method in all research studies. For example, at the exploratory stage, rather than questionnaire, it is advisable to use a more unstructured interview. Secondly, when the number of respondents is small and one has to collect more subjective data, then a questionnaire is not advisable.

Criteria for designing a questionnaire

There are certain criteria that must be kept in mind while designing the questionnaire. The first and foremost requirement is that the spelt-out research objectives must be converted into clear questions which will extract answers from the respondent. This is not as easy as it sounds, for example, if one wants to know how many times your teacher praised you in a week? It is very difficult to give an exact number. The second requirement is, it should be designed to engage the respondent and encourage a meaningful response. For example, a questionnaire measuring stress cannot have a voluminous set of questions which fatigue the subject. The questions, thus, should encourage response and be easy to understand. Lastly, the questions should be self-explanatory and not confusing as then the person will answer the way he understood the question and not in terms of what was asked. This will be discussed in detail later, when we discuss the wording of the questions.

11.2.1 Types of Questionnaire

There are many different types of questionnaire available to the researcher. The categorization can be done on the basis of a variety of parameters. The two criteria

- Examine the advantages and disadvantages of the questionnaire method
- Discuss the types, uses, and advantages of secondary data sources
that are most frequently used for designing purposes are the degree of structure and the degree of concealment. Structure refers to the degree to which the response category has been defined. Concealment refers to the degree to which the purpose of the study is explained to the respondent.

Instead of considering them as individual types, most research studies use a mixed format. Thus, they will be discussed here as a two-by-two matrix (Figure 11.1).

![Fig. 11.1 Types of Questionnaire](image)

Let us discuss the types of questionnaires. Questionnaires can be categorized on the basis of their structure or method of administration.

Based on the structure, questionnaires can be divided into the following categories:

**Formalized and unconcealed questionnaire:** This is the one that is the most frequently used by all management researchers. For example, if a new brokerage firm wants to understand the investment behaviour of people, they would structure the questions and answers as follows:

1. Do you carry out any investment(s)?
   Yes ________ No ________
   If yes, continue, else terminate.
2. Out of the following options, where do you invest? (tick all that apply).
   Precious metals ________, real estate ________, stocks ________,
   government instruments ________, mutual funds ________,
   any other ________.

   This kind of structured questionnaire is easy to administer, and has both the questions as self-explanatory and the answer categories clearly defined.

**Formalized and concealed questionnaire:** These questionnaires have a formal method of questioning; however the purpose is not clear to the respondent. The research studies which are trying to find out the latent causes of behaviour and cannot rely on direct questions use these. For example young people cannot be asked direct questions on whether they are likely to indulge in corruption at work. Thus, the respondent has to be given a set of questions that can give an indication of what are his basic values, opinions and beliefs, as these would influence how he would react to issues.
Self-Instructional Material

Questionnaire Design

NOTES

Non-formalized and unconcealed: Some researchers argue that rather than giving the respondents pre-designed response categories, it is better to give them unstructured questions where they have the freedom of expressing themselves the way they want. Some examples of these kinds of questions are given below:

1. Why do you think Maggi noodles are liked by young children?

2. How do you generally decide on where you are going to invest your money?

3. Give THREE reasons why you believe that the show Satyamev Jayate has affected the common Indian person?

The data obtained here is rich in content, but quantification cannot go beyond frequency and percentages to represent the findings.

Non-formalized and concealed: If the objective of the research study is to uncover socially unacceptable desires and subconscious and unconscious motivations, the investigator makes use of questions of low structure and disguised purpose. However, these require interpretation that is highly skilled. Cost, time and effort are also much higher than others.

Another useful way of categorizing questionnaires is on the method of administration. Thus, the questionnaire that has been prepared would necessitate a face-to-face interaction. In this case, the interviewer reads out each question and makes a note of the respondent’s answers. This administration is called a schedule. It might have a mix of the questionnaire type as described in the section above and might have some structured and some unstructured questions. The other kind is the self-administered questionnaire, where the respondent reads all the instructions and questions on his own and records his own statements or responses. Thus, all the questions and instructions need to be explicit and self-explanatory.

The selection of one over the other depends on certain study prerequisites.

Population characteristics: In case the population is illiterate or unable to write the responses, then one must as a rule use the schedule, as the questionnaire cannot be effectively answered by the subject himself.

Population spread: In case the sample to be studied is large and widely spread, then one needs to use the questionnaire. When the resources available for the study are limited, then schedules become expensive to use and the self-administered questionnaire is better.

Study area: In case one is studying a sensitive topic like harassment at work- a self administered questionnaire is suggested. However, in case the study topic needs additional probing then in that case a schedule is better.

There is another categorization that is based upon the mode of administration; this would be discussed in later sections of the unit.
Check Your Progress
1. State the first and foremost requirement for designing a questionnaire.
2. What do you mean by the degree of structure and the degree of concealment?
3. Name the categories of questionnaires on the basis of method of administration.

11.3 PROCESS OF QUESTIONNAIRE DESIGNING

Even though the questionnaire method is most used by researchers, designing a well-structured instrument needs considerable skill. Presented below is a standardized process that a researcher can follow.

Figure 11.2 summarizes the steps involved in questionnaire design.

Fig. 11.2 Questionnaire Design Process
1. Convert the research objectives into information areas

This is the first step of the design process. By this time the researcher is clear about the research questions; research objectives; variables to be studied; research information required and the characteristics of the population being studied. Once these tasks are done, one can prepare a tabled framework so that the questions which need to be developed become clear. This step-wise process is explained with an example in Table 11.1.

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Research Objectives</th>
<th>Variables to be Studied</th>
<th>Information (Primary Required)</th>
<th>Population to be Studied</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the nature of plastic bag usage amongst people in the NCR (National Capital Region)?</td>
<td>To identify the different uses of plastic bags.</td>
<td>Usage behaviour, Demographic details</td>
<td>Uses of plastic bags, Disposal of plastic bags</td>
<td>Consumers, Retailers</td>
</tr>
</tbody>
</table>

2. Method of administration

Once the researcher has identified his information area; he needs to specify how the information should be collected. The researcher usually has available to him a variety of methods for administering the study. The main methods are personal schedule (discussed earlier in the unit), self-administered questionnaire through mail, fax, e-mail and web-based questionnaire. There are different preconditions for using one method over the other (Table 11.2).

<table>
<thead>
<tr>
<th>Administrative control</th>
<th>Administrative control</th>
<th>Sensitive issues</th>
<th>Sensitive issues</th>
<th>Large sample</th>
<th>Large sample</th>
<th>Cost/time taken</th>
<th>Cost/time taken</th>
<th>Question structure</th>
<th>Question structure</th>
<th>Sampling control</th>
<th>Sampling control</th>
<th>Response rate</th>
<th>Response rate</th>
<th>Interviewer bias</th>
<th>Interviewer bias</th>
</tr>
</thead>
</table>
3. Content of the questionnaire

The next step is to determine the matter to be included as questions in the measure. The researcher needs to do an objective quality check in order to see what research objective/information need the question would be covering before using any of the framed questions.

*How essential is it to ask the question?* You must remember that the time of the respondent is precious and it should not be wasted. Unless a question is adding to the data needed for getting an answer to the research problem, it should not be included. For example, if one is studying the usage of plastic bags, then demographic questions on age group, occupation, education and gender might make sense but questions related to marital status, family size and the state to which the respondent belongs are not required as they have no direct relation with the usage or attitude towards plastic bags.

Sometimes, especially in self-administered questionnaires, one may ask some neutral questions at the beginning of the questionnaire to establish an involvement and rapport. For example, for a biofertilizer usage study, the following question was asked:

- Farming for you is a:
  - noble profession
  - ancestral profession
  - profession like any other
  - profession that is not money making
  - any other

*Do we need to ask several questions instead of a single one?* After deciding on the significance of the question, one needs to ascertain whether a single question will serve the purpose or should more than one question be asked. For example, in a TV serial study, one may give ten popular serials to be ranked as 1 to 10 in order of preference. Then the second question after the ranking question is:

‘Why do you like the serial _________ (the one you ranked No. 1/prefer watching most)?’

Incorrect

Here, one lady might say, ‘Everyone in my family watches it’. While another might say, ‘It deals with the problems of living in a typical Indian joint family system’ and yet another might say, ‘My friend recommended it to me’.

Thus, we need to ask her:

‘What do you like about _________?’

‘Who all in your household watch the serial?’

and

‘How did you first hear about the serial?’

Correct
4. Motivating the respondent to answer

The questionnaire should be designed in a manner that it involves the respondent and motivates him/her to give information. There are different situations which might lead to this. Each of these is examined separately here:

**Does the person have the required information?** It has been found that the person has had no experience with the issue being studied. Look at the following question:

> How do you evaluate the negotiation skills module, viz., the communication and presentation skill module?  
>(Incorrect)

In this case it might be that the person has not undergone one or even both the modules, so how can he compare? Thus, certain qualifying or filter questions must be asked. Filter questions enable the researcher to filter out the respondents who are not adequately informed. Thus, the correct question would have been:

Have you been through the following training modules?
- Negotiation skills module  
  Yes/no
- Communication and presentation skills  
  Yes/no

In case the answer to both is yes, please answer the following question, or else move to the next question.

> How do you evaluate the negotiation skills module, viz., the communication and presentation skill module?  
>(Correct)

**Does the person remember?** Many a times, the question addressed might be putting too much stress on an individual’s memory. For example, consider the following questions:

> How much did you spend on eating out last month?  
>(Incorrect)

Such questions are beyond any normal individual’s memory bank. Thus, the questions listed above could have been rephrased as follows:

When you go out to eat, on an average your bill amount is:
- Less than ₹100
- ₹101–250
- ₹251–500
- More than ₹500

> How often do you eat out in a week?  
> 1–2 times
> 3–4 times
> 5–6 times
> Every day  
>(Correct)

**Can the respondent articulate?** Sometimes the respondent might not know how to put the answer in clear words. For example, if you ask a respondent to:

- Describe a river rafting experience.
Most respondents would not know what phrases to use to give an answer. Thus, in the above case, one can provide answer categories to the person as follows:

<table>
<thead>
<tr>
<th></th>
<th>Unexciting</th>
<th>Exciting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Describe the river rafting experience. (Correct)

Sensitive information: There might be instances when the question being asked might be embarrassing to the respondents and thus they would not be comfortable in disclosing the data required.

For example, questions such as the following will not get any answers.

- Have you ever used fake receipts to claim your medical allowance? (Incorrect)
- Have you ever spit tobacco on the road (to tobacco consumers)? (Incorrect)

However, in case the socially undesirable habit is in the context of a third person, the chances of getting some correct responses are possible. Thus the questions should be rephrased as follows:

- Do you associate with people who use fake receipts to claim their medical allowance? (Correct)
- Do you think tobacco consumers spit tobacco on the road? (Correct)

5. Determining the type of questions

Available to the researcher are different kinds of question-response options (Figure 11.3)

![Fig. 11.3 Types of Question-Response Options](image-url)
Open-ended questions

In open-ended questions, the openness refers to the option of answering in one’s own words. They are also referred to as unstructured questions or free-response or free-answer questions. Some illustrations of this type are listed below:

- What is your age?
- Which is your favourite TV serial?
- I like Nescafe because ________________________
- My career goal is to ________________________

Closed-ended questions

In closed-ended questions, both the question and response formats are structured and defined. There are three kinds of formats as we observed earlier—dichotomous questions, multiple-choice questions and those that have a scaled response.

i. Dichotomous questions: These are restrictive alternatives and provide the respondents only with two answers. These could be ‘yes’ or ‘no’, like or dislike, similar or different, married or unmarried, etc.

- Are you diabetic? Yes/No
- Have you read the new book by Dan Brown? Yes/No
- What kind of petrol do you use in your car? Normal/Premium

Dichotomous questions are the easiest type of questions to code and analyse. They are based on the nominal level of measurement and are categorical or binary in nature.

ii. Multiple-choice questions: Unlike dichotomous questions, the person is given a number of response alternatives here. He might be asked to choose the one that is most applicable. For example, this question was given to a retailer who is currently not selling organic food products:

Will you consider selling organic food products in your store?

- Definitely not in the next one year
- Probably not in the next one year
- Undecided
- Probably in the next one year
- Definitely in the next one year

Sometimes, multiple-choice questions do not have verbal but rather numerical options for the respondent to choose from, for example:

How much do you spend on grocery products (average in one month)?

- Less than ₹2500/-
- Between ₹2500–5000/-
- More than ₹5000/-
Most multiple-choice questions are based upon ordinal or interval level of measurement. There could also be instances when multiple options are given to the respondent and he can select all those that apply in the case. These kinds of multiple-choice questions are called checklists. For example, in the organic food study, the retailer who does not stock organic products was given multiple reasons as follows:

You do not currently sell organic food products because (Could be ≥ 1)
- You do not know about organic food products.
- You are not interested.
- Organic products do not have attractive packaging.
- Organic food products are not supplied regularly.
- Any other __________________________

iii. Scales: Scales refer to the attitudinal scales that were discussed in detail in Unit 10. Since these questions have been discussed in detail in the earlier unit, we will only illustrate this with an example. The following is a question which has two sub-questions designed on the Likert scale. These require simple agreement and disagreement on the part of the respondent. This scale is based on the interval level of measurement.

Given below are statements related to your organization. Please indicate your agreement/disagreement with each:

<table>
<thead>
<tr>
<th>(1-Strongly Disagree</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5 Strongly Agree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The people in my company know their roles very clearly.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. I want to complete my current task by hook or by crook.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11. Criteria for question designing

Step six of the questionnaire involves translating the questions identified into meaningful questions. There are certain designing criteria that a researcher should keep in mind when writing the research questions.

Clearly specify the issue: By reading the question, the person should be able to clearly understand the information need.

Which newspaper do you read? (Incorrect)

This might seem to be a well-defined and structured question. However, the ‘you’ could be the person filling the questionnaire or the family. He could be reading different newspapers. He might be reading different papers at home and may be the college library. A better way to word the question would be:

Which newspaper or newspapers did you personally read at home during the last month? In case of more than one newspaper, please list all that you read. (Correct)
Questionnaire Design

NOTES

Use **simple terminology**: The researcher must take care to ask questions in a language that is understood by the population under study. Technical words or difficult words that are not used in everyday communication must be avoided.

- Do you think thermal wear provides immunity? (Incorrect)
- Do you think that thermal wear provides you protection from the cold? (Correct)

**Avoid ambiguity in questioning**: The words used in the questionnaire should mean the same thing to all those answering the questionnaire. A lot of words are subjective and relative in meaning. Consider the following question:

How often do you visit Pizza Hut?
- Never
- Occasionally
- Sometimes
- Often
- Regularly (Incorrect)

These are ambiguous measures, as *occasionally* in the above question, might be three to four times in a week for one person, while for another it could be three times in a month. A much better wording for this question would be the following:

In a typical month, how often do you visit Pizza Hut?
- Less than once
- 1 or 2 times
- 3 or 4 times
- More than 4 times (Correct)

**Avoid leading questions**: Any question that provides a clue to the respondents in terms of the direction in which one wants them to answer is called a leading or biasing question. For example, ‘Do you think that working mothers should buy ready-to-eat food when that might contain some chemical preservatives?’

- Yes
- No
- Don’t know (Incorrect)

The question would mostly generate a negative answer, as no working mother would like to buy something that is convenient but might be harmful. Thus, it is advisable to construct a neutral question as follows:

Do you think that working mothers should buy ready-to-eat food?
- Yes
- No
- Don’t know (Correct)
Avoid loaded questions: Questions that address sensitive issues are termed as loaded questions and the response to these questions might not always be honest, as the person might not wish to admit the answer. For example, questions such as follows will rarely get an affirmative answer:

Will you take dowry when you get married?  

(SIncorrect)  
Sensitive questions like this can be rephrased in a variety of ways. For example, the question could be constructed in the context of a third person as follows:

Do you think most Indian men would take dowry when they get married?  

(CCorrect)  

Avoid double-barrelled questions: Questions that have two separate options separated by an ‘or’ or ‘and’ like the following:

Do you think Nokia and Samsung have a wide variety of touch phones?  

Yes/No (Incorrect)  

The problem is that the respondent might believe that Nokia has better phones or Samsung has better phones or both. These questions are referred to as double-barrelled and the researcher should always split them into two separate questions. For example,

A wide variety of touch phones is available for:

Nokia  
Samsung  
Both  

(CCorrect)  

7. Determine the questionnaire structure

The questions now have to be put together in a proper sequence.

Instructions: The questionnaires always, even the schedules, begin with standardized instructions. These begin by greeting the respondent and then introducing the researcher and then the purpose of questionnaire administration. For example, in the study on organic food products, the following instructions were given at the beginning of the questionnaire:

‘Hi. We __________ are carrying out a market research on the purchase behaviour of grocery products/organic food. We are conducting a survey of consumers, retailers and experts in the NCR for the same. As you are involved in the purchase and/or consumption of food products, we seek your cooperation for providing the following relevant information for our research. Thank you very much.’

