CENTRE FOR DISTANCE & ONLINE EDUCATION

UNIVERSITY OF JAMMU JAMMU

SELF LEARNING MATERIAL OF

BUSINESS RESEARCH METHODS

FOR M.COM SEM III

For the Year 2023,2024,2025

COURSE NO. MCOMC351

UNIT: I - IV
LESSON: 1 - 20

Course Coordinator
Prof. Sandeep Kour Tandon
CDOE, University of Jammu, Jammu

Teacher In charge
Dr. Deepti Abrol
Assistant Professor, CDOE

Printed and Published on behalf of the Centre for Distance & Online Education, University of Jammu, Jammu by the Director, CDOE, University of Jammu, Jammu.
http:/www.centrefordistanceandonlineeducationju.in

COURSE NO. M.COMC351

Written By:

I & II Unit
Mr.Gourav Arora
Lecturer,Computer Software
Application at National Skill

Training Institute,Jammu

III & IV Unit
Dr.Ritika Sambyal
Lecturer,Commerce
Udhampur Campus
University of Jammu,Jammu

Editing and
Proof reading by:

Prof.Sandeep Kour Tandon
Co-ordinator,M.Com
Room No.111,Ist Floor
DD & OE,University of
Jammu

Ms.Shriya Gupta
Teacher Incharge,M.Com
Room No.205,IInd Floor
DD&OE ,University of
Jammu

Ms.Avantika Bakshi
Room No.205
IInd Floor,DD&OE
University of Jammu

Printed By :- Jandiyal Printing Press / 2023 / 650

CENTRE FOR DISTANCE & ONLINE EDUCATION
UNIVERSITY OF JAMMU
SYLLABUS
M.COM. THIRD SEMESTER (NON CBCS)
BUSINESS RESEARCH METHODS
(Core Course)

Course: MCOMC351 **Max Marks: 100 Marks**

Credit: 4 **External: 80 Marks**

Time: 3.00 Hrs **Internal: 20 Marks**

(Syllabus for the examination to be held in December 2023, 2024, 2025)

COURSE OBJECTIVES

To understand the nature and process of business research.

To explain the types of research design and methods of sampling.

To discuss the tools of data analysis and errors in interpretation.

To discuss the structure of research report and research ethics.

COURSE OUTCOMES

After the completion of this course, the student will be able to:

acquaint students with the basics of business research.

understand the different research designs and need of sampling in business research.

application of tools of data analysis and interpretation.

prepare the detailed research report by following research ethics and plagiarism.

apply the statistical techniques for analysing the data of various organisations

UNIT I INTRODUCTION TO BUSINESS RESEARCH

Concept, characteristics, role and process of business research; Problem identification and usefulness of objectives and hypothesis in business research; Data collection methods for business research - Observation, interviews, questionnaire and schedules; Designing of questionnaire; Choice between primary and secondary data; Scaling techniques in business research; Precautions in business research.

UNIT II RESEARCH DESIGN

Concept, nature and classification of research designs – Descriptive, exploratory, case study method, quantitative vs qualitative research, cross-sectional and longitudinal studies, errors affecting research design; relationship among descriptive, exploratory and causal research, advantages and disadvantages of research design, concept and relevance of sampling design in business research.

UNIT IV RESEARCH REPORT WRITING

Processing of data: Feeding and coding of data; Reliability and validity of data Overview of Univariate; Bivariate and Multivariate data analysis through SPSS, Concept and applications of exploratory factory analysis, confirmatory factor analysis and SEM in business research, Testing of hypothesis; Concept, importance and pre-requisites of data interpretation; Errors in data interpretation.

UNIT IV RESEARCH REPORT WRITING

Concept and types of research reports; Essentials of a good research report; Structure of the research report: Preliminary section, Main report, Managerial implications; References and bibliography and annexure; Effectiveness of research report; Research ethics: Meaning and relevance of ethics in business research; Researchers' ethical code; Need for plagiarism; Publication and presentation of business research.

SUGGESTIVE READINGS

Cooper, D. R., Schindler, P. S. *Business Research Methods*. Tata Mc Graw Hill, New Delhi.

Sekaran U., Bougie R. *Research Methods for Business: A Skill Building Approach*. Wiley, India.

Srivastava, T. N., Rego, *S. Business Research Methodology*. Tata Mc Graw Hill, New Delhi.

William G. Z. *Business Research Methods*. Thomson, India.

Harper, W. B., Ralph W., Stanley F. *S. Marketing Research: Text and Cases,* Homewood, Irwin.

Note: Latest edition of the readings may be used.

NOTE FOR PAPER SETTING

The paper consists of two sections. Each section will cover the whole of the syllabus without repeating any question in the entire paper.

Section A: It will consist of eight short answer type questions, selecting two from each unit. A candidate has to attempt any six and answer within 200 words. Each question carries 4 marks, total weightage to this section shall be 24 marks.

Section B: It will consist of six essay type questions. The answer of each question should be within 800 words. One questions shall be set atleast from each unit and the candidate has to attempt four. Each question will carry 14 marks and total weightage shall be 56 marks.

MODELQUESTION PAPER

BUSINESS RESEARCH METHODS

Time : 3hrs **M. Marks : 80**

SECTION-A

Note:- Attempt any six questions. Each question carries 4 marks. Answer to each question should be within 200 words.

Explain the meaning and role of business research.

Highlight the steps involved in defining the business research problem.

What are different type of interviews?

What is the difference between qualitative and quantitative research?

What is univariate, bivariate and multivariate data analysis?

What are applications of exploratory Factor Analysis in business research?

What are essentials of a good research report?

What are different types of research reports?

SECTION-B

Note:- Attempt any four questions. Each question carries 14 marks. Answer to each question should be within 800 words.

What are the different methods of data collection? Briefly explain the advantages and disadvantages of each?

Define the steps involved in the research design proun. Explain the different types of research designs.

What are the two types of sampling design? Explain its merits & demerits?

Disuss with help of suitable examples the steps involved in data processing.

What is hypothesis testing? Explain the procedure for hypothesis testing?

What are components of effective research report? What are the elements of researchers ethical code?

CONTENT

CONTENT

INTRODUCTION TO BUSINESS RESEARCH

M.COM III SEM **UNIT- I**

MCOMC351 **LESSON - 1**

CONCEPT, CHARACTERISTICS, ROLE AND PROCESS OF BUSINESS RESEARCH

STRUCTURE

1.1 INTRODUCTION

When managers use research, they are applying the methods of science to the art of management. All business undertakings operate in the world of uncertainty. There is no unique method that can eliminate the uncertainty. But business research methodology, more than any other procedure, can minimize the degree of uncertainty. Thus, it reduces the probability of making a wrong choice among alternative courses of action. This is particularly significant in the light of increasing competition and growing size which make the task of choosing the best course of action difficult for any business enterprise. In this unit, we will discuss at length the importance of research in decision-making by delineating all its relevant elements.

1.2 OBJECTIVES

After going through this lesson, you should be able to describe:-

The meaning of research in the context of making intelligent decisions.

The characteristics and role of research.

The steps involved in the process of research.

1.3 MEANING OF BUSINESS RESEARCH

Research is a purposeful investigation. It provides a structure for decision-making. There are three parts involved in any investigation: (i) the implicit question posed (ii) the explicit answer proposed (iii) collection, analysis, and interpretation of the information leading from the question to the answer. This third part is the defense that justifies the recommendation and is viewed as research.The word research is comprised of two syllables: ***re*** and ***search***. The dictionary defines the meaning of 're' as again, a new or over again and the 'search' as to examine closely and carefully, to test and try, or to probe. Together they form a word that describes a careful,systematic, patient study and investigation in some field of knowledge, undertaken to establish facts or principles.