Opening questions: After instructions come the opening questions, which lead the reader into the study topic. For example, a questionnaire on understanding the consumer’s buying behavior in malls can ask an opening question that is generic in nature, such as:

What is your opinion about shopping at a mall?
Study questions: After the opening questions, the bulk of the instrument needs to be devoted to the main questions that are related to the specific information needs of the study. Here also, the general rule is that the simpler questions, which do not require a lot of thinking or response time should be asked first as they build the tempo for answering the more difficult/sensitive questions later on. This method of going in a sequential manner from the general to the specific is called the funnel approach.

Classification information: This is the information that is related to the basic socio-economic and demographic traits of the person. These might include name (kept optional in some cases), address, e-mail address and telephone number.

Acknowledgement: The questionnaire ends by acknowledging the inputs of the respondent and thanking him for his cooperation and valuable contribution.

8. Physical characteristics of the questionnaire

The researcher must pay special attention to the look of the questionnaire. The first thing is the quality of the paper on which the questionnaire is printed which should be of good quality. The font style and spacing used in the entire document should be uniform. One must ensure that every question and its response options are printed on the same page. Surveys for different groups could be on different coloured paper. For example, if Delhi is being studied as five zones, then the questionnaire used in each zone could be printed on a differently coloured paper. Each question and section must be numbered properly. In case there is any response instruction for an individual question, it must be before the question. In case the questionnaire is going to be administered by the investigator and if there are any probing question then they should be clearly written as instructions for the investigator.

9. Pilot testing of the questionnaire

Pilot testing refers to testing and administering the designed instrument on a small group of people from the population under study. This is to essentially cover any errors that might have still remained even after the earlier eight steps. For example the question wording may not be clear, the sequence of questions may not be correct or the question is not needed as it does not solve any purpose. Thus these aspects need to be corrected. Every aspect of the questionnaire has to be tested and one must record all the experiences of the conduction, including the time taken to administer it. Sometimes, the researcher might also get the questionnaire whetted by academic or industry experts for their inputs. As far as possible, the pilot should be a small scale replica of the actual survey that would be subsequently conducted.

10. Administering the questionnaire

Once all the nine steps have been completed, the final instrument is ready for conduction and the questionnaire needs to be administered according to the sampling plan.
Advantages and Disadvantages of the Questionnaire Method

The questionnaire has many advantages over the other data collection methods discussed earlier.

- Probably the greatest benefit of the method is its adaptability. There is, actually speaking, no domain or branch for which a questionnaire cannot be designed. It can be shaped in a manner that can be easily understood by the population under study. The language, the content and the manner of questioning can be modified suitably. The instrument is particularly suitable for studies that are trying to establish the reasons for certain occurrences or behaviour.

- The second advantage is that it assures anonymity if it is self-administered by the respondent, as there is no pressure or embarrassment in revealing sensitive data. A lot of questionnaires do not even require the person to fill in his/her name. Administering the questionnaire is much faster and less expensive as compared to other primary and a few secondary sources as well. There is considerable ease of quantitative coding and analysis of the obtained information as most response categories are closed-ended and based on the measurement levels as discussed in Unit 5. The chance of researcher bias is very little here.

- Lastly, there is no pressure of immediate response, thus the subject can fill in the questionnaire whenever he or she wants.

- The questionnaire is the most economical method as it can be administered simultaneously to a number of respondents. Thus a large amount of data can be collected within a short time through a questionnaire.

However, the method does not come without any disadvantages.

- The major disadvantage is that the inexpensive standardized instrument has limited applicability, that is, it can be used only with those who can read and write.

- The questionnaire is an impersonal method and sometimes for a sensitive issue it may not reveal the actual reasons or answers to the questions that you asked. The return ratio, i.e., the number of people who return the duly filled in questionnaires are sometimes not even 50 per cent of the number of forms distributed.

- Skewed sample response could be another problem. This can occur in two cases; one, if the investigator distributes the same to his friends and acquaintances and second, because of the self-selection of the subjects. This means that the ones who fill in the questionnaire and return it might not be the representatives of the population at large. In case the person is not clear about a question, clarification with the researcher might not be possible.
Check Your Progress

4. Mention some other names of open-ended questions.
5. Name the method in which sampling control is the highest.
6. What are loaded and double-barrelled questions?
7. State the greatest benefit of the questionnaire method.

11.4 SECONDARY DATA SOURCES AND PRECAUTIONS

We have already discussed what secondary data is. Let us see what are its uses, types and sources.

11.4.1 Uses of Secondary Data

Secondary data can be used for multiple purposes as follows:

- **Problem identification and formulation stage**: Existing information on the topic under study is useful to help develop the research question.
- **Hypotheses designing**: Previous research studies done in the area could help in hypothesizing about expected results.
- **Sampling considerations**: There might be respondent related databases available to seek respondent statistics and relevant contact details. These would help during sampling for the study.
- **Primary base**: The secondary information collected can be used to design the primary data collection instruments, in order to phrase and design the right questions.
- **Validation board**: Earlier records and studies can also be used to support or validate the information collected through primary sources.

Before we examine the wide range of the secondary sources available to the business researcher, it is essential that one is aware of the advantages and disadvantages of using secondary sources.

11.4.2 Advantages and Disadvantages of Secondary Data

There are multiple advantages of using secondary data.

- **Resource advantage**: Any research that is making use of secondary information will be able to save immensely in terms of both cost and time.
- **Accessibility of data**: The other major advantage of secondary sources is that it is very easy to access this data.
- **Accuracy and stability of data**: Data from recognized sources has the additional advantage of accuracy and reliability.
• **Assessment of data:** It can be used to compare and support the primary research findings of the present study. However, there is need for caution as well because in using secondary data, there might be some disadvantages like:
  
  • **Applicability of data:** The information might not be directly suitable for our study. Also since it is past data it might not be applicable today.
  
  • **Accuracy of data:** All data that is available might not be reliable and accurate.

11.4.3 Types and Sources of Secondary Data

Secondary data can be divided into internal and external sources. Internal, as the name implies, is organization- or environment-specific source and includes the historical output and records available with the organization which might be the backdrop of the study. The data that is independent of the organization and covers the larger industry-scape would be available in the form of published material, computerized databases or data compiled by syndicated services. Discussed below are three major sources of data – internal, external, computer-stored data and syndicated databases.

1. **Internal sources of data**

Compilation of various kinds of information and data is mandatory for any organization that exists. Some sources of internal information are presented in Figure 11.4.

![Fig. 11.4 Internal Sources of Data](image)

- **Company records:** This includes all the data about the inception, the owners, and the mission and vision statements, infrastructure and other details, including both the process and manufacturing (if any) and sales, as well as a historical timeline of the events.
- **Employee records:** All details regarding the employees (regular and part-time) would be part of employee records.
- **Sales data:** This data can take on different forms:
  1. **Cash register receipt**
  2. **Salespersons’ call records:** This is a document to be prepared and updated every day by each individual salesperson.
(iii) **Sales invoices:** Customer who has placed an order with the company, his complete details including the size of the order, location, price by unit, terms of sale and shipment details (if any).

- **Financial records and sales reports**
  Besides this, there are other published sources like warranty records, CRM data and customer grievance data which are extremely critical in evaluating the health of a product or an organization.

2. **External data sources**
   As stated earlier, information that is collected and compiled by an outside source that is external to the organization is referred to as external source of data. External sources of data include the following:

**Published data:** There could be two kinds of published data—one that is from the official and government sources and the other kind of data is that which has been prepared by individuals or private agencies or organizations.

**Government sources:** The Indian government publishes a lot of documents that are readily available and are extremely useful for the purpose of providing background data. A brief snapshot of some government data is given in Table 11.3.

<table>
<thead>
<tr>
<th>Sub-type</th>
<th>Sources</th>
<th>Data</th>
<th>Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Census data conducted every ten years throughout the country</td>
<td>Registrar General of India conducting census survey <a href="http://censusindia.gov.in/">http://censusindia.gov.in/</a></td>
<td>Size of the population and its distribution by age, sex, occupation and income levels. 2010 census is taking many more variables to get a better picture of the population</td>
<td>Population information is significant as the forecasts of purchase, estimates of growth and development, as well as policy decisions can be made on this base</td>
</tr>
<tr>
<td>2. Statistical Abstract India – annually</td>
<td>CSO (Central Statistical Organization) for the past 5 years <a href="http://www.mospi.gov.in/cso_test1.htm">http://www.mospi.gov.in/cso_test1.htm</a></td>
<td>Education, health, residential information at the state level is part of this document</td>
<td>Making demand, estimations and a state level assessment of government support and policy changes can be made</td>
</tr>
</tbody>
</table>
NOTES

4. Annual Survey of Industries – all industries
   CSO No. of units, persons employed, capital output ratio, turnover, etc.
   http://www.mospi.gov.in/cs_o_test1.htm
   Information on existing units gives perspective on the Industrial development and helps in creating the employee profile

5. Monthly survey of selected industries
   CSO
   http://www.mospi.gov.in/cs_o_test1.htm
   Production statistics in detail
   Demand–supply estimations

6. Foreign Trade of
   Director General of Commercial Intelligence
   Exports and imports
   Forecast, manufacturing

7. Wholesale price index—weekly all India Consumer Price Index
   Ministry of Commerce and Industry
   http://india.gov.in/sectors/commerce/ministry_commerce.php
   Reporting of prices of products like food articles, foodgrains, minerals, fuel, power, lights, lubricants, textiles, chemicals, metal, machinery and transport
   Establishing price bands of product categories; pricing estimations for new products; determining consumer spend

8. Economic Survey – annual publication
   Dept. of Economic Affairs, Ministry of Finance, patterns, currency and finance
   http://finmin.nic.in/the_ministry/dept_eco_affairs/
   Descriptive reporting of the current economic status
   Estimations of the future and evaluation of policy decisions and extraneous factors in that period

9. National Sample Survey (NSS)
   Ministry of Planning
   http://www.planningcommission.gov.in/
   Social, economic, demographic, industrial and agricultural statistics.
   Significant for making policy decisions as well as studying sociological patterns

Other data sources: This source is the most voluminous and most frequently used, in every research study. The information could be

- Books and periodicals
- Guides: including Industry guides
- Directories and indices
- Standard non-governmental statistical data: Some non-governmental data sources are presented in Table 11.4.
### Table 11.4 Secondary Data—Non-government Publications

<table>
<thead>
<tr>
<th>Sub-type</th>
<th>Sources</th>
<th>Data</th>
<th>Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Status reports by various commodity boards</td>
<td>The commodity board or the industry associations like Jute Board, Cotton Industry, Sugar Association, Pulses Board, Metal Board, Chemicals, Spices, Fertilizers, Coir, Pesticides, Rubber, Handicrafts, Plantation Boards, etc.</td>
<td>Detailed information on current assets-in terms of units, current production figures and market condition</td>
<td>These are useful for individual sectors in working out their plans as well as evaluating causes of success or failure</td>
</tr>
<tr>
<td>4. Export related data — commodity wise</td>
<td>Leather Exports Promotion Council, Apparel Export Promotion Council, Handicrafts, Spices Tea, etc., Exim Bank etc. <a href="http://www.leatherindia.org/">http://www.leatherindia.org/</a> <a href="http://www.aepcindia.com/">http://www.aepcindia.com/</a></td>
<td>Product and country wise data on the export figures as well as information on existing policies related to the sector</td>
<td>To estimate the demand; gauge opportunities for trade and impetus required in terms of manufacturing and policy changes</td>
</tr>
<tr>
<td>5. Retail Store Audit on pharmaceutical, veterinary, consumer products</td>
<td>ORG (Operations Research Group). Monthly reports on urban sector. Quarterly reports on rural sector <a href="http://www.orgindia.org/">http://www.orgindia.org/</a></td>
<td>The touch point for this data is the retailer, who provides the figures related to product sales; the data is very comprehensive and covers most brands. The data is region specific and covers both inventory and goods sold</td>
<td>Market analysis and market structure mapping with estimations of market share of leading brands. The audit can also be used to study consumption trends at different time periods or subsequent to sales promotion or other activities</td>
</tr>
</tbody>
</table>
Questionnaire Design

NOTES

6. National Reader-ship Surveys (NRS)
   - IMRB survey of reading behaviour for different segments as well as different products
   - http://www.imrbint.com/

   Today these surveys are done by various bodies with different sample bases. Today the survey base has become younger, with the age of the reader lowered to 12+.

   Media planning and measuring exposure as well as reach for product categories

7. THOMPS ON INDICES:
   - Urban market index
   - Rural market index

   All towns with population of more than one lakh are covered and information of demographic and socio-economic variables are given for each city with Mumbai as base. The rural index similarly covers about 400 districts with socio-economic indicators like value of agriculture output, etc.

   The inclinations to purchase consumer products are directly related to socio-economic development of communities in general. The indices provide barometers to measure such potentials for each city and has implications for the researcher in terms of data collection sources

3. Computer-stored data

   Information today is also available in an electronic form. The databases available to the researcher can be classified on the basis of the type of information or by the method of storage and recovery as described below. Figure 11.5 gives a classification of the sources of computerized data.

   - Reference databases: These refer users to the articles, research papers, abstracts and other printed news contained in other sources. They provide online indices and abstracts and are thus also called bibliographic databases.
   - Source databases: These provide numerical data, complete text, or a combination of both.

   Based on storage and recovery mechanisms: Another useful way of classifying databases is based on their method of storage and retrieval.

   - Online databases: These can be accessed in real time directly from the producers of the database or through a vendor. Examples include ABI/Inform, EBSCO and Emerald.
   - CD-ROM databases: Here information is available on a CD-ROM.
4. Syndicated data sources

Syndicated service agencies are organizations that collect organization/product-category-specific data from a regular consumer base and create a common pool of data that can be used by multiple buyers, for their individual purpose.

There are different ways to classify syndicate sources.

- **Household/individual data**: These could be in the form of surveys or panel data available through reputed agencies.
- **Surveys**: Surveys are usually one-time assessments conducted on a large representative respondent base. Like opinion polls before elections, best business school to study.
- **Product purchase panels**: These specially selected respondent groups specifically record certain identified purchases, generally related to household products and groceries.
- **Media-specific panels**: Panels are also created for collecting information related to promotion and advertising. The task of the media panel is to make use of different kinds of electronic equipment to automatically record consumer viewing behaviour. These are used to calculate the television rating performance (TRP) of different programs.
- **Scanner devices and individual source systems**: To overcome the problems of panel data, a new service is provided by research agencies through electronic scanner devices—e.g. sales volume tracking data.
- **Institutional syndicated data**: The syndicated data can also be available at the institutional level. Retailer and wholesaler audits are examples of this kind. Usually the records are noted as:

  \[
  \text{Beginning stocks} + \text{deliveries} - \text{ending inventory} = \text{sales for the period}
  \]
Check Your Progress

8. How is secondary information collected helpful in for primary data collection instruments?
9. Cash registers and sales invoices are the examples of which type of secondary data sources?
10. Define syndicated service agencies.

11.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The first and foremost requirement of a questionnaire is that the spelt-out research objectives must be converted into clear questions which will extract answers from the respondent.

2. The degree of structure and the degree of concealment are the two criteria that are frequently used for designing purposes. Structure refers to the degree to which the response category has been defined. Concealment refers to the degree to which the purpose of the study is explained to the respondent.

3. The categories of questionnaires on the basis of method of administration are schedule and self-administered questionnaire.

4. The other names for open-ended questions are unstructured questions or free-response or free-answer questions.

5. Schedule is the method in which sampling control is the highest.

6. Questions that address sensitive issues are termed as loaded questions and the response to these questions might not always be honest, as the person might not wish to admit the answer. However, double-barred questions have two separate options which are separated by an ‘or’ or ‘and’. The researcher always splits them into two separate questions.

7. The greatest benefit of the questionnaire method is its adaptability.

8. The secondary information collected can be used to design the primary data collection instruments, in order to phrase and design the right questions.

9. Cash registers and sales invoices are the examples of sales data type of internal data sources.

10. Syndicated service agencies are organizations that collect organization/product-category-specific data from a regular consumer base and create a common pool of data that can be used by multiple buyers, for their individual purpose.
11.6 SUMMARY

- The questionnaire is a research technique that consists of a series of questions asked to respondents, in order to obtain statistically useful information about a given topic. It is one of the most cost-effective methods of collecting primary data, which can be used with considerable ease by most individual and business researchers.

- There are many different types of questionnaire available to the researcher. The categorization can be done on the basis of a variety of parameters. The two criteria that are most frequently used for designing purposes are the degree of structure and the degree of concealment.

- Formalized and unconcealed questionnaire is the most frequently used by all management researchers. This kind of structured questionnaire is easy to administer, and has both the questions as self-explanatory and the answer categories clearly defined.

- Formalized and concealed questionnaire have a formal method of questioning; however the purpose is not clear to the respondent. The research studies which are trying to find out the latent causes of behaviour and cannot rely on direct questions use these.

- The non-formalized and unconcealed questionnaire comes with unstructured questions where they have the freedom of expressing themselves the way they want to.

- In non-formalized and concealed questionnaire the researcher makes use of questions of low structure and disguised purpose.

- The questionnaire that has been prepared would necessitate a face-to-face interaction. In this case, the interviewer reads out each question and makes a note of the respondent’s answers. This administration is called a schedule. It might have a mix of the questionnaire type as described in the section above and might have some structured and some unstructured questions.