1.4 CONCEPT OF BUSINESS RESEARCH

According to **Robert Ross**, "research is essentially an investigation, a recording and an analysis of evidence to gain knowledge." The word research identifies a process by which the organization attempts to supply the information required for making sound management decisions. Research is not synonymous with common sense. The difference revolves around words such as "systematic," "objective," and "reproducible". Both research and common sense depend on information; the distinction between them lies in the procedures and methods adopted by which the information is obtained and used in arriving at conclusions. Research cannot address itself to the complete information on a particular subject. Hence two secondary characteristics of research specify **"relevance"**, and **"control."**

A systematic approach is an essential approach for good research. Each step must be so planned that it leads to the next step. It is usually very difficult to go back and correct the mistakes of the previous step; sometimes it is impossible. Even when it is possible, it will involve a loss of time and money. The authors have divided the research into several steps. Both the number of steps and the names are somewhat arbitrary, however, the recognition of a sequence is crucial. Planning and organization are part of this systematic approach with a lot of emphases given to the interdependence of the various steps.

While planning, one of the very common mistakes that are committed is the separation of data collection and data analysis. First, we collect the data; then, we decide what analysis is appropriate. This approach invites a disaster. In one of the research projects, depth interviews of fresh college students were carried out at a very high cost and the necessary data were all collected. The data were still unanalyzed because no one knew how to proceed. Our point is that considerable thought should have been given at the planning stage itself as to what kind of analysis will be required for the project which will satisfy the needs of the decision-maker.

Objectivity warrants an approach that is independent of the researcher's personal views and opinions about the answer to the problem under investigation. It is possible to have honest differences for the proper definition or collection procedure, but the one selected must not be chosen to verify a prior position.

A reproducible research procedure is one that an equally competent researcher could duplicate and from it obtain approximately the same results. To achieve reproducibility, all procedures must be stated unambiguously. The precise wording of questions, method of sampling, collection method, interviewer instructions, and all other details must be clearly stated. Even if the environment changes, the research is at least "conceptually" reproducible in the sense that the steps could be mentally duplicated. The interviewer should avoid the temptation of rephrasing the question for the respondent to preserve the reproducibility aspect. Poor and vague sampling procedures can also lead to non-reproducibility. If procedures are vague and not stated clearly, you cannot expect consistency even from the same interviewers.

Relevancy accomplishes two important tasks. First, it avoids the collection of unnecessary information along with the accompanying cost. In the second place, it forces the comparison of the data collected with the decision maker's criteria for action. Before the start of the research project, you should ask the question "what action would you take if the research answer were?" This approach enables both the investigator and the decision-maker to know whether the project is in the right direction.

The control aspect is particularly elusive in research. We must be aware that the results of our study are due to the presence of some factor other than those we are investigating. It is impossible to have control over all other factors; the best we can do is to have control over those we think are most likely to cause us difficulty. Suppose we study the relationship between shopping behavior and income without controlling for education and age, it will be the height of folly since our findings may reflect the effect of education or age rather than income. The control must consider two aspects.

Those variables that are true within your control must be varied according to the nature of your investigation. (2) Those variables beyond your control should be recorded.

1.5 CHARACTERISTICS OF RESEARCH

Good business research generates dependable data that are derived from professionally conducted practices and that can be used reliably for business decision-making. But, poor research is carelessly planned and conducted, resulting in data that a manager cannot use to reduce his or her decision-making risks. Good business research follows the standards of the scientific method with the following characteristics:

Purpose clearly defined: The purpose of the business research, the problem involved or the decision to be made should be clearly defined in terms as unambiguous as possible. The statement of the decision problem should include its scope, its limitations, and the precise meanings of all words and terms significant to the research. Failure of the researcher to do this adequately may raise legitimate doubts in the minds of research report readers as to whether the researcher has sufficient understanding of the problem to make a sound proposal attacking it.

Detailed research process : The research procedures used should be described in sufficient detail to permit another researcher to repeat the research. This includes the steps to acquire participants, informed consent, sampling methods and representativeness, and data gathering procedures. The omission of significant procedural details makes it difficult or impossible to estimate the validity and reliability of the data and justifiably weakens the confidence of the reader in the research itself as well as any recommendations based on the research.

Research design thoroughly planned: The procedural design of the research, and its choice among competing designs, should be clearly described and carefully planned to yield results that are as objective as possible. A survey of opinions or recollections ought not to be used when more reliable evidence is available from documentary sources or by direct observation. Bibliographic searches should be as thorough and complete as possible. Experiments should have satisfactory controls, reducing threats to internal validity and enhancing the probability of external validity (generalizability). Direct observations should be recorded as soon as possible after the event. Efforts should be made to minimize the influence of personal bias in selecting and recording data.

High ethical standards applied: Researchers often work independently and have significant latitude in designing and executing projects. A research design that includes safeguards against causing mental or physical harm to participants and makes data integrity a priority should be highly valued. Ethical issues in research reflect important moral concerns about the practice of responsible behavior in society. Careful consideration must be given to those research situations in which there is a possibility for physical or psychological harm, exploitation, invasion of privacy, and/or loss of dignity. The research need must be weighed against the potential for these adverse effects. Typically, you can redesign a study, but sometimes you cannot. The researcher should be prepared for this dilemma.

Limitations including: The researcher should report, with complete frankness, flaws in procedural design and estimate their effect on the findings. There are very few perfect research designs. Some of the imperfections may have little effect on the validity and reliability of the data; others may invalidate them entirely. A competent researcher should be sensitive to the effects of imperfect design. The researcher's experience in analyzing data should provide a basis for estimating the influence of design flaws. As a decision-maker, you should question the value of research about which no limitations are reported.

Adequate analysis for decision makers needs: Analysis of the data should be extensive enough to reveal its significance, what managers call insights. The methods of analysis used should be appropriate. The extent to which this criterion is met is frequently a good measure of the competence of the researcher. Adequate analysis of the data is the most difficult phase of research for the novice. The validity and reliability of data should be checked carefully. The data should be classified in ways that assist the researcher in reaching pertinent conclusions and reveal the findings that have led to those conclusions. When statistical methods are used, appropriate descriptive and inferential techniques should be chosen, the probability of error should be estimated, and the criteria of statistical significance applied.

Findings presented unambiguously: Some evidence of the competence and integrity of the researcher may be found in the report itself. For example, language that is restrained, clear, and precise; assertions that are carefully drawn and hedged with appropriate reservations; and an apparent effort to achieve maximum objectivity tend to leave a favorable impression of the researcher with the decision-maker. The presentation of data should be comprehensive, reasonably interpreted, easily understood by the decision-maker, and organized so that the decision-maker can readily locate critical findings.

Conclusions justified: Conclusions should be limited to those for which the data provide an adequate basis. Researchers are often tempted to broaden the basis of induction by including personal experiences and their interpretations of data not subject to the controls under which the research was conducted. Equally undesirable is the all-too-frequent practice of drawing conclusions from a study of a limited population and applying them universally. Good researchers always specify the conditions under which their conclusions seem to be valid.

Researcher's experience reflected: Greater confidence in the research is warranted if the researcher is experienced, has a good reputation in research, and is a person of integrity. Were it possible for the reader of a research report to obtain sufficient information about the researcher, this criterion perhaps would be one of the best bases for judging the degree of confidence a piece of research warrants and the value of any decision based upon it. The research report should contain information about the qualifications of the researcher.

1.6 ROLE OF BUSINESS RESEARCH

Through research, an executive can quickly get a synopsis of the current scenario which improves his information base for making sound decisions affecting future operations of the enterprise. The following are the major areas in which research plays a key role in making effective decisions :

Marketing: Marketing research has become very crucial in making sound marketing decisions. Marketing research involves the process of systematic collection, compilation, analysis, and interpretation of relevant data for marketing decisions. Research tools are applied effectively for studies involving demand forecasting, consumer buying behavior, measuring advertising effectiveness, media selection, test marketing, product positioning, and new product potential.