- The other kind is the self-administered questionnaire, where the respondent reads all the instructions and questions on his own and records his own statements or responses. Thus, all the questions and instructions need to be explicit and self-explanatory.

- The steps involved in questionnaire designing are conversion of research objectives into information areas, method of administration, content of the questionnaire, motivating the respondent to answer, determining the type of questions, criteria for question designing, determining the questionnaire structure, physical characteristics of the questionnaire, pilot testing of the questionnaire, and administrating the questionnaire.
Secondary data can be used for multiple purposes such as problem identification and formulation, hypotheses designing, sampling considerations, primary base, validation board.

Some of the advantages of using secondary data include resource advantage, accessibility of data, accuracy and stability of data, assessment of data.

Besides advantages, there might be some disadvantages of secondary data such as applicability of data and accuracy of data.

Secondary data can be divided into internal and external sources. Internal, as the name implies, is organization-or environment-specific source and includes the historical output and records available with the organization which might be the backdrop of the study. The data that is independent of the organization and covers the larger industry-scape would be available in the form of published material, computerized databases or data compiled by syndicated services. There are three major sources of data – internal, external, computer-stored data and syndicated databases.

Company records, employee records, sales data, financial records and sales reports are some of the sources of internal information.

Information that is collected and complied by an outside source that is external to the organization is referred to as external source of data. Published data, government sources, other data sources are the examples of external data source.

Reference databases refer users to the articles, research papers, abstracts and other printed news contained in other sources. They provide online indices and abstracts and are thus also called bibliographic databases.

Sources databases provide numerical data, complete text, or a combination of both.

Online databases can be accessed in real time directly from the producers of the database or through a vendor. Examples include ABI/Inform, EBSCO and Emerald.

Syndicated service agencies are organizations that collect organization/product-category-specific data from a regular consumer base and create a common pool of data that can be used by multiple buyers, for their individual purpose.

**11.7 KEY WORDS**

- **Questionnaire**: It refers to a research tool that consists of a series of questions asked to respondents, in order to obtain statistically useful information about a given topic.

- **Dichotomous Questions**: It refers to the questions with restrictive alternatives that provide the respondents only with two answers.
• **Secondary data**: It refers to the information which is not topical or research-specific and has been collected and compiled by some other researcher or investigative body.

• **Syndicated data**: It refers to the information gathered by a service or company for public release and sold by subscription.

### 11.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

#### Short Answer Questions

1. Define questionnaire.
2. Write short notes on the following:
   - (a) Formalized and concealed questionnaire
   - (b) Non-formalized and unconcealed questionnaire
   - (c) Non-formalized and concealed questionnaire
   - (d) Open-ended questions
3. What are closed-ended questions?
4. Briefly discuss the criteria for designing a questionnaire.
5. What are the advantages and disadvantages of secondary data?

#### Long Answer Questions

1. Describe the three kinds of formats of closed-ended questions.
2. What are the steps involved in the questionnaire design? Explain in detail the questionnaire design process.
3. What are the advantages and disadvantages of the questionnaire method? Illustrate with suitable examples.
4. Discuss the three major sources of secondary data in detail.

### 11.9 FURTHER READINGS

UNIT 12 ANALYSIS OF DATA

12.0 INTRODUCTION

Data analysis is the process of systematically applying statistical or logical techniques to describe, illustrate and evaluate data. The appropriate analysis of research findings is essentially important in order to maintain data integrity. Analysis of data plays a vital role in any research as it summarizes collected data and discuss the interpretation of outcomes using multiple data sources.

In this unit, you will study the process of data analysis and the techniques involved in the process. You will also be able to differentiate between descriptive analysis and inferential analysis and their types with appropriate examples. The unit goes on discussing the analysis of multiple responses and ordinal scaled questions. You will also learn the technique of analysing data by calculating rank order and using data transformation. In addition to this, the unit will also explore the significant role and meaning of research reporting.

12.1 OBJECTIVES

After going through this unit, you will be able to:

- Discuss the techniques used in the analysis of data
- Differentiate between descriptive and inferential analysis
- Describe the types of descriptive analysis
12.2 DATA ANALYSIS

Once the raw data is collected from both primary and secondary sources, the next step is to analyse the same so as to draw logical inferences from them. The data collected in a survey could be voluminous in nature, depending upon the size of the sample. In a typical research study there may be a large number of variables that the researcher needs to analyse. The analysis could be univariate, bivariate and multivariate in nature. In the univariate analysis, one variable is analysed at a time. In the bivariate analysis two variables are analysed together and examined for any possible association between them. In the multivariate analysis, the concern is to analyse more than two variables at a time.

The type of statistical techniques used for analysing univariate and bivariate data depends upon the level of measurements of the questions pertaining to those variables. Further, the data analysis could be of two types, namely, descriptive and inferential. Below is mentioned a list of illustrative set of questions which are answered under both descriptive and inferential analysis.

Descriptive vs Inferential Analysis

Descriptive Analysis

Descriptive analysis refers to transformation of raw data into a form that will facilitate easy understanding and interpretation. Descriptive analysis deals with summary measures relating to the sample data. The common ways of summarizing data are by calculating average, range, standard deviation, frequency and percentage distribution. The first thing to do when data analysis is taken up is to describe the sample. Below is a set of typical questions that are required to be answered under descriptive statistics:

- What is the average income of the sample?
- What is the average age of the sample?
- What is the standard deviation of ages in the sample?
- What is the standard deviation of incomes in the sample?
- What percentage of sample respondents are married?
- What is the median age of the sample respondents?
- Which income group has the highest number of user of product in question in the sample?
• Is there any association between the frequency of purchase of product and income level of the consumers?
• Is the level of job satisfaction related with the age of the employees?
• Which TV channel is viewed by the majority of viewers in the age group 20–30 years?

Types of descriptive analysis

The type of descriptive analysis to be carried out depends on the measurement of variables into four forms—nominal, ordinal, interval and ratio. Table 12.1 presents the type of descriptive analysis which is applicable under each form of measurement.

<table>
<thead>
<tr>
<th>Type of Measurement</th>
<th>Type of Descriptive Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>Frequency tables, Proportions, Percentages, Mode</td>
</tr>
<tr>
<td>Ordinal</td>
<td>Median, Quartiles, Percentiles, Rank order correlation</td>
</tr>
<tr>
<td>Interval</td>
<td>Arithmetic mean, Correlation coefficient</td>
</tr>
<tr>
<td>Ratio</td>
<td>Index numbers, Geometric mean, Harmonic mean</td>
</tr>
</tbody>
</table>

It is assumed that readers are acquainted with the methods of descriptive analysis as the material could be found in any elementary text on descriptive statistics. Here only a brief review of some of the methods is mentioned. In an inferential analysis, inferences are drawn on population parameters based on sample results. A necessary condition is that the sample should be drawn at random.

Inferential Analysis

After descriptive analysis has been carried out, the tools of inferential statistics are applied. Under inferential statistics, inferences are drawn on population parameters based on sample results. The researcher tries to generalize the results to the population based on sample results. The analysis is based on probability theory and a necessary condition for carrying out inferential analysis is that the sample should be drawn at random.

The following is an illustrative list of questions that are covered under inferential statistics.
• Is the average age of the population significantly different from 35?
• Is the average income of population significantly greater than ₹25,000 per month?
• Is the job satisfaction of unskilled workers significantly related with their pay packet?
• Do the users and non-users of a brand vary significantly with respect to age?
• Is the growth in the sales of the company statistically significant?
• Does the advertisement expenditure influences sale significantly?
• Are consumption expenditure and disposable income of households significantly correlated?
• Is the proportion of satisfied workers significantly more for skilled workers than for unskilled workers?
• Do urban and rural households differ significantly in terms of average monthly expenditure on food?
• Is the variability in the starting salaries of fresh MBA different with respect to marketing and finance specialization?

**Descriptive Analysis of Univariate**

Univariate procedures deal with analysis of one variable at a time. In this chapter only a brief review of various techniques is given. The first step under univariate analysis is the preparation of frequency distributions of each variable. The frequency distribution is the counting of responses or observations for each of the categories or codes assigned to a variable. Consider a nominal scale variable—gender of respondents.

Table 12.2 shows both the raw frequency and the percentages of responses for each category in case of the variable gender.

<table>
<thead>
<tr>
<th>Gender of the respondent</th>
<th>Frequency</th>
<th>Per cent</th>
<th>Valid Per cent</th>
<th>Cumulative Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>301</td>
<td>72.7</td>
<td>72.7</td>
<td>72.7</td>
</tr>
<tr>
<td>Female</td>
<td>113</td>
<td>27.3</td>
<td>27.3</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>414</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This tabulation process can be done by hand using tally marks. However, in case of large sample, the frequency distribution table is prepared using computer software. In the present case, SPSS software is used. The results indicate that out of a sample of 414 respondents, 301 are male and 113 are female. The raw frequencies are often converted into percentages as they are more meaningful. In the present case, for example, there are 72.7 per cent male and 27.3 per cent female respondents.

**Missing Data**

There are situations when certain questions knowingly or unknowingly are not answered by the respondents. The responses corresponding to such respondents are treated as 'missing data'. The frequency distribution in case of the variable ‘marital status’ is presented in Table 12.3.

<table>
<thead>
<tr>
<th>Marital status of respondents</th>
<th>Frequency</th>
<th>Per cent</th>
<th>Valid Per cent</th>
<th>Cumulative Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>285</td>
<td>68.8</td>
<td>68.0</td>
<td>68.0</td>
</tr>
<tr>
<td>Married</td>
<td>129</td>
<td>31.2</td>
<td>91.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>414</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>414</td>
<td>100.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The respondent who did not answer the question on ‘marital status’ is coded as nine, which is being treated as the missing data. The missing value could as well be coded with another number. The only precaution to be kept in mind is that a missing observation should be assigned a number that should not be equal to the value of the variable obtained as part of the survey. If the value of the missing observation was available; it could perhaps lead to different research conclusions. The intensity of the deviation of the actual results from the observed depends upon the number of missing observations and the extent to which the missing data would be different from actual observation.

In case of Table 12.3, it may be noted that out of a sample of 414 respondents, 285 are single, 128 are married and one observation is missing. In the column on ‘per cent’ in this table, it is indicated that 68.8 per cent are single, 30.9 per cent are married and 0.2 per cent are missing observation. Here, the percentages are computed on a total sample of 414. As it is known that one observation is missing, the actual sample for this variable should be 413. Therefore, a column named ‘valid per cent’ has been included, where the percentages are computed based on a sample of 413. The result using the ‘valid per cent’ column indicates that 69.0 per cent of respondents are single, whereas 31 per cent are married. The results in both cases are almost similar. This is so because there was only one single missing value. Generally, if the volume of missing data is small, it is unlikely to affect the conclusion from the analysis. This may not always be the case. It is for this reason that the ‘valid per cent’ column should be used for interpreting the results.

Table 12.4 gives the frequency distribution of time of the day preferred to use café. It may be noted from this table that the number of missing observations in this case is 48, amounting to 11.6 per cent of the sample. As a consequence of this, the results of ‘per cent’ and ‘valid per cent’ vary, especially for ‘afternoon’, ‘evening’ and ‘night’ response categories.

Table 12.4 Preferred time of the day for using cyber café

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Per cent</th>
<th>Valid Per cent</th>
<th>Cumulative Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>Morning</td>
<td>176</td>
<td>4.3</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>Noon</td>
<td>191</td>
<td>4.6</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>Afternoon</td>
<td>111</td>
<td>2.7</td>
<td>13.1</td>
</tr>
<tr>
<td></td>
<td>Evening</td>
<td>178</td>
<td>43.0</td>
<td>75.1</td>
</tr>
<tr>
<td></td>
<td>Night</td>
<td>111</td>
<td>27.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Missing</td>
<td>Total</td>
<td>486</td>
<td>11.8</td>
<td></td>
</tr>
</tbody>
</table>

It may be worth considering a variable where the cumulative frequencies in percentages may be very useful in interpretation of the results. Table 12.5 presents
the frequency distribution of monthly household income of 414 respondents. It may be noted that there are 10 missing observations in this table. Therefore, the analysis should be applicable using a sample of 404 respondents. As discussed earlier the ‘valid per cent’ column should be used for interpretation of the results.

For example, the results indicate that 20.5 per cent of the respondents have a monthly household income of ₹10,000 to ₹19,999, whereas 4.7 per cent of respondents have monthly income of ₹65,000 and more. The last column of Table 12.5 presents cumulative per cent. The results in Table 12.5 indicate that while 27 per cent of the respondents have a monthly household income less than or equal to ₹19,999, there are 95.3 per cent of them that have income less than or equal to ₹64,999.

### Table 12.5 Monthly household income of cyber café users

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Per cent</th>
<th>Valid Per cent</th>
<th>Cumulative Per cent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than ₹10,000</td>
<td>25</td>
<td>5.3</td>
<td>5.3</td>
</tr>
<tr>
<td>₹10,000 to ₹19,999</td>
<td>33</td>
<td>20.2</td>
<td>20.5</td>
</tr>
<tr>
<td>₹20,000 to ₹59,999</td>
<td>129</td>
<td>31.2</td>
<td>31.9</td>
</tr>
<tr>
<td>₹60,000 to ₹99,999</td>
<td>171</td>
<td>39.7</td>
<td>41.4</td>
</tr>
<tr>
<td>₹100,000 to ₹199,999</td>
<td>24</td>
<td>5.9</td>
<td>39.3</td>
</tr>
<tr>
<td>₹200,000 and above</td>
<td>19</td>
<td>4.7</td>
<td>44.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>404</strong></td>
<td><strong>97.6</strong></td>
<td><strong>100.0</strong></td>
</tr>
<tr>
<td>Missing</td>
<td>9</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>414</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>

### Analysis of Multiple Responses

At times, the researcher comes across multiple category questions where respondents could choose more than one answer. In such a case, the preparation of frequency table and its interpretation is slightly different. If the question in the research study is multiple category question and the respondents are allowed to tick more than one choice, the percentage in such a case may not add up to 100. For example, one may consider the following question:

**When accessing the internet at a cyber café, tick up to frequently used applications for which you use the cyber café.**

1. E-mail
2. Chat
3. Browsing
4. Downloading
5. Shopping
6. Net telephony
7. Business and Commerce (e-commerce)
8. Entertainment
9. Adult sites  
10. Astrology and Horoscope  
11. Education  
12. Any other, please specify.

The resulting frequency table for the above-mentioned question is as presented in Table 12.6.

Table 12.6 Frequently used applications at cyber café

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Application</th>
<th>Frequencies</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>E-mail</td>
<td>399</td>
<td>94.9</td>
</tr>
<tr>
<td>2</td>
<td>Chatting</td>
<td>216</td>
<td>70.3</td>
</tr>
<tr>
<td>3</td>
<td>Browsing</td>
<td>108</td>
<td>26.7</td>
</tr>
<tr>
<td>4</td>
<td>Downloading</td>
<td>197</td>
<td>47.8</td>
</tr>
<tr>
<td>5</td>
<td>Shopping</td>
<td>30</td>
<td>7.2</td>
</tr>
<tr>
<td>6</td>
<td>Net Banking</td>
<td>30</td>
<td>7.2</td>
</tr>
<tr>
<td>7</td>
<td>E-commerce</td>
<td>51</td>
<td>12.3</td>
</tr>
<tr>
<td>8</td>
<td>Leatherware</td>
<td>19</td>
<td>4.6</td>
</tr>
<tr>
<td>9</td>
<td>Adult sites</td>
<td>50</td>
<td>14.3</td>
</tr>
<tr>
<td>10</td>
<td>Astrology and horoscopes</td>
<td>258</td>
<td>62.2</td>
</tr>
<tr>
<td>11</td>
<td>Education</td>
<td>109</td>
<td>26.6</td>
</tr>
<tr>
<td>12</td>
<td>Any Other</td>
<td>14</td>
<td>3.4</td>
</tr>
<tr>
<td>TOTAL RESPONDENTS</td>
<td>416</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Total exceeds 100% because of multiplicity of answers.

In Table 12.6 the percentages are computed on the total sample size of 414. If these percentages are added up, they would exceed more than 100 per cent. This is because of multiplicity of answers as respondents were given the chance to choose more than one answer. The interpretation of the table would be based on a sample of 414 and is given as:

- The most used application at a cyber café is e-mail. It is seen that 94.9 per cent of the users make use of this.
- The second popular application is chatting, and 76.3 per cent of the sample respondents make use of it.
- Similarly, other applications in order of preference are browsing (56 per cent), downloading (47.6 per cent), education 35.4 per cent), entertainment (32.6 per cent) and so on.

Analysis of Ordinal Scaled Questions

It is quite likely that there may be some respondents who might have used more than one brand of toothpaste in the last one year. These could be Colgate, Pepsodent, Close up, Neem, Sensodyne etc. The respondents could be asked to rank their preference for toothpaste. The question before the researcher is how to tabulate and interpret the responses to such questions. It could be done in two
ways as would be shown in the following example. The questions asked of the respondents in such a case could be:

- Rank the following five attributes while choosing a restaurant for dinner. Assign a rank of 1 to the most important, 2 to the next important . . . and 5 to the least important.
  - Ambience
  - Food quality
  - Menu variety
  - Service
  - Location

From a sample of 32, the responses obtained are given in Table 12.7. To construct univariate tables out of the given data, one can take up one column at a time from Table 12.7 and prepare the separate frequency tables. For example, distribution of rank assigned to attribute food quality may be considered in Table 12.8.

Table 12.7  Ranking of various attributes while selecting a restaurant for dinner

<table>
<thead>
<tr>
<th>Respondent No.</th>
<th>Ambience</th>
<th>Food Quality</th>
<th>Menu Variety</th>
<th>Service</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>18</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>1</td>
<td>9</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>20</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>27</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>29</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>31</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>
It is seen from Table 12.8 that out of 32 respondents, 16 (50 per cent) have assigned rank one, 13 (40.6 per cent) have assigned rank two, 2 (6.3 per cent) have assigned rank three and 1 (3.1 per cent) has assigned rank four to food quality. This shows that food quality is given a lot of importance by the respondents. Similar analysis could be carried out for other attributes.

The other way of preparing a univariate table could be to find distribution of attribute which got various ranks. Table 12.9 indicates the distribution of attributes that received rank one.

Table 12.9 indicates that 50 per cent of the respondents gave food quality rank one, whereas 21.88 per cent gave menu variety as rank one, followed by ambience that was ranked one by 12.5 per cent of the respondents. Similar analysis could be carried out corresponding to the remaining attributes.

Calculating Rank Order

In survey research, it is generally observed that respondents may be asked to indicate a rank ordering of various attributes of a product or rank ordering of brand preference or some other variable of interest. For example, data presented in Table 12.7 gives the ranking by 32 respondents on five attributes while choosing a restaurant for dinner. The data given in Table 12.7 can be used to prepare the summarized rank ordering of various attributes. The rankings of attributes given in Table 12.7 can be presented in the form of frequency distribution in Table 12.10.
Analysis of Data

Table 12.10 Frequency table of the rankings of the attributes while selecting a restaurant for dinner

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Ambience</td>
<td>4</td>
</tr>
<tr>
<td>Food Quality</td>
<td>16</td>
</tr>
<tr>
<td>Menu Variety</td>
<td>5</td>
</tr>
<tr>
<td>Service</td>
<td>2</td>
</tr>
<tr>
<td>Locations</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
</tr>
</tbody>
</table>

To calculate a summary rank ordering, the attribute with the first rank was given the lowest number (1) and the least preferred attribute was given the highest number (5).

The summarized rank order is obtained with the following computations as:

\[
\text{Ambience} = (4 \times 1) + (5 \times 2) + (3 \times 3) + (3 \times 4) + (5 \times 5) = 90
\]

\[
\text{Food Quality} = (16 \times 1) + (13 \times 2) + (2 \times 3) + (2 \times 4) + (9 \times 5) = 120
\]

\[
\text{Menu Variety} = (5 \times 1) + (2 \times 2) + (3 \times 3) + (1 \times 4) + (5 \times 5) = 110
\]

\[
\text{Service} = (2 \times 1) + (3 \times 2) + (11 \times 3) + (11 \times 4) + (3 \times 5) = 121
\]

The total lowest score indicates the first preference ranking. The results show the following rank ordering:

1. Food quality
2. Service
3. Ambience
4. Menu variety
5. Location

Data Transformation

Under data transformation, the original data is changed to a new format for performing data analysis so as to achieve the objectives of the study. This is generally done by the researcher through creating new variables or by modifying the values of the scaled data. The following illustrations show how it is carried out:

(a) It is usually believed by researchers that the response bias will be less if instead of asking the question on the exact age, the question is asked on the date of birth. This does not create any problem in data analysis as having known the date of birth, it is always possible to compute the exact age of the respondent.

(b) At times it may become essential to collapse or combine adjacent categories of a variable so as to reduce the number of categories of original variables. In a 5-point Likert scale, having categories like strongly agree, agree, neither...
agree nor disagree, disagree and strongly disagree can be clubbed into three categories. One can combine strongly agree and agree category into one category. Similarly, disagree and strongly disagree responses could be clubbed into a separate category and neither agree nor disagree could be treated as a separate category. This is how a five-category scale can be collapsed into a three-category one.

(c) The researcher could create new variables by re-specifying the data with numeric or logical transformation. Suppose a multiple-item Likert scale designed to measure the perception of a customer towards the bank has 10 items. The total score of a respondent can be computed as:

\[
\text{Total score of } i^{\text{th}} \text{ respondent} = \text{Score of } i^{\text{th}} \text{ respondent on item 1} + \text{Score of } i^{\text{th}} \text{ respondent on item 2} + \ldots + \text{Score of } i^{\text{th}} \text{ respondent on item 10}.
\]

Once the total score for each of the respondent is computed, the average score can be obtained by dividing it by the number of items. It can be further categorized as favourable, neutral and unfavourable perception that could be related to various demographic variables depending upon the objectives of research.

12.3 RESEARCH REPORTING

The research report has a very important role to play in the entire research process. It is a concrete proof of the study that was undertaken. It is a one-way communication of the researcher’s study and analysis to the reader/manager, and thus needs to be all-inclusive and yet neutral in its reporting. The significant role that a research report can play is as follows:

- The research report list the steps beginning with the framing the research question to the interpretation of the study findings.
- Each step also includes details on how and why that step was conducted, i.e. the justification for choosing one technique over the other.
- It also serves to authenticate the quality of the work carried out and establishes the strength of the findings obtained.
- The report gives a clear direction in terms of the implication of the results for the decision maker. This could be academic or applied depending on the orientation.
- The report serves as a very important framework for anyone who would like to do research in the same area or topic.

12.3.1 Brief Reports

These kinds of reports are not formally structured and are generally short, sometimes not running more than four to five pages. The information provided has limited scope and is a prelude to the formal structured report that would subsequently follow. These reports could be designed in several ways.
12.3.2 Detailed reports

These are more formal and could be academic, technical or business reports.

- **Technical reports**: These are major documents and would include all elements of the basic report, as well as the interpretations and conclusions, as related to the obtained results. This would have a complete problem background and any additional past data/records that are essential for understanding and interpreting the study results. All sources of data, sampling plan, data collection instrument(s), data analysis outputs would be formally and sequentially documented.

- **Business reports**: These reports include conclusions as understood by the business manager. The tables, figures and numbers of the first report would now be pictorially shown as bar charts and graphs and the reporting tone would be more in business terms. Tabular data might be attached in the appendix.

12.3.3 Interpretation

This section comes after the main report and contains interpretations of results and suggested recommendations. It presents the information in a summarized and numerical form.

Sometimes, the research results obtained may not be in the direction as found by earlier researchers. Here, the skill of the researcher in justifying the obtained direction is based on his/her individual opinion and expertise in the area of study. After the interpretation of results, sometimes, the study requirement might be to formulate indicative recommendations to the decision-makers as well. Thus, in case the report includes recommendations, they should be realistic, workable and topical related to the industry studied.

12.3.4 Presentation

It has been observed that report and presentation are the media of conveying one’s ideas to a group of people. However, the most significant difference between the two of them is that the individual making the presentation becomes visible and is able to connect in a better way with his audience.
A report is also a structured presentation of the results of a research that interprets facts and figures and becomes valuable for suggestions. Its foundation is guided by research followed by collection, development and validation of knowledge. Once all the information/data has been collected it is presented in a structured or logical manner in the form of a presentation. A presentation is also developed in the same manner like the report; however, the most prominent difference is that of the human element.

Let us briefly study the essential elements of a presentation.

1. **Data:** The making of a presentation begins with collection of information to be put before an audience. However, in the case of a report, the audience has the advantage of going through the information at his or her own pace while this is not the same in a presentation.

2. **Structure:** The presenter must ensure that the data or information being presented should have a logical flow, i.e., proper beginning, middle and end. The main challenge for a presenter is to keep the audience engaged in the presentation. In a report, while the reader is going through its contents, he is guided by the endnotes, footnotes and suggested readings which is not the case of a presentation.

3. **Presentation:** Another significant difference between a presentation and a report is that a report can be read and read again but not a presentation. In case, a linking aspect is missed in the presentation then the logical sequencing of the flow of content get eliminated for the audience thereby, affecting the understanding of the audience.

4. **Human element:** The most significant differentiating factor is that of the presenter making the presentation. Largely, a presentation is remembered by the audience due to the presenter making the presentation. Here it needs to be mentioned that the presenter through his voice quality, accent and pronunciation grasps the attention of the audience. Also, the body language of the presenter plays a vital role in making the presentation successful.

### Check Your Progress

1. What is univariate, bivariate and multivariate analysis?
2. Define descriptive analysis.
3. Name the types of descriptive analysis.
4. On which theory is the inferential analysis based?
5. What is research reporting?
6. What is the purpose of working papers?
12.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

NOTES

1. In the univariate analysis, one variable is analysed at a time. In the bivariate analysis two variables are analysed together and examined for any possible association between them. In the multivariate analysis, the concern is to analyse more than two variables at a time.

2. Descriptive analysis refers to transformation of raw data into a form that will facilitate easy understanding and interpretation. Descriptive analysis deals with summary measures relating to the sample data. The common ways of summarizing data are by calculating average, range, standard deviation, frequency and percentage distribution.

3. The type of descriptive analysis to be carried out depends on the measurement of variables into four forms – nominal, ordinal, interval and ratio.

4. The inferential analysis is based on probability theory and a necessary condition for carrying out inferential analysis is that the sample should be drawn at random.

5. The research report has a very important role to play in the entire research process. It is a concrete proof of the study that was undertaken. It is a one-way communication of the researcher’s study and analysis to the reader / manager, and thus needs to be all-inclusive and yet neutral in its reporting.

6. Working papers or basic reports are written for the purpose of recording the process carried out in terms of scope and framework of the study, the methodology followed and instrument designed. The results and findings would also be recorded here. However, the interpretation of the findings and study background might be missing, as the focus is more on the present study rather than past literature.

12.5 SUMMARY

- Once the raw data is collected from both primary and secondary sources, the next step is to analyse the same so as to draw logical inferences from them. The data collected in a survey could be voluminous in nature, depending upon the size of the sample.

- In a typical research study there may be a large number of variables that the researcher needs to analyse. The analysis could be univariate, bivariate and multivariate in nature. In the univariate analysis, one variable is analysed at a time. In the bivariate analysis two variables are analysed together and
examined for any possible association between them. In the multivariate analysis, the concern is to analyse more than two variables at a time.

- The type of statistical techniques used for analysing univariate and bivariate data depends upon the level of measurements of the questions pertaining to those variables.

- Descriptive analysis refers to transformation of raw data into a form that will facilitate easy understanding and interpretation. Descriptive analysis deals with summary measures relating to the sample data. The common ways of summarizing data are by calculating average, range, standard deviation, frequency and percentage distribution.

- The type of descriptive analysis to be carried out depends on the measurement of variables into four forms—nominal, ordinal, interval and ratio.

- In an inferential analysis, inferences are drawn on population parameters based on sample results. A necessary condition is that the sample should be drawn at random.

- Univariate procedures deal with analysis of one variable at a time. In this chapter only a brief review of various techniques is given. The first step under univariate analysis is the preparation of frequency distributions of each variable.

- The intensity of the deviation of the actual results from the observed depends upon the number of missing observations and the extent to which the missing data would be different from actual observation.

- At times, the researcher comes across multiple category questions where respondents could choose more than one answer. In such a case, the preparation of frequency table and its interpretation is slightly different. If the question in the research study is multiple category question and the respondents are allowed to tick more than one choice, the percentage in such a case may not add up to 100.

- Under data transformation, the original data is changed to a new format for performing data analysis so as to achieve the objectives of the study. This is generally done by the researcher through creating new variables or by modifying the values of the scaled data.

- The researcher could create new variables by re-specifying the data with numeric or logical transformation.

- The research report has a very important role to play in the entire research process. It is a concrete proof of the study that was undertaken. It is a one-way communication of the researcher’s study and analysis to the reader/manager, and thus needs to be all-inclusive and yet neutral in its reporting.
The report gives a clear direction in terms of the implication of the results for the decision maker. This could be academic or applied depending on the orientation.

Brief reports are not formally structured and are generally short, sometimes not running more than four to five pages. The information provided has limited scope and is a prelude to the formal structured report that would subsequently follow.

Basic reports could be designed in several ways, namely working papers or basic reports and survey reports.

Detailed reports are more formal and could be academic, technical or business reports.

Technical reports are major documents and would include all elements of the basic report, as well as the interpretations and conclusions, as related to the obtained results. This would have a complete problem background and any additional past data/records that are essential for understanding and interpreting the study results. All sources of data, sampling plan, data collection instrument(s), data analysis outputs would be formally and sequentially documented.

Business reports include conclusions as understood by the business manager. The tables, figures and numbers of the first report would now be pictorially shown as bar charts and graphs and the reporting tone would be more in business terms.

This section comes after the main report and contains interpretations of results and suggested recommendations. It presents the information in a summarized and numerical form.

In the management report, the information on the sampling techniques follows the research intention, and the questionnaire design details need not be reported. The review of past literature would be perfunctory in the management report; however, they would be detailed and accompanied with the bibliography in the technical report.

Then come the background section, which includes the problem statement, introduction, study background, scope and objectives of the study and the review of literature (depends on the purpose).

In the management report, the sequencing of the report might be reversed to suit the needs of the decision-maker, as here the reader needs to review and absorb the findings.

Thus, the entire research project needs to be recorded either as a single written report or into several reports, depending on the need of the readers.
12.6 KEY WORDS

- **Data Analysis**: It refers to the process of inspecting, cleansing, transforming, and modelling data with the goal of discovering useful information, informing conclusions and supporting decision-making.
- **Likert scale**: A Likert scale is a psychometric scale commonly involved in research that employs questionnaires.
- **Rank Order**: It refers to a scale which gives the respondent a set of items and asks them to put the items in some form of order.
- **Data Transformation**: It refers to the process of converting data from one format or structure into another format or structure.
- **Research Report**: It refers to a document prepared by an analyst or strategist who is a part of investment research team.

12.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**
1. What do you mean by data analysis?
2. What is inferential analysis?
3. How does a researcher analyse multiple responses?
4. Briefly explain calculation of rank order with the help of an example.
5. Define data transformation.
6. Write short note on brief reports.

**Long Answer Questions**
1. How is data transformation carried out? Explain.
2. Discuss the role of research reporting.
4. Examine the types of detailed reports.

12.8 FURTHER READINGS

Analysis of Data

NOTES


UNIT 13 STATISTICAL ANALYSIS SOFTWARE

Structure
13.0 Introduction
13.1 Objectives
13.2 Statistical Analysis Software: Overview
13.3 Statistical Package for Social Sciences (SPSS)
13.4 Statistical Analysis Software: PSPP and SOFA
13.5 Answers to Check Your Progress Questions
13.6 Summary
13.7 Key Words
13.8 Self Assessment Questions and Exercises
13.9 Further Readings

13.0 INTRODUCTION

Statistical Analysis involves the collecting and examining every data sample. In every sector, organizations depend on data and analysis to offer and gain new insights to the growing market trend and gain a competitive edge. A statistical analysis is aimed to identify growing trends and statistical analysis software provides wide-ranging statistical capabilities that aim to meet the expanding requirements. Statistical software are specialized computer programs that are able to solve complex business and statistical problems.

In this unit, you will study about the number of specific software programs used in statistical analysis. You will also study about SPSS which is one of the most popular software package for statistical analysis. Apart from SPSS, you will also learn about PSPP, the alternative for SPSS and also about SOFA which is an open source user friendly statistical software package considered as an innovative analysis tool for statistical calculation and reporting.

13.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the System for Statistical Analysis (SAS)
- Carry out statistical preparation of data
- Discuss the popular software packages to perform statistical analysis like SPSS
- Describe the usage, installation of SOFA statistical software
13.2 STATISTICAL ANALYSIS SOFTWARE: OVERVIEW

Researchers have advantage of a wide array of statistical programmes to assist them in both data management and data analysis. In this section we will briefly discuss only the most frequently used packages.

**MS Excel:** The simplest and most widely used method of presenting and tabulating data is on Excel. The basic mathematical functions can be calculated here. Secondly, the software is easy to understand and used by most computer users. The data entered on Excel can be transported to most statistical packages for a higher level analysis.

**Minitab:** Minitab Inc. was developed more than 20 years ago at the Pennsylvania State University. It can be used with considerable ease and effectiveness in all business areas. It was originally used by statisticians. However, today it is used for multiple applications—especially quality control, six sigma and the design of experiments. The URL for Minitab is http://www.minitab.com/. The researcher can utilize the products and help the guide to undertake a quantitative research analysis.

**Statistical Analysis System (SAS):** SAS was created in the late 1960s at North Carolina State University. It has been actively and extensively used in managing, storing and analysing information. It has the advantage of being able to manage really bulky data sets with considerable ease. Linear models (Regression, Analysis of variance, Analysis of covariance), Generalized linear models (including Logistic regression and Poisson regression), multivariate methods (MANOVA, Canonical correlation, Discriminant analysis, Factor analysis, Clustering), categorical data analysis (including log-linear models), and all the standard techniques for descriptive and confirmatory statistical analysis are possible with SAS. The statistical analyses may be interfaced with the graphical products to produce relevant plots such as q-q plots, residual plots, and other relevant graphical descriptions of the data. Forecasting and trend series can also be carried out using the package. It finds a higher usage amongst industry than students who are more comfortable with SPSS. The URL for package is http://www.sas.com/.