Production: Research enables an organization to decide on what to produce, how much to produce when to produce, and for whom to produce in the field of production. Research tools are also of immense help in quality control and setting up optimum inventory levels.

Banking: Banking institutions have found it useful to setup research departments to gather and analyze information both for their internal operations and for making in-depth studies on the economic conditions of business. Reserve Bank of India has setup an excellent research department for planning and management reporting.

Materials: The materials department uses research to frame suitable policies regarding where to buy, how much to buy, when to buy, and at what price to buy.

Human resource development: The human resource development

department uses research to study wage fates, incentive schemes, cost of living, employee turnover rates, employment trends, and performance appraisal. It also uses research effectively for its most important activity namely manpower planning.

Government: Research lays the foundation for all government policies in our economic system. For example, research is applied to evolving the union finance budget and railway budget every year. Research is used for economic planning and optimum utilization of resources for the development of the nation. Research is also needed for the systematic collection of information on the economic and social structure of the nation. Such information indicates what is happening to the economy and what changes are taking place.

1.7 PROCESS OF BUSINESS RESEARCH

Before embarking on the details of the research methodology and techniques, it seems appropriate to present a brief overview of the research process. The research process consists of a series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps.One should remember that the various steps involved in a research process are neither mutually exclusivenor separate and distinct. They do not necessarily follow each other in any specific order and the researcher has to be constantly anticipating at each step in the research process the requirements of the subsequent steps. However, the following order concerning various steps provides a useful procedural guideline regarding the research process: (1) formulating the research problem; (2) extensive literature survey; (3) developing the hypothesis; (4) preparing the research design; (5) determining sample design; (6) collecting the data;

execution of the project; (8) analysis of data; (9) hypothesis testing; (10) generalizations and interpretation, and (11) preparation of the report or presentation of the results, i.e., formal write-up of conclusions reached.

A brief description of the above-stated steps of the process of business research are as follows :-

Formulating the Research Problem: There are two types of research problems, viz., those which relate to states of nature and those which relate to relationships between variables. At the very outset the researcher must single out the problem he wants to study, i.e., the researcher must decide the general area of interest or aspect of a subject matter that he would like to inquire into. Initially, the problem may be stated in a broad general way and then the ambiguities, if any, relating to the problem be resolved. Then, the feasibility of a particular solution has to be considered before a working formulation of the problem can be set up.

The formulation of a general topic into a specific research problem, thus, constitutes the first step in a scientific inquiry. Essentially two steps are involved in formulating the research problem, viz.,understanding the problem thoroughly, and rephrasing the same into meaningful terms from ananalytical point of view.The researcher must at the same time examine all available literature to get himself acquainted with the selected problem. Researcher may review two types of literature; the conceptual literature concerning the concepts and theories, and the empirical literature consisting of studies made earlier which are similar to the one proposed. The basic outcome of this review will be the knowledge of what data and other materials are available for operational purposes which will enable the researcher to specify his research problem in a meaningful context. The problem to be investigated must be defined unambiguously for that will help discriminate relevant data from irrelevant ones.

Extensive Literature Survey: Once the problem is formulated, a summary of it should be written down. It is compulsory for a research worker writing a thesis for a Ph.D. degree to write a synopsis of the topic and submit it to the necessary Committee or the Research Board for approval.At this juncture, the researcher should undertake an extensive literature survey connected with the problem. For this purpose, abstracting and indexing journals and published or unpublished bibliographies are the first place to go. Academic journals, conference proceedings, government reports, books, etc., must be tapped depending on the nature of the problem.

Development of Working Hypotheses: After an extensive literature survey, the researcher should state in clear terms the working hypothesis or hypotheses. The working hypothesis is the tentative assumption made to draw out and test its logical or empirical consequences. As such how research hypotheses are developed is particularly important since they provide the focal point for research. They also affect how tests must be conducted in the analysis of data and indirectly the quality of data that is required for the analysis. In most types of research, the development of a working hypothesis plays an important role. The hypothesis should be very specific and limited to the piece of research in hand because it has to be tested. The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track.