**SPSS:** Amongst the student community as well as with most research agencies, this is the most widely used package. It is adaptable to most business problems.

There are a number of specific software programs like E Views for business forecasting and LISREL (Linear Structural Relations) for structural equation modelling. However, for most purposes, SPSS is the most widely used software.
13.3 STATISTICAL PACKAGE FOR SOCIAL SCIENCES (SPSS)

Statistical Package for Social Sciences (SPSS) is one of the most popular software packages to perform statistical analysis on survey data. Its first version was released in 1968 and since then, it has come a long way. It is used by researchers in educational institutes, research organizations, government, marketing firms, etc.

Launching SPSS

To start SPSS, go to Start -> Programs -> SPSS followed by its version. For example, SPSS 12, SPSS 14, SPSS 16, SPSS 17.

A dialog box will open in front of SPSS grid listing several options to choose from. The following options will appear in the dialog box:

- Run the tutorial
- Type in data
- Run in existing query
- Create new query using Database Wizard
- Open an existing data source
- Open another type of file

For the moment, we will concentrate on the second option, i.e., Type in data. Select this option and click Ok. By default, the Data Editor view is initially selected.

SPSS Data Editor

The SPSS Data Editor Window has two views: Data View and Variable View. Variable View is used to define variables that will store the data. Data View contains the actual data.
The first step is to open the 'Variable View' window of the Data Editor and define variables. Let us consider an example where Employee Data of an organization needs to be saved and analysed. The objective is to create a small data file for employees that consist of six variables as given in the following Table.

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Variable type</th>
</tr>
</thead>
<tbody>
<tr>
<td>EmpID</td>
<td>Numeric</td>
</tr>
<tr>
<td>EmpName</td>
<td>String</td>
</tr>
<tr>
<td>Gender</td>
<td>Numeric</td>
</tr>
<tr>
<td>Age</td>
<td>Numeric</td>
</tr>
<tr>
<td>Income</td>
<td>Numeric</td>
</tr>
<tr>
<td>MaritalStatus</td>
<td>Numeric</td>
</tr>
</tbody>
</table>

There are different types of variables in SPSS, the default one being numeric. To change variable type, in Variable View click on the variable in the column Type. A window similar to one below will open. Create all the variables and select appropriate Type as given in the table above.

**Note:**

While defining variable names empty spaces are not allowed.

- E.g., Marital Status – Not allowed
- MaritalStatus or Marital_Status – Correct

The third column in Variable view is **Width**, which specifies the number of characters allowed to be entered in the column. By default the width is 8 characters and can be modified depending upon the data being entered.

The fourth column is **Decimals**, which represents the number of decimal places. For numeric data type the default value is 0. Say, for example, EmpID does not require decimal places, therefore, it can be set to 0.

The fifth column is **Label**, which describes the variable.
The sixth column is *Values*. For example, Gender contains two categories (Female = 1 and Male = 2). In Data View, the gender will be entered as either 1 or 2. But what 1 or 2 represents is given in the Values as 1 represents Female and 2 Male.

The seventh column is *Missing*. Often while collecting data, you will have missing values within your data. This column is used in cases where no data is provided by a respondent. A missing value is chosen as an impossible value for that column. For example, the missing value for age can be entered as 1000 or -100 which are impossible entries for age. The objective of giving a missing value is to exclude that record while analysing the data.

The eighth column is *Columns*. It represents the width of the column. Default value is 8 and can be changed.

The ninth column is *Align*, which aligns the data at the left, centre or right of cell.

The last column is *Measure*. It can take values of Nominal, Ordinal or Scale.

The table below shows the different types of measurement, with examples:

<table>
<thead>
<tr>
<th>Nominal</th>
<th>Category</th>
<th>Discrete</th>
<th>Eye colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordinal</td>
<td>Ranking</td>
<td>Discrete</td>
<td>Ranking preference for various soft drinks</td>
</tr>
<tr>
<td>Interval</td>
<td>Scale</td>
<td>Continuous</td>
<td>Temperature</td>
</tr>
<tr>
<td>Ratio</td>
<td>Scale</td>
<td>Continuous</td>
<td>Age, years of education</td>
</tr>
</tbody>
</table>

**Nominal Data**: Discrete/category variable (limited number of values), e.g., Gender (Male or Female), Days of the week, Yes/No response in a questionnaire.

**Ordinal Data**: Discrete/category variable (limited number of ranks).

**Interval Data**: Continuous Data.

**Ratio Data**: Continuous Data.

Category or discrete measure consists of values that can be grouped into categories, for example, gender, which can be grouped into male and female. A category variable can be a string variable or a numeric variable but it is recommended that categorical variables should be numeric because strings contain letters which cannot be numerically analysed. Therefore, rather than representing female as ‘f’ and male as ‘m’, it is recommended as stated earlier in the chapter, where possible, use numeric values instead of letters when coding and entering data, e.g., use 1 for female and 2 for male.

Continuous measure is not restricted to specific values and is usually measured on a continuous scale, such as distance from home to office (in km). It will vary from individual to individual on a scale as given below.

<table>
<thead>
<tr>
<th>0 km</th>
<th>Distance between home and office (in km)</th>
<th>100 km</th>
</tr>
</thead>
</table>
Enter some data for the variables created in the Variable View. The Data View grid will look something like shown below:

**Recoding Variables**

Recode is a very important feature in SPSS, which is used to convert continuous data into discrete or category data. One can recode values within the existing variable into a new variable.

**Notes**: If you recode the values into the existing variable, the old values are lost. So it is recommended to recode a variable into a new variable wherever possible, so that your original values are retained.

Recode is available under Transform menu. There are three ways to recode the data.

1. Recode into same variables
2. Recode into new variables
3. Automatic recode

Now suppose, the variable income is to be categorized into three income categories based upon the below logic.

- \(< 10000 \rightarrow 1 \text{ (Low income)}
- 10000 \rightarrow 30000 \rightarrow 2 \text{ (Middle income)}
- \(> 30000 \rightarrow 5 \text{ (High income)}

Go to Transform -> Recode into new variable. The variable income will be recoded into a new variable (IncomeRe) labeled as Income Redefined which is the Output Variable.

Click on the button Old and New Values. A window will open divided into two parts. Left side will be Old Value and right side shows New Value.

Since the first category is 10000, the Old Value option to be selected will be Range, Lowest through value: 10,000. New Value is 1.

The second category is a range >10000 and 30,000, the Old Value option to be selected is a Range, i.e., 10,000 through 30,000. New Value is 2.

The third category is > 3000, the Old value option to be selected is Range, value through Highest: 30,000. New Value is 3.

A snapshot of the recode screen is given below for reference. Click on Continue and Ok.

A new variable IncomeRe will be created based upon the income variable. Next, we need to label what are 1, 2 and 3 values. Go to Variable View and give the labels for the new variable IncomeRe.

13.4 STATISTICAL ANALYSIS SOFTWARE: PSPP AND SOFA

PSPP is considered as an alternative for the data management and analysis tool SPSS. It is an open source application which is used for statistical analysis of sampled data. The software provides a complete set of capabilities including frequencies, cross-tabs comparison of means (t-tests and one-way ANOVA).
The PSPP product was launched at the end of the 1990s as a free software replacement for SPSS. It is considered a powerful tool for data visualization, data pre-processing, hypothesis testing and data analysis.

PSPP is developed using the C programming language and can work on most operating systems and computer hardware. This software can work on most operating systems and it comes with one unique feature which is its interoperability with other software applications. The application is capable of performing various tasks such as:

- Data re-ordering
- Linear and logistic regression
- Reliability testing
- Factor analysis
- Principal components analysis

SOFA is an open source user friendly statistical software package considered as a pioneering analysis tool for statistical calculation and reporting. The name SOFA stands for ‘Statistics Open For All’. SOFA has a graphical user interface (GUI) and can directly connect to MySQL, PostgreSQL, SQLite, MS Access, Microsoft SQL Server, and CUBRID. In addition, the user can also import the required ‘Data’ from CSV or (common separated values) and tab-separated files or spreadsheets (Microsoft Excel, OpenOffice.org Calc, Gnumeric, Google Docs).

The key statistical tests that are accessible or included in SOFA statistical package are independent and paired t-Tests, Wilcoxon Signed Ranks, Mann–Whitney U Test, Pearson’s R Test, Chi-Square Test, Kruskal Wallis H Test, One-Way ANOVA, Spearman’s R Test, etc. The SOFA statistical software can produce ‘Nested Tables’ with rows and columns for calculating percentages, totals, standard deviation, mean, median, lower and upper quartiles, and sum.

The SOFA statistical software is compatible with latest technologies and can be installed on several operating systems, such as Microsoft Windows, Ubuntu, ArchLinux, Linux Mint, and MacOS (Leopard upwards).

SOFA Statistics is specifically written in Python computer language, and the widget toolkit used is wxPython. The statistical analyses are based on functions that are obtainable through the Scipy stats module.

**Definition:** SOFA or Statistics Open For All is the user-friendly, open-source statistics, analysis, and reporting package.

**Using SOFA**

SOFA software is used for:

- Making charts, e.g., Pie Charts
- Producing eye-catching and systematic report tables of the data, e.g., Table Income vs Expenditure
- Running basic statistical tests, e.g., one-way ANOVAs

SOFA is excellent software for preliminary research and exploratory analysis also termed as ‘Statistical/Mathematical Doodling’.

When you install SOFA and start the software the following Window will appear.

To analyse the data using the SOFA Statistics, Click on the “Enter/Edit Data” button to get started as shown below.

To open the existing data tables or make new ones click on ‘Open’. To existing tables have extension as ‘tbl’.

---

*NOTES*

**Statistical Analysis Software**

**Self-Instructional Material**

209
You can view the data and can do the data analyses using SOFA Statistics. Note that the background colour in the row or column indicates that the field is read-only. Characteristically, read-only fields are auto numbered or timestamps.

Click on ‘Close’ when completed.

Making a Simple Report Table

On the SOFA main form, click on ‘Report Tables’ and select the variables from the predefined list as shown below.
For example, you can create a simple report table of Age Group vs Weight as shown below.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>&lt; 20</th>
<th>20-29</th>
<th>30-39</th>
<th>40-64</th>
<th>65+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>48.53</td>
<td>67.91</td>
<td>75.00</td>
<td>76.72</td>
<td>83.67</td>
</tr>
<tr>
<td>Median</td>
<td>44.0</td>
<td>65.0</td>
<td>72.0</td>
<td>77.6</td>
<td>83.0</td>
</tr>
<tr>
<td>N</td>
<td>300</td>
<td>185</td>
<td>176</td>
<td>390</td>
<td>436</td>
</tr>
<tr>
<td>Std Dev</td>
<td>22.64</td>
<td>17.35</td>
<td>16.60</td>
<td>10.60</td>
<td>20.25</td>
</tr>
<tr>
<td>Sum</td>
<td>14996.8</td>
<td>12835.4</td>
<td>13206.4</td>
<td>29021.2</td>
<td>38478.0</td>
</tr>
</tbody>
</table>

**ANOVA Test**

Once you have created the statistical table in SOFA you can perform the required statistical test that are available in the software.

To perform statistical test, Click on the ‘Statistics’ button on the main SOFA form as shown below.

Then click on the ‘CONFIGURE TEST’ button after selecting ANOVA from the list.

ANOVA or Analysis of Variance Test is a statistical method which is used to test differences between two or more means in a sample.

Click on ‘Run’ button to view results.
Check Your Progress

1. What is Minitab software and why is it used?
2. What is the purpose of SAS?
3. Define SPSS software.
4. What are the two views of SPSS Data Editor?
5. Why recode is considered a very important feature of SPSS?

13.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Minitab software was developed more than 20 years ago at the Pennsylvania State University. It was originally used by statisticians for teaching statistics. It can be used with considerable ease and effectiveness in all business areas. However, today it is used for multiple applications—especially quality control, six sigma and the design of experiments.

2. System for Statistical Analysis (SAS) has been actively and extensively used in managing, storing and analysing information. It has the advantage of being able to manage really bulky data sets with considerable ease.

3. Statistical Package for Social Sciences (SPSS) is one of the most popular software packages to perform statistical analysis on survey data. Its first version was released in 1968 and since then, it has come a long way. It is used by researchers in educational institutes, research organizations, government, marketing firms, etc.

4. The SPSS Data Editor Window has two views: Data View and Variable View. Variable view is used to define variables that will store the data. However, data view contains the actual data.

5. Recode is a very important feature in SPSS as it is used to convert continuous data into discrete or category data. One can recode values within the existing variable into a new variable.

13.6 SUMMARY

- MS Excel is the simplest and most widely used method of presenting and tabulating data. The basic mathematical functions can be calculated here. The data entered on Excel can be transported to most statistical packages for a higher level analysis.

- Minitab Inc. was developed more than 20 years ago at the Pennsylvania State University. It can be used with considerable ease and effectiveness in all business areas. It was originally used by statisticians. However, today it
Statistical Analysis Software

is used for multiple applications—especially quality control, six sigma and the design of experiments.

- System for Statistical Analysis (SAS) was created in the late 1960s at North Carolina State University. It has been actively and extensively used in managing, storing and analysing information. It has the advantage of being able to manage really bulky data sets with considerable ease. The statistical analyses may be interfaced with the graphical products to produce relevant plots such as q-q plots, residual plots, and other relevant graphical descriptions of the data.
- Amongst the student community as well as with most research agencies, this is the most widely used package. It is adaptable to most business problems and is extremely user friendly.
- Statistical Package for Social Sciences (SPSS) is one of the most popular software packages to perform statistical analysis on survey data. Its first version was released in 1968 and since then, it has come a long way. It is used by researchers in educational institutes, research organizations, government, marketing firms, etc.
- The SPSS Data Editor Window has two views: Data View and Variable View. Variable View is used to define variables that will store the data. Data View contains the actual data.
- A category variable can be a string variable or a numeric variable but it is recommended that categorical variables should be numeric because strings contain letters which cannot be numerically analysed.
- Recode is a very important feature in SPSS, which is used to convert continuous data into discrete or category data. One can recode values within the existing variable into a new variable.
- PSPP is considered as an alternative for the data management and analysis tool SPSS. It is an open source application which is used for statistical analysis of sampled data. The software provides a complete set of capabilities including frequencies, cross-tabs comparison of means (t-tests and one-way ANOVA).
- PSPP software can work on most operating systems and it comes with one unique feature which is its interoperability with other software applications.
- SOFA is an open source user friendly statistical software package considered as a pioneering analysis tool for statistical calculation and reporting. The name SOFA stands for ‘Statistics Open For All’.
- SOFA has a graphical user interface (GUI) and can directly connect to MySQL, PostgreSQL, SQLite, MS Access, Microsoft SQL Server, and CUBRID.
- The SOFA statistical software is compatible with latest technologies and can be installed on several operating systems, such as Microsoft Windows, Ubuntu, ArchLinux, Linux Mint, and MacOS (Leopard upwards).
• SOFA or Statistics Open For All is the user-friendly, open-source statistics, analysis, and reporting package.

13.7 KEY WORDS

• **Statistical Analysis**: It involves collecting and scrutinizing every data sample in a set of items from which samples can be drawn.

• **SPSS (Statistical Package for Social Sciences)**: It refers to a software package used for the analysis of statistical data.

• **Interoperability**: It refers to the ability of computer systems or software to exchange and make use of information.

• **SOFA (Statistical Open For All)**: It refers to the user-friendly, open-source statistics, analysis and reporting software package.

• **ANOVA**: It refers to a statistical technique that is used to check if the means of two or more groups are significantly different from each other.

13.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Write a short note on SAS.

2. Give a brief overview of few features of SPSS.

3. What is PSPP software?

4. Define ANOVA test.

**Long Answer Questions**

1. Discuss the significance of statistical analysis software in today’s businesses.

2. Describe SOFA software and its uses.

13.9 FURTHER READINGS


UNIT 14 REPORT WRITING

Structure
14.0 Introduction
14.1 Objectives
14.2 Structure and Components of Report Writing
14.3 Style Manuals: APA, MLA, Chicago, MS Word and Zotero
14.4 Answers to Check Your Progress Questions
14.5 Summary
14.6 Key Words
14.7 Self Assessment Questions and Exercises
14.8 Further Readings

14.0 INTRODUCTION

On completion of the research study and after obtaining the research results, the real skill of the researcher lies in analysing and interpreting the findings and linking them with the propositions formulated in the form of research hypotheses at the beginning of the study. The statistical or qualitative summary of results would be little more than numbers or conclusions unless one is able to present the documented version of the research endeavour.

One cannot overemphasize the significance of a well-documented and structured research report. Just like all the other steps in the research process, this requires careful and sequential treatment. In this unit, we will be discussing in detail the documentation of the research study. The format and the steps might be moderately adjusted and altered based on the reader’s requirement. Thus, it might be for an academic and theoretical purpose or might need to be clearly spelt and linked with the business manager’s decision dilemma.

14.1 OBJECTIVES

After going through this unit, you will be able to:
- Discuss the structure and components of report writing
- Examine the guidelines for presenting tabular data and visual representations
- Describe the various types of style guides

14.2 STRUCTURE AND COMPONENTS OF REPORT WRITING

Whatever the type of report, the reporting requires a structured format and by and large, the process is standardized. The major difference amongst the types of
The process of report formulation and presentation is presented in Figure 14.1. As can be observed, the preliminary section includes the title page, followed by the letter of authorization, acknowledgements, executive summary and the table of contents. Then comes the background section, which includes the problem statement, introduction, study background, scope and objectives of the study and the review of literature (depends on the purpose). This is followed by the methodology section, which, as stated earlier, is again specific to the technical report. This is followed by the findings section and then come the conclusions. The technical report would have a detailed bibliography at the end.

In the management report, the sequencing of the report might be reversed to suit the needs of the decision-maker, as here the reader needs to review and absorb the findings. Thus, the last section on interpretation of findings would be presented immediately after the study objectives and a short reporting on methodology could be presented in the appendix.
Listed below are some features of a good research study that should be kept in mind while documenting and preparing the report.

- **Clear report mandate:** While writing the research problem statement and study background, the writer needs to be absolutely clear in terms of why and how the problem was formulated.

- **Clearly designed methodology:** Any research study has its unique orientation and scope and thus has a specific and customized research design, sampling and data collection plan. In researches, that are not completely transparent on the set of procedures, one cannot be absolutely confident of the findings and resulting conclusions.

- **Clear representation of findings:** Complete honesty and transparency in stating the treatment of data and editing of missing or contrary data is extremely critical.

- **Representativeness of study finding:** A good research report is also explicit in terms of extent and scope of the results obtained, and in terms of the applicability of findings.

Thus, some guidelines should be kept in mind while writing the report.

- **Command over the medium:** A correct and effective language of communication is critical in putting ideas and objectives in the vernacular of the reader/decision-maker.

- **Phrasing protocol:** There is a debate about whether or not one makes use of personal pronoun while reporting. The use of personal pronoun such as ‘I think…’ or ‘in my opinion…’ lends a subjectivity and personalization of judgement. Thus, the tone of the reporting should be neutral. For example: ‘Given the nature of the forecasted growth and the opinion of the respondents, it is likely that the……’

Whenever the writer is reproducing the verbatim information from another document or comment of an expert or published source, it must be in inverted commas or italics and the author or source should be duly acknowledged. For example:

Sarah Churchman, Head of Diversity, PricewaterhouseCoopers, states ‘At PricewaterhouseCoopers we firmly believe that promoting work–life balance is a ‘business-critical’ issue and not simply the ‘right thing to do’. The writer should avoid long sentences and break up the information in clear chunks, so that the reader can process it with ease.

**Simplicity of approach:** Along with grammatically and structurally correct language, care must be taken to avoid technical jargon as far as possible. In case it is important to use certain terminology, then, definition of these terms can be provided in the glossary of terms at the end of the report.

**Report formatting and presentation:** In terms of paper quality, page margins and font style and size, a professional standard should be maintained. The font style must be uniform throughout the report. The topics, subtopics, headings and
subheadings must be construed in the same manner throughout the report. The researcher can provide data relief and variation by adequately supplementing the text with graphs and figures.

**Guidelines for Presenting Tabular Data**

Most research studies involve some form of numerical data, and even though one can discuss this in text, it is best represented in tabular form. The data can be given in simple summary tables, which only contain limited information and yet, are, essentially critical to the report text.

The mechanics of creating a summary table are very simple and are illustrated below with an example in Table 14.1. The illustration has been labelled with numbers which relate to the relevant section.

**Table identification details:** The table must have a title (1a) and an identification number (1b). The table title should be short and usually would not include any verbs or articles. It only refers to the population or parameter being studied. The title should be briefly yet clearly descriptive of the information provided. The numbering of tables is usually in a series and generally one makes use of Hindu Arabic numbers to identify them.

**Data arrays:** The arrangement of data in a table is usually done in an ascending manner. This could either be in terms of time, as shown in Table 14.1 (column-wise) or according to sectors or categories (row-wise) or locations, e.g., north, south, east, west and central. Sometimes, when the data is voluminous, it is recommended that one goes alphabetically, e.g., country or state data. Sometimes there may be subcategories to the main categories, for example, under the total sales data—a columnwise component of the revenue statement—there could be subcategories of department store, chemists and druggists, mass merchandisers and others. Then these have to be displayed under the sales data head, after giving a tab command as follows:

**Table 14.1** Automobile Domestic Sales Trends

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial Vehicles</td>
<td>112,862</td>
<td>114,122</td>
<td>114,522</td>
<td>107,367</td>
<td>110,912</td>
</tr>
<tr>
<td>Three-wheelers</td>
<td>23,521</td>
<td>24,876</td>
<td>21,078</td>
<td>30,079</td>
<td>33,905</td>
</tr>
<tr>
<td>Two-wheelers</td>
<td>4,102,526</td>
<td>3,304,240</td>
<td>6,208,765</td>
<td>5,223,301</td>
<td>7,072,334</td>
</tr>
<tr>
<td>Grand Total</td>
<td>4,133,633</td>
<td>3,330,466</td>
<td>6,229,536</td>
<td>5,223,301</td>
<td>7,072,334</td>
</tr>
</tbody>
</table>

*Does not include second hand car sales.*

Source: Siam
Total sales
Mass market
Department store
Drug stores
Others (including paan beedi outlets)

Measurement unit: The unit in which the parameter or information is presented should be clearly mentioned.

Spaces, leaders and rulings (SLR): For limited data, the table need not be divided using grid lines or rulings, simple white spaces add to the clarity of information presented and processed. In case the number of parameters are too many, it is advisable to use vertical ruling. Horizontal lines are drawn to separate the headings from the main data, as can be seen in Table 14.1. When there are a number of subheadings as in the sales data example, one may consider using leaders (…….) to assist the eye in reading the data.

Total sales
Mass market ………
Department store ………
Drug stores ………
Others (including paan beedi outlets)………

Assumptions, details and comments: Any clarification or assumption made, or a special definition required to understand the data, or formula used to arrive at a particular figure, e.g., total market sale or total market size can be given after the main tabled data in the form of footnotes.

Data sources: In case the information documented and tabled is secondary in nature, complete reference of the source must be cited after the footnote, if any.

Special mention: In case some figure or information is significant and the reader should pay special attention to it, the number or figure can be bold or can be highlighted to increase focus.

Guidelines for Visual Representations: Graphs
Similar to the summarized and succinct data in the form of tables, the data can also be presented through visual representations in the form of graphs.

Line and curve graphs: Usually, when the objective is to demonstrate trends and some sort of pattern in the data, a line chart is the best option available to the researcher. It is also possible to show patterns of growth of different sectors or industries in the same time period or to compare the change in the studied variable across different organizations or brands in the same industry. Certain points to be kept in mind while formulating line charts include:

- The time units or the causal variable being studied are to be put on the X-axis, or the horizontal axis.
- If the intention is to compare different series on the same chart, the lines should be of different colours or forms (Figure 14.2).
Too many lines are not advisable; an ideal number would be five or less than five lines on the chart.

The researcher also must take care to formulate the zero baseline in the chart as otherwise, the data would seem to be misleading. For example, in Figure 14.3a, in case the zero baseline is (as shown in the chart) the expected change in the number of hearing aids units to be sold over the time period 2002–03 to 2007–08, it can be accurately understood. However, in Figure 14.3b, where the zero is at 1,50,000 units, the rate of growth can be misjudged to be more swift.
Area or stratum charts: Area charts are like the line charts, usually used to demonstrate changes in a pattern over a period of time. What is done is that the change in each of the components is individually shown on the same chart and each of them is stacked one on top of the other. The areas between the various lines indicate the scale or volume of the relevant factors/categories (Figure 14.4).

Pie charts: Another way of demonstrating the area or stratum or sectional representation is through the pie charts. The critical difference between a line and pie chart is that the pie chart cannot show changes over time. It simply shows the
There are certain rules that the researcher should keep in mind while creating pie charts.

- The complete data must be shown as a 100 per cent area of the subject being graphed.
- It is a good idea to have the percentages displayed within or above the pie rather than in the legend as then it is easier to understand the magnitude of the section in comparison to the total. For example, Figure 14.5 shows the brand-wise sales in units for the sample of existing brands of hearing aids in the North Indian market.

![Figure 14.5 Brandwise Sales (units) of Hearing Aids in the North Indian Market (2002–03)](image)

- Showing changes over time is difficult through a pie chart, as stated earlier. However, the change in the components at different time periods could be demonstrated as in Figure 14.6, showing a sample share of the car market in India in 2009 and the expected market composition of 2015.

![Figure 14.6 Sample Structure of the Indian Car Market (2009) and the Forecasted Structure for 2015](image)
Bar charts and histograms: A very useful representation of quantum or magnitude of different objects on the same parameter are bar diagrams. The comparative position of objects becomes very clear. The usual practice is to formulate vertical bars; however, it is possible to use horizontal bars as well if none of the variable is time related [Figure 14.7(a)]. Horizontal bars are especially useful when one is showing both positive and negative patterns on the same graph [Figure 14.7(b)]. These are called bilateral bar charts and are especially useful to highlight the objects or sectors showing a varied pattern on the studied parameter.

![Bar Chart per Day, Unit Sales (thousands) at Fast Food Outlets in Mumbai](image1)

![Bilateral Bar Chart—the Brand Recall and Brand Purchase Response for Pizza Joints in the NCR](image2)

Another variation of the bar chart is the histogram (Figure 14.8) here the bars are vertical and the height of each bar reflects the relative or cumulative frequency of that particular variable.
Check Your Progress
1. What does the preliminary section of the research report include?
2. How should numerical data be represented in the report?
3. What is the difference between line and pie chart?

14.3 STYLE MANUALS: APA, MLA, CHICAGO, MS WORD AND ZOTERO

A style guide (or manual of style) is a set of measures for the composition and plan of reports, either for general use or for a particular distribution, association, or field. (It is regularly called a template, however that term has different implications.) Style guides are common for general and specialized use, for the general reading and writing audience, and for students and scholars of various academic disciplines, medicine, journalism, the law, government, business, and specific industries. House style refers to the internal style manual of a particular publisher or organization.

Bibliography can be defined as ‘a list of books or article, being considered or used by an author for the preparation of particular article’. We can also term the bibliography as reference list depending upon the style manual being chosen.
Style manuals refer to such kind of books which help us to understand the methods through which we can format and record the desired information in a correct way. The usage of these manuals are in much demand due to information available which encompass the location for the bibliography, punctuation rules in general, how to format footnotes, how to give spacing, indentation, form of author’s name and many more.

The choice of the right citation style is important for publishing the article in a correct way. Style manuals are found at:

1. Reference section of academic libraries
2. Bookstores
3. The reference section of some public libraries

For writing a paper and preparing the bibliography, the professor guides a student about the style manual to be used for that particular article. This choice of style guide can depend upon the domain on which the research is being conducted or paper written on the particular subject area. These style manuals are updated frequently, so it becomes necessary to add or modify the content accordingly.

There are various style manuals, but let us focus on some of them in a detailed manner to understand the features being used in them. Some of the major style guides are:

1. APA
2. MLA
3. Chicago
4. MS Word
5. Zotero

I. APA

The American Psychological Association (APA) got involved in journal publishing in 1923. The American Psychological Association style originated in 1929 through the article published in the Psychological Bulletin. In 1944, a 32-page guide appeared as an article in the same journal. It is used by academic writers to cite sources in various disciplines like psychology, education and the social sciences. The first edition of the APA Publication Manual was published in 1952 as a 61-page supplement to the Psychological Bulletin making the beginning of a recognized APA Style. In response to the growing complexities of scientific reporting, subsequent editions were released in 1974, 1983, 1994, 2001, and 2009.

APA Style acts as a tool for formatting the papers and essays written on the above mentioned fields of subject. The guidelines of APA Style were eventually expanded into the APA Publication Manual.

The importance of using APA style in writing the content for psychology and social sciences is that researchers and students can easily express their ideas and
can also explain the experiments by routing the information in the particular format. By writing in the consistent format, readers can quickly find out the content they are looking for in journals or articles.

APA style was developed by social and behavioural scientists to standardize the scientific writing.

APA style is used for
- Terms Papers
- Research Reports
- Empirical Studies
- Literature Reviews
- Theoretical Articles
- Methodological Articles
- Case Studies

APA Style uses five level of headings
i. Level 1 – Centred, Boldface, Uppercase and Lowercase Heading
ii. Level 2 – Flush Left, Boldface, Uppercase and Lowercase Handling
iii. Level 3 – Indented, Boldface, Lowercase paragraph heading ending with a period
iv. Level 4 – Indented, Boldface, Italicized, Lowercase paragraph heading ending with a period
v. Level 5 – Indented, Italicized, Lowercase paragraph heading ending with a period.

The four major sections to be kept in considerations while writing a paper in the APA style are:

1. Title Page: This is the most important part of APA style as it gives a conclusion of the paper like Article title, Author’s name, Author’s school affiliation, Running head and Page Number. Title should be specific and general so that it can easily represent what paper is all about. The APA publication manual suggests that the title should not be more than 12 words. Always remember that while writing the Author’s name it should be by listing first name, middle initial’s and the last name. Abbreviations usage such as Dr. or Ph.D. is strictly avoidable. School affiliation is the location where the particular research is being conducted. A running head should be on the upper left-hand corner on all the pages including the title page. Similarly, all the pages should have a page number on the upper right-hand corner. Title, name, and school should be double-spaced and centred on the page.

2. Abstract: It is an important part of any article or paper, which is being immediately followed by the title. It serves as overall brief description of the contents and the information covered in the paper. Abstract should not be
Report Writing

more than 150 to 250 words depending upon the requirement of the paper. The APA publication manual suggests that the abstract should much function like the title page—which allows the person to quickly determine what the article is all about. The APA manual states that the abstract is the single most important paragraph in the entire paper. As the reader reads the abstract first, then proceed on to the further pages, so the first impression should be interesting for the reader to read the full paper. According to the official guidelines of the American Psychological Association, a good abstract should be short but packed with information, objective and accurate.

3. The Main Body: It is just like an essay which has various sections such as introduction, method, results and discussion sections.

4. References: It includes all the list of sources from where the particular information or content is being taken for the consideration. It one cites any piece of information, anywhere in a paper, it needs to be properly referenced. It should be started on a new page. All the entries should be in alphabetical order. The reference section should be double-spaced. Title of the books, journals, magazines, and newspapers should appear in italics.

General Guidelines for APA Style are as follows:

1. Use 8 ½ x 11” paper.
2. Make 1 inch margins on the top, bottom, and sides.
3. The first word in every paragraph should be intended one half inch.
4. Use Times New Roman font size 12.
5. Double space the entire paper.
6. Include a page header known as the “running head” at the top of every page

The sixth edition of the Publication Manual of the American Psychological Association is the most current. It was released in July 2009 after four years of development. The sixth edition is accompanied by a style website, apastyle.org as well as the APA Style Blog, which answers many common questions from users.

Citation and References

Citations are necessary parts of APA style, as they show readers from where the ideas and research came from. While writing the paper, we need to refer to other publications or journals used for the content or information shared and any of the research conducted. APA style uses an author-date reference citation system in the text with an accompanying reference list. That means that to cite any reference in a paper, the writer should write the author and year of the work, either by putting both in parentheses separated by a comma (parenthetical citation) or by putting the author in the narrative of the sentence and the year in parentheses (narrative citation).
II. MLA

It can be defined as a system for documenting the sources in scholarly writing. It is being used throughout the world by scholars, journal publishers, academic and commercial presses. MLA style was updated in 2016 to overcome the difficulties faced by the researchers in today’s date. This style is being developed by an organization named ‘The Modern Language Association’. It helps in using a uniform way to format papers and assignments which in turns lead to easy reading. It is a documentation style used in scholarship like humanities, especially in English Studies, modern languages, literatures, comparative literature, literature criticism, media studies, cultural studies, and other related disciplines. The MLA Handbook, and the MLA Style both were preceded by a slim booklet titled MLA style sheet, first published in 1951 and revised one in 1970.

General Guidelines for writing in MLA Style

- Use white 8½ x 11” paper
- Make 1 inch margins on the top, bottom and sides
- The first word in every paragraph should be indented one half inch.
- Indent set-off quotations one inch from the left margin.
- Use any type of font that is easy to read
- Font size should be 12.

The following elements should be kept in consideration while writing in MLA Style:

1. Author: Always begin with the author’s last name, followed by a comma and rest of the name, as presented in the work by the ending this element with a period.

2. Title of source: The title of the source should be followed by the author’s name depending upon the type of source list the title in Italics, or quotation marks.

3. Title of container: It is usually italicized and followed by a comma, since the information that follows next describes the container. The eight edition refers to “containers” which are the larger wholes in which the source is located.

4. Other Contributors: Other than the author, any other contributor who has a role in contributing to the source like editors, illustrators, translators, etc., should be given a credit in the documentation.

5. Version: If a source is listed on any edition or version of work, it should be included in the citation.
6. Number: If a source is part of a numbered sequence, such as a multi-volume book, or journal with both volume and issue numbers, those numbers must be listed in the citation.

7. Publisher: The publisher is the one who distributes or produces the book to the public. So, if there is more than one publisher who are relevant to the research work, then their names should be cited by using a forward slash for the separation.

8. Publication date: It can happen that the same source may have been published on more than one date, then in that case, use the one which is most relevant to the research. If not sure about the date, then go with source original publication date.

9. Location: We should always be specific to identify the work location. Page numbers should be enlisted.

**In-text Citations**

It is brief reference within the text which indicates the source used. It should be such which can attribute any ideas, paraphrases, or direct quotations to the source. A reader should follow the source without becoming distracted by any other flow of information.

**III. Chicago Style Manual**

The Chicago Style manual is the guide for American English published since 1906 by the University of Chicago Press. The guidelines for writing and citation styles has been best described in its seventeen editions. It is ‘one of the most widely used and respected style guides in the United States’.

There are two ways for citing a source in the Chicago Style-

1. Notes and Bibliography: It is most commonly used for humanities like literature, history and the arts. In this sources are cited in numbered footnotes or endnotes.

2. Author-date: It is for sciences and social sciences. In this, sources are cited in the form of text like parentheses, by author’s last name and year of publication.

Now let us understand each one of them in detailed.

Notes: Here an endnote or footnote which is being used every time source is being cited. It can be through a direct quote, paraphrase or summary. Footnotes will be at the end of the page on which the source is referred and endnotes will be at the end of each chapter or document. A superscript number corresponding to a
note with the bibliographic information for that source should be placed in the text following the end of the sentence or clause in which the source is referenced.

Bibliographies: It provides an alphabetical list of all sources used in a given work. It is placed at the end of the work preceded by Index. It should include all sources cited within the work and may sometimes include other relevant sources that were not cited but provide further reading.

Although bibliographic entries for various sources may be formatted differently, all included sources (books, articles, websites, etc.) are arranged alphabetically by author’s last name. If no author or editor is listed, the title or, as a last resort, a descriptive phrase may be used.

IV. MS – Word

MS-Word is the new blank document just like a normal template where we can add a text by using the normal style. It can be modified by font name, font size, line spacing, indentation, text alignment and many more.

In Word 2013, style designs are located on the design tab which are a combinations of character, heading and other paragraph styles.

We can apply different set of formatting styles by creating a new style in the style gallery. It is always easier to format with MS-Word style. It is just like a paint which can be applied again and again on the document content.

Advantages of MS-Word Style: -

1. All the headings, sub-headings, paragraphs and tables are in the consistent format.
2. It is always faster to apply a style rather than formatting the entire document.
3. It helps in navigating by clicking on headers and subheads so that content required can easily be located.
4. We can insert a table of contents by using the inbuilt feature of heading and subheading styles.

MS-Word style has two types of document formatting styles:

1. Paragraph Style: It applies to entire paragraph by formatting through alignments, indents and character formatting such font size, colour etc.
2. Character Style: It applies to the smaller ones like one or two more characters.

Different types of headings are

1. Major Headings – It includes important headings like Abstract, Table of Contents, Acknowledgements, References, Chapter headings etc.
2. Subheadings – It can have divided as primary which is first level subheadings and secondary which is second level subheadings etc.

The important points to remember while formatting the headings
1. All the heading size and styles should be written in the same format throughout the content. Also all the levels of subheadings including primary and secondary level should be in the same style.
2. The font size is always larger than the size of the normal text and consistency should be maintained in the entire document.
3. It should be ensured that all the major headings are included in the table of contents.

Key note for the requirements to format the particular chapter/sections
1. Each chapter/sections should start from the new page.
2. Page numbers should be in continuity while starting a new page or section.
3. Each chapter must be with the introductions and conclusions.

V. Zotero Style Manual

Zotero is one of the first Firefox extensions to take benefit of the new function in Firefox 2.0 named mozStorage. The Zotero Style repository is automatically updated within the hour whenever we make changes in our GitHub style repository. It offers a lot of ways to capture, transfer and store item, content or information. It has an important feature that it automatically senses the particular item we are locating through the web. It checks the information through site translators. It can be used with various library catalogues and popular websites such as Amazon, New York Times. We can also add the data manually in Zotero style repository. It creates a copy of the database on every shutdown of the system. Site translators are the discrete pieces of code that are extracted from web-based metadata and then finally inserted into a Zotero item.

Zotero uses the Citation style language to configure citation formatting. CSL is new language for citation formatting and designed in such a simple way which can be easily be used. We can view and edit CSL files by URL chrome://zotero/content/tools/csledit.xul into the Firefox location. Zotero can import bibliographic data stored in a variety of standardized formats used by databases and other reference management tools. The most popular formats are RIS, Bib(La)Tex, and MODS.

How to add a new style to Zotero:
(i) Go to the EndNote styles page and download the style you want.
NOTES
(ii) Open Zotero. Go to Preferences (under Zotero, Edit, or Tools).
(iii) Click the "Cite" button.
(iv) Click the "Styles" tab.
(v) Click the + button at the bottom right.
(vi) Select the style file you saved in step 1.

Check Your Progress
4. What is a style guide?
5. How is the American Psychological Association style originate?

14.4 ANSWERS TO CHECK YOUR PROGRESS 
QUESTIONS
1. The preliminary section includes the title page, followed by the letter of authorization, acknowledgements, executive summary and the table of contents.
2. The numerical data in a research report is best represented in tabular form.
3. One critical difference between line and pie chart is that the pie chart cannot show changes over time.
4. A style guide (or manual of style) is a set of measures for the composition and plan of reports, either for general use or for a particular distribution, association, or field.
5. The American Psychological Association style originated in 1929 through the article published in the *Psychological Bulletin*.

14.5 SUMMARY
- Whatever the type of report, the reporting requires a structured format and by and large, the process is standardized.
- The major difference amongst the types of reports is that all the elements that make a research report would be present only in a detailed technical report in comparison to management report.
- The preliminary section includes the title page, followed by the letter of authorization, acknowledgements, executive summary and the table of contents. Then come the background section, which includes the problem statement, introduction, study background, scope and objectives of the study and the review of literature (depends on the purpose).
The background is followed by the methodology section, which, as stated earlier, is again specific to the technical report. This is followed by the findings section and then come the conclusions.

The technical report would have a detailed bibliography at the end.

There must be no ambiguity in either presenting the findings or representativeness of the findings.

Visual relief for the written can be provided through figures, tables and graphs

Style guides are common for general and specialized use, for the general reading and writing audience, and for students and scholars of various academic disciplines, medicine, journalism, the law, government, business, and specific industries.

There are various style manuals which include APA, MLA, Chicago, MS Word and Zotero.

14.6 KEY WORDS

- **Pie Chart:** It is a type of graph in which a circle is divided into sectors that each represent a proportion of the whole.
- **Style Manual:** It is a manual detailing the house style of a particular publisher, publication, etc.
- **Footnote:** It is a note printed at the bottom of a page that gives extra information about something that has been written on that page.

14.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short-Answer Questions**

1. What is the process of report formulation?

**Long-Answer Questions**

1. Examine the various types of style guides.
2. Discuss some of the guidelines for visual representations in reports.
14.8 FURTHER READINGS


UNIT 15 INFORMETRICS AND WEBOMETRICS

Structure
15.0 Introduction
15.1 Objectives
15.2 Informetrics
15.3 Bibliometrics
15.4 Scientometrics
15.5 Webometrics
15.6 Answers to Check Your Progress Questions
15.7 Summary
15.8 Key Words
15.9 Self Assessment Questions and Exercises
15.10 Further Readings

15.0 INTRODUCTION

Informetrics is the study of the quantitative aspects of information in any form, not just records or bibliographies and any social group, not just scientists. You will also learn about the genesis, scope and definition of informetrics. The major areas of research in the category of information systems are information quality, usability research and retrieval evaluation. The main thrust of retrieval evaluation is the study of a retrieval system’s ability for recall and the precision with which it is done. The key objectives of usability research are tests of accessibility, accuracy, relevance, completeness, objectivity and timeliness of users. This unit will also discuss about the significance of librametry. Bibliometrics is associated with written communication’s publication pattern, which also comprises their authorship. Moreover, you will about the bibliometrics. Finally, you will learn about the basic concept of scientometrics and Webometrics.

15.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the genesis and scope of informetrics
- Define bibliometrics
- Discuss the basic concept of scientometrics and Webometrics
According to Tague – Sutclifle, “Informetrics is the study of the quantitative aspects of information in any form, not just records or bibliographies and any social group, not just scientists.”

The study of quantitative aspects of information is referred to as Informetrics. Informetrics has its scope in the production or creation, dissemination or dispersal, and the use or employment of each and every information form, irrespective of its origin or form. The following fields are the part of informetrics:

- **Scientometrics**: The study of the quantitative aspects of science.
- **Webometrics**: The study of the quantitative aspects of the World Wide Web (WWW).
- **Cybermetrics**: It is similar to Webometrics but it differs in that it has a broader definition which also encompasses electronic resources.
- **Bibliometrics**: The study of the quantitative aspects of the recorded information.
- **Librametrics**: This is associated with the study of the aspects pertaining to library science.

In 1979, the term informetrics (French: Informetrie) was coined by Nacke. The 20th century Infometrics as seen in the western region has its basis largely in Lotka’s law (Alfred J. Lotka), Zipf’s law (George Kingsley Zipf), Bradford’s law (Samuel C. Bradford) and on works of persons, such as Eugene Garfield, Manfred Bonitz, Peter Ingwersen, Olle Persson, Ronald Rousseau, Tibor Braun, Leo Egghe, Gerard Salton and Derek J. de Solla Price.

The one pioneer bibliographic data’s Quantitative analysis was Robert K. Merton in his article titled ‘Science, Technology, and Society’ published in ‘Seventeenth Century England’. It has originally been published in 1938 by Merton.

Usage of ‘Informetrics’ in the manner put forth by the two German Scientists Blackert and Zyge and refers to a recently developed branch of science, which employs the use of statistical and mathematical methods for the investigation of technical as well as scientific information on their practical information activities and their theoretical level.

As has been mentioned above, the first to propose Informetrics was none other than Otto Nacke of Germany in 1979. FID formed a committee on informetrics in the year 1984. The committee had its objectives as the making available of data which was reliable and would support research and development, planning and policy making and also aid in the management of activities, programmes, projects and institutions. In the year 1987, Belfium hosted the first international conference on bibliometrics and in 1989, London hosted the second conference on bibliometrics, scientometrics and informetrics. It was suggested by
B.C. Brookes that the generic term ‘informetrics’ should be employed to imply both the use and the development of various types of measures for the study and analysis of the various properties associated with documents in particular and information in general. In 1991, Bangalore hosted the third conference and at this conference the term informetrics was used. Germany hosted the fourth conference in 1993 and it was here that use of all three terms was made. The fifth conference at Illinois, inclusion has been made of the terms scientometrics and informetrics in the title of the conference itself. So, it would appear that there now is a universal acceptance of the term ‘informetrics’.

The term informetrics is now been employed to impartially encompass scientometrics as well as bibliometrics. While by itself it has no such idea of its own which is in any way distinctively new, yet it implicitly covers both documentary and electronic forms of communication. It does appear to be an important term that could continue into the future. Informetrics, when looked at with the angle of its being a general field of study, is seen to encompass the earlier fields of both scientometrics and bibliometrics.

There has been a huge widening of the scope of informetrics ever since the beginning of the series of conference being held on informetrics. Research and researchers in this field are playing an active role which is amply visible with the publishing of a great number of papers on the subject in several journals on the subject of information science. Figure 15.1 shows a diagrammatic representation of the subject areas and research areas of informetrics.

**Fig. 15.1 Subject and Research Areas of Informetrics**

**Information User and Usage:** A major area of research done as ‘User research’ is done in the following areas:
- Library service’s usage
- Information seeking behavior on Web (specifically search engine usage)
- Humans and their information behavior

**Information System:** The major areas of research in the category of information systems are information quality, usability research and retrieval evaluation. The main thrust of retrieval evaluation is the study of a retrieval system’s ability for
recall and the precision with which it is done. The key objectives of usability research are tests of accessibility, accuracy, relevance, completeness, objectivity and timeliness of users.

Information Itself: Bibliometrics, scientometrics and Webometrics are main research area in information itself.

Another categorization of the main areas of study in the field of informetrics is as follows:

1. Statistical facets of language, word, and phrase frequencies, in natural language text as well as in indexes, for the electronic media as well as the printed media.
2. Characteristics of the productivity of authors measured in respect to the degree of collaboration, number of papers or other means, to name a few.
3. Publication sources’ characteristics, chiefly the distribution of papers in a discipline across journals.
4. Citation analysis which is the distribution over authors, papers, institutions, journals, countries. It is used for evaluating and is a co-citation based mapping of disciplines.
5. Working with recorded information, such as library circulation and use of in-house journal and book and use of databases.
6. Literature obsolescence measured in terms citation as well as use.

The following are two such phenomena which in the past were not looked upon as being part of scientometrics or bibliometrics yet they have comfortably fitted into the scope of informetrics:

1. Definition and measurement of information.
2. Types and characteristics of retrieval performance measures.

The milestones in the history of informetrics are the various regulations that have been formulated and implemented. Some of these are as follows:

- Lotka’s law (1926) pertaining to author productivity studies
- Bradford’s law (1934) pertaining to scattering of journals
- Zipf’s law (1935) pertaining to word occurrence

Since the formulation of the above laws, there has been a huge proliferation of papers associated with them and their refinements. The result of such writings led to the creation of new laws like Sengupta’s law of bibliometric, Garfield’s law of concentration and Price’s square root law of scientific productivity.
Various studied that have been conducted resulted in the creation of models, which can be applied to quantitative studies. One of the most important of these models is Price’s cumulative advantage distribution model. Some other is as follows:

- Brook’s logarithmic model
- Bookstein’s model linking the three basic laws
- Mandelbrot’s model of Zipf’s law
- Vickery’s model of Bradford’s law
- Leimkuhler’s model of journal productivity
- Garfield’s citation analysis
- Impact factor for scientific journal
- Sengupta’s weightage formula
- Narayan’s power law model
- Goffman’s epidemic process
- Groos model connecting Keenan – Atherton data to Bradford’s law
- Kessler’s bibliographic coupling
- Small’s co-citation and clustering
- Ravichandra Rao’s negative binomial model

Informetric analysis is the study of science policy studies and it is mostly undertaken in such nations, such as the Netherlands, UK and USA. Informetric analysis is of great help for the evaluation of activities connected with research and development associate with decision-makers with measures, such as research productivity, impact and quality. Even in other countries, there is a rise in the interest of people in the use of quantitative indicators in the field of both science and technology. Such evaluation proves to be helpful for policy making bodies and for national funding agencies in the process of proper fund allocation. Yet, these studies are not without their fair share of criticism. For example, they have faced criticism in the case of citation studies for scientific publication evaluation. Nevertheless, one can rely upon this method when it is combined with certain other measures for decision-making.

Scientometric and infometrics can be distinguished with their extensive employing of qualitative and quantitative techniques. Recent times have also seen extensive use of such quantitative methods for the purpose of studying the structure of literature. These techniques have proved to be of great help in facilitating of the mapping of disciplines and also in the study of a discipline’s various transitions of its composition and structure. As a matter of fact, over the past decade, various indices, techniques and tools have been formulated and developed with this purpose in mind. They have also been applied very successfully to several and various areas.
M. Bonitz worked with the goal of investigating the existing correlations and relationships that existed in informetrics, scientometrics and bibliometrics on a theoretical level. His work attempted to identify the limits of their competence. According to his theory:

- Bibliometrics is a methodological sub-discipline of library science, including the complex of mathematical and statistical methods, used for analysis of scientific and nonscientific documents and library network.
- Scientometrics is a complex of mathematical and statistical methods, used to investigate aspects, such as research staff, economics, history, science, and to define evolutionary prospects of science.
- Infometrics is a methodological sub-discipline of information science, including the complex of mathematical and statistical methods used to investigate indexing languages, information systems, communication systems, information networks, and so on.

Various authors have put out their own solutions by defining the limits of competence of the above three metrics. Some such scientists are A. Lara (Spanish), I. Marschakova (Russian), D. Schmidmaye (German), Belorussian V. Lazare and Lithuanian D. Voveriene.

In conjunction to the various other fields, Leo Egghe and Ronald Rousseau have identified a place for informetrics as shown in the Figure 15.2.

Fig. 15.2 Relationship of Informetrics with Other Fields
Check Your Progress

1. What do you understand by the term informetrics?
2. Write the key objectives of usability research.

15.3 BIBLIOMETRICS

The statistical analysis of written publications, like articles and books is referred to as bibliometrics. Bibliometric methods are frequently used in the field of library and information science, including scientometrics.

In the recent times, bibliometric analysis has taken on a position of being a well-established component of information research. It is also fast becoming popular as the quantitative approach to the description of documents and examination of services both in practice and in research. The first use of bibliometrics was made in early 1969, and it was during the 1980s that it increasingly gained in popularity. Bibliometrics makes use of bibliometric techniques for the analytical study of the literature. In other words, it is that analysis which is conducted based on the bibliographic phenomena, such as pattern of citation and development of literature.

The term bibliometrics is formed of two Latin and Greek words – ‘biblio’ and ‘metrics’. Therefore, etymologically the words bibliometrics means ‘application of mathematics to the study of bibliography.’ The British Standards Institution define bibliometrics as, ‘…The study of the use of documents and pattern of publication in which mathematical and statistical methods have been applied.’

The field of bibliometrics is a branch of information theory, that quantitatively analyses recorded knowledge to know the behavior and properties of the same. The recorded knowledge that is being analysed could be in any of the various forms, such as dissertations, doctoral theses, journals and books, and so on. For nearly every subject discipline, there has been creation and development by bibliometric research of a body of theoretical knowledge and a group of techniques which has arisen fundamentally from bibliographic data elements. Such studies are employed for identifying the pattern of citations, authorship, publication and/or secondary journals coverage, expecting that these regularities will provide some insight into the dynamics of the field of consideration of information resources that will possibly be employed effectively.

There is now a wide use of and acceptance of ‘bibliometrics’ as a term. Various authors have provided their own interpretation of the term. According to Alan Pritchard, bibliometrics is that which deals with all studies that seek to quantify the processes of written communication. He defines bibliometrics as 'The application of mathematical methods to books and other media of communication.'
According to Fairthorne, bibliometrics is the ‘…Quantitative treatment of properties of recorded discourse and behavior appertaining to it.’

In the year 1977, during his bibliometric study conducted online, Hawkins made his own interpretation of the term. According to him, bibliometrics is the ‘Quantitative analysis of the bibliographic features of a body of literature.’

In recent times, Potter put out his own interpretation of bibliometrics and described it as ‘…The study and measurement of the publication patterns of all forms of written communication and their authorship.’

A historic overview of bibliometrics’s various definitions was presented by Broadus. He also put out his own definition as an alternative, defined as, ‘Bibliometrics is the quantitative study of physical published units or of bibliographic or of surrogates of either.’

Broadly speaking, there are two basic groups into which all aspects of bibliometric studies are divided. While one group deals with describing the characteristics or features of literature (descriptive studies), the other deals with examining those relationships that get created amongst the components of a literature (behavioral studies, sometimes referred to as citation studies but not restricted to them). Possibly, it is the easiest to understand descriptive studies and can be looked upon as being in various manners similar to population censuses.

While there are two groups of bibliometric studies, they both have to make available accurate as well as comprehensive data pertaining to the rapidly changing population, while one will do it with respect to literature, the other will do it with respect to people. Literature descriptions get put together with the help of the bibliographic characteristics of each individual member of the literature. Nevertheless, there are some specific studies connected with bibliographic description that are pertinent to bibliometric study which are the ones that make such data available which is for the character or condition of the literature as a whole. Thus, a vital difference is present between bibliometric and bibliographic data.

When considered as a technique, bibliometry has extensive applications for identification of research trends in a subject, trends in authorship and collaboration in research. Furthermore, it applies to core journals, obsolescence, and dispersion of scientific literature useful in estimating the comprehensiveness of secondary journals studying the author productivity, characteristics of subject literature including structure of knowledge, historical and sociological aspects of science and helpful in formulation of need based collection development policy, weeding and stacking policy, science policy studies and many others.

15.4 SCIENTOMETRICS

Scientometric research is the quantitative mathematical study of science and technology. It comprises economic as well as bibliometric analysis. In the recent
times, there has been expansion in this field as a pace that can only be called enormous. All of the leading journals of the field carry ‘research covering article output, citation relationships between disciplines and linked geographical analysis.

Bibliometric research started in the 19th century and it originated possibly in the field of law. According to Shapiro (1999) the aspects associated with bibliometrics were practiced in the legal field long before being introduced into scientific literature. Researches during the 1880s had been reported by Delmas (1992), when he noted and elaborated on the documentation being conducted in France. Nevertheless, it has been seen that it was the field of psychology where the pioneering studies associated with quantitative and qualitative analysis of science originated (Godin 2006). The work done by Buchner is cited by Godin while describing the concept of ‘scientific psychology to be factual, inductive, measurable and experimental’. In the year 1920, research was presented by Boring pertaining to geographical and subject analysis of psychologists.

The research that is looked upon as being the earliest and is the, most definable one in the field of scientometric and is the one from which the laws of bibliometrics were derived. Of these laws, the first law was the one which is named after Alfred Lotka and is referred to as the Lotka’s Law, and its origins go as far back as 1926. According to Lotka’s Law, in any defined area, when it is studied for a specific period, it is seen that a small number of authors are responsible for a large percentage of publications associated with that area. Then, in the year 1935 the work done by George Kingsley Zipf, which was done with keeping in mind the frequency of occurrence of words in a text and that came to be known as Zipf’s Law. Over time, the research done by Zipf got further refined and two specific laws emerged from it which pertained to the occurrence of low frequency and the high frequency words in a text. During the year 1948, the analysis by Samuel Clement Bradford arrived at the fact that for a specific time for a specific area a small number of journals are seen to publish a high percent of articles for that specific area while at the same time there are many journals in which just a few articles pertaining to that area are published. This is known as Bradford’s Law.

Even in the current times, these are studied at all times and are looked upon as being the very base on which the scientometric literature of these modern rests.

Lehman, in the year 1944, had explained the relationship that existed between quality and quantity in the sphere of scientific writing. Later, in the year 1952, Dennis made an analysis of the effect that the age of a scientist would have on the quantity and quality of writing. Even now such analyses are discussed and described in literature of present times, and are responsible for steering the process of thinking in the direction of averaged metrics which lies in the scope of bibliometrics.

15.5 WEBOMETRICS

One can define Webometrics, a component of information science as ‘the study of the quantitative aspects of the construction and use of information...
resources, structures and technologies on the Web drawing on bibliometric and informetric approaches’. In more generic term, it is ‘the study of Web-based content with primarily quantitative methods for social science research goals using techniques that are not specific to one field of study’. Though in the first definition, the stress is on the informetric heritage of several of the bibliometric methods, in the second definition, the stress is on the value that Webometrics could provide to the wider social sciences, displaying a move to more applied studies and away from theoretical studies, by keeping alive the focus on methods development. Currently, in Webometrics there are various software and various methods to implement different types of quantitative analyses of the Web. In spite of the concerns that had arisen initially in regard to the idea that it was easy to manipulate Web data since such data had no quality-control, it is claimed by Webometrics advocates that it proves to be useful for both studies of the various aspects like hyperlink between the different Websites geared to being academic Websites, and also to studies of offline phenomena which could be reflected online, like blogs reflecting political attitudes.

In 1997, Tomas Almind and Peter Ingwersen coined the term Webometrics on the basis of the concept that it is possible to apply informetric analyses to the Web. There was a huge spur in the growth of this field when the Web Impact Factor (WIF) was introduced. WIF is a metric used for assessing the impact that a specific Website or other area of the Web will have on the basis of the number of hyperlinks that point to it. There seems to be a lot of sense in WIFs and this is so because such areas on the Web which prove to be more important or useful will certainly attract more hyperlinks as compared to such areas that are of average value. This metric’s logic has its basis in the significance of the citations in journal impact factors, yet there is that advantage that WIFs had of being calculated by making use of the new advanced search queries by the commercial search engine AltaVista. Over time, Webometrics has risen to acquire the position of being a vast, coherent field that is a subset of information science from the perspective of bibliometry. It combines in itself various quantitative techniques which are Web based, such as Web citation analysis and link analysis. Furthermore, the usefulness of Webometrics increased in several applied contexts, like in creation of world Webometrics ranking of universities and on the scientometric investigations or evaluations of areas of research or bodies of research.

Link analysis: impact measurement and network as it is the early research of Webometrics got conducted mainly via combination of the growth of both better applications and methods for a wide range of varying contexts. This gave rise to specific types of studies which were link network analyses and link impact analyses. Mostly, the link impact studies emphasizing in comparing how many hyperlinks, in terms of numbers, point to each Website within a pre-defined set. For example, all of the departments of a specific discipline within a country and all of the universities of a specific country. Links that point to the Websites of the university and possibly in a few cases to the Websites of departments were realized to
Informetrics and Webometrics significantly correlate with the research productivity measures or prestige. Thereby, providing evidence regarding how valid it was to employ the link impact metrics as an indicator for research-related activities. Link impact metrics are used in such a manner to make available the clue to the most important Websites or organizations in any given group. Furthermore, the breakdown in the sources of links which have been used for the calculations show the sources of the impact, like the organization type or the country which host the majority of the links.

With link network research, there has been the generation of network diagrams based upon the links in a specific Website group for the purpose of connectivity pattern identification. Besides working with such networks that are based upon a direct links between site pairs, use has also been made of co-inlinks to show the connection which exists between site pairs. To understand what a co-inlink to a pair of Websites is, consider Websites A and B and if there is a third Website, let it be called C, which maintains a hyperlink to A and also to B, then C is the co-inlink to the Websites A and B. Such a relation can be considered to draw a parallel with the concept of co-citation in the field of bibliometrics. Co-inlink is particularly used in the investigation of Websites that might be similar yet are not necessarily having hyperlink with each other. In a network diagram made for direct links, there is the possibility of the exclusion of links that are there between pairs of sites that though in some way similar are not related with one another directly.

For link analysis, it has already been said that there exist two components that are of major importance and these are the methods and the software which are used for extract data pertaining to links. For several years, researchers managed to collect information regarding hyperlinks with the help of search engines that were commercial, such as Yahoo!, AltaVista and Bing with the help of their associated advanced link search commands. However, all of these tools finally were withdrawn. Data on links is still obtainable through making use of Web crawlers designed for specialist link analysis, including free programs, such as IssueCrawler and SocSciBot. In the field of information science, network diagrams that are hyperlink-based are employed for the purpose of investigating what interconnections exist between large groups of organizations, like the specific knowledge sector organizations and universities within Europe.

There are a number of link analysis researches whose main focus is on the links themselves and with respect to the links they study the reason for their creation, and the reason why some pages and sites of the sites pull in more links than other pages and sites. While these studies have focused on just links, they have also focused on those links in the contexts of academics. Some of the content analyses that have been conducted have revealed that the links that exist between Websites are academic are mostly created for purposes educational or scholarly. This is partially the same reason as given for citation analysis. Researchers have even conducted statistical tests for the purpose of studying which of the attributes possessed by an owner of a Website (besides the already known research
productivity or production attribute) happen to associate with higher inlink counts. For example, one of the findings is that the gender of the owner of a research Website is inconsequential. Of late, a very comprehensive study was conducted which employed an approach for statistical modeling which was extremely advanced and used it on a massive dataset to obtain such insights which was very significant in view of the factors that played a role in academic Website interlinking within Europe. This study found that it was not country, region, domain specialization and that held the most importance in predicting hyperlinks, but rather reputation in the case of top universities.

From Web Citation Analysis to Altmetrics

Another form of Webometrics that gained a lot of popularity was web citation analysis. Web citation analysis refers to the counting of the online citations to published academic documents, such as refereed journal articles. The rationale given for conducting of early research was making an assessment of whether or not the Web would be able to replace traditional citation databases for assessing the impact that was felt of articles in open access online journals and following that even in the case of all journals. Through this research, it was revealed that even while the counts of Web citations were in correlation with citation counts found in the traditional databases. Several of the Web citations were taken from non-academic sources, like online library catalogues. In this manner, it was gleaned that the Web was a lower or inferior source of citation impact evidence for journals or individual journal articles.

Then the investigations that follow were much more specialized and geared towards specific kinds of Web citations to academic publications, such as citations from Google Books, online syllabi and PowerPoint presentations. These investigations were based on the restricted domains, such as the Web-based citation counts, which show the various different kinds of impact that would differ from the scholarly impact observed in the traditional citation counts. For example, the citations of online syllabus would reveal the education value or impact of articles. After a while, even this line of search handed over the reins and this time it was to the altmetrics initiative.

The analysis of mentione phrases or keywords is yet another relatively less studied yet promising type of Webometrics. It does not necessarily deal with citations. This form of analysis began when investigation was being done on the context of online mentions of academics. Research into keywords is also employed for the purpose of mapping concepts online as also interactions between concepts online by tracking similar words in Web pages.

Theoretical Perspectives and Information-Centered Research

Webometrics was and has remained to be a field that is method-centered, and it continues to create more and more methods for the purpose of gathering and analysing data gathered from the Web. Probably due to such focus, majority of
the Webometric studies’ theoretical component typically was derived from citation analysis instead of Web data. For example, several of the studies conducted earlier made an assessment of whether or not Web citation or Web link counts has any correlation with traditional citation counts. In this way, they were based upon Robert Merton’s theoretical discussion of citation norms in science. Therefore, to an extent such studies were able to find the extent of Web data built-in Merton’s theory. The area of link analysis is the most developed area of Webometrics and has no special theories to show that the Web is a space which is much more complex and varied than academic journal databases, with theory development in the latter being recognized as problematic and controversial.

One reason that is given for there being paucity of native theory for the field of Webometrics is that research in this field is information-centered, since it is specifically apt for this area. Any research study that is information-centered will have as its key focus some new information source. For example, some form of Web data, and the study will try to detect the social science research problems for which specific data is best placed for addressing instead of employing a priori intuitions and matching the data with a research problem and then to assess the value of the data for the problem. Specifically this particular theory was employed for the purpose of justifying the growth of a whole range of varying methods for analysing Web data and for matching the methods with a varied range of problems related with areas of social science.

**Web Data Analysis as a Service for the Social Sciences**

There is an increase of the Webometrics research as its moving from general to the area of academic Web analyses for the investigations of social Websites. It can be achieved by downloading and obtaining data automatically from those Websites with the help of a Web crawler or data requests sent via routes that are permitted.

For example, RSS feeds and blogs have been analysed in same manner for the purpose of detecting the fears of public with respect to science and there has been investigation of social network sites to find out patterns for detecting friendship and language use. Analysis has been done using Twitter to know the public sentiments caused as a reaction to major media events. There has been an analysis of YouTube for such factors that are linked to discussions which are associated with online videos. The research method for all cases was Webometrics – data gathering and analysis on a massive scale for the purpose of social science – with the research findings being used for disciplines other than information science, like-scientific communication, and politics and media studies.

---

**Check Your Progress**

3. What are bibliometrics?
4. Define the term scientometrics.
5. What do you understand by the term Webometrics?
15.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

NOTES

1. Informetrics is the study of the quantitative aspects of information in any form, not just records or bibliographies and any social group, not just scientists.
2. The key objectives of usability research are tests of accessibility, accuracy, relevance, completeness, objectivity and timeliness of users.
3. Bibliometrics is the quantitative study of physical published units or of bibliographic or of surrogates of either.
4. Scientometrics is defined as a sub-field study of science as an information process. When the information model is considered, all of the publications will be looked upon as being communication promoters and information carriers.
5. Webometrics, a component of information science is defined as ‘the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web drawing on bibliometric and informetric approaches’. In more generic term, it is ‘the study of Web-based content with primarily quantitative methods for social science research goals using techniques that are not specific to one field of study’.

15.7 SUMMARY

- Informetrics is the study of the quantitative aspects of information in any form, not just records or bibliographies and any social group, not just scientists.
- Informetrics has its scope in the production or creation, dissemination or dispersal, and the use or employment of each and every information form, irrespective of its origin or form.
- The statistical analysis of written publications, like articles and books is referred to as bibliometrics. Bibliometric methods are frequently used in the field of library and information science, including scientometrics.
- The term bibliometrics is formed of two Latin and Greek words – ‘biblio’ and ‘metrics’. Therefore, etymologically the words bibliometrics means ‘application of mathematics to the study of bibliography.
- Bibliometric study is considered to be an important tool for checking the effectiveness of a system. It is also considered to be a distinctive research technique which helps in measuring science on the basis of citation data.
- Scientometric research is the quantitative mathematical study of science and technology. It comprises economic as well as bibliometric analysis.
The key purpose of scientometrics is to determine the prospect and the state of a subject and how it will develop in the future. Various indicators of scientometry are employed to determine all of this.

Webometrics, a component of information science, is defined as ‘the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the Web drawing on bibliometric and informetric approaches’.

15.8 KEY WORDS

- **Informetrics**: It is the study of the quantitative aspects of information in any form, not just records or bibliographies and any social group, not just scientists.
- **Scientometrics**: It is the study of the quantitative aspects of science.
- **Webometrics**: It is the study of the quantitative aspects of the World Wide Web (WWW).
- **Bibliometrics**: It is the statistical analysis of written publications, like articles and books.

15.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Write a short note on the genesis of informetrics.
2. What are the two basic groups in which all aspects of bibliometric studies are divided?

**Long Answer Questions**

1. Discuss the main areas of study in the field of informetrics.
2. Discuss web citation analysis.

15.10 FURTHER READINGS


Carpenter, R.L. and Vasu, E.S. 1979 *Statistical Methods for Librarian*. Chicago: ALA.