Preparing the Research Design: The research problem has been formulated in clear-cut terms, the researcher will be required to prepare a research design, i.e., they have to state the conceptual structure within which research would be conducted. The preparation of such a design facilitates research to be as efficient as possible yielding maximal information. In other words, the function of research design is to provide for the collection of relevant evidence with minimal expenditure of effort, time and money. But how all these can be achieved depends mainly on the research purpose. Research purposes may be grouped into four categories, viz., Exploration, Description, Diagnosis and Experimentation. A flexible research design that provides the opportunity for considering many different aspects of a problem is considered appropriate if the purpose of the research study is that of exploration. But when the purpose happens to be an accurate description of a situation or an association between variables, the suitable design will minimize bias and maximize there liability of the data collected and analyzed.

Determining Sample Design: All the items under consideration 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 the highest accuracy is obtained. But in practice, this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observations increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or the use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Not only this, census inquiry is not possible in practice under any circumstances. For instance, blood testing is done only on a sample basis. Hence, quite often we select only a few items from the universe for our study purposes. The items so selected constitute what is technically called a sample. The sample design to be used must be decided by the researcher taking into consideration the nature of the inquiry and other related factors.

Collecting the Data: In dealing with any real-life problem it is often found that the data at hand are inadequate, and hence, it becomes necessary to collect appropriate data. There are several ways of collecting the appropriate data which differ considerably in the context of money costs, time and other resources at the disposal of the researcher. Primary data can be collected either through experiments or through asurvey. If the researcher experiments and observes some quantitative measurements, or the data, with the help of which the researcher examines the truth contained in his hypothesis. The researcher should select one of these methods of collecting the data taking into consideration the nature of investigation, objective and scope of the inquiry, financial resources, available time and the desired degree of accuracy.

Execution of the Project: Execution of the project is a very important step in the research process. If the execution of the project proceeds on the correct lines, the data to be collected would be adequate and dependable. The researcher should see that the project is executed in a systematic manner and on time. If the survey is to be conducted by using structured questionnaires, data can be readily machine-processed. In such a situation, questions, as well as the possible answers, may be coded. If the data are to be collected through interviewers, arrangements should be made for proper selection and training of the interviewers. The training may be given with the help of instruction manuals which explain clearly the job of the interviewers at each step. Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently. A careful watch should be kept for unanticipated factors to keep the survey as much realistic as possible. This, in other words, means that steps should be taken to ensure that the survey is under statistical control so that the collected information is as per the pre-defined standard of accuracy.

Analysis of Data: After the data have been collected, the researcher turns to the task of analyzing them. The analysis of data requires several closely related operations such as the establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences. The unwieldy data should necessarily be condensed into a few manageable groups and tables for further analysis. Thus, the researcher should classify the raw data into some purposeful and usable categories. Coding operation is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted. Editing is the procedure that improves the quality of the data for coding. With coding, the stage is ready for tabulation. Tabulationis a part of the technical procedure wherein the classified data are put in the form of tables. The mechanical devices can be made use of at this juncture. A great deal of data, especially in large inquiries, is tabulated by computers. Computers not only save time but also make it possible to study the large number of variables affecting a problem simultaneously. Analyzing work after tabulation is generally based on the computation of various percentages, coefficients, etc., by applying various well-defined statistical formulae.

Testing of Hypothesis: After analyzing the data as stated above, the researcher is in a position to test the hypotheses, if any, he had formulated earlier. Do the facts support the hypotheses or do they happen to be contrary? This is the usual question that should be answered while testing hypotheses. Various tests, such as the Chi-square test, t-test and F-test, have been developed by statisticians for this purpose. The hypotheses may be tested through the use of one or more of such tests, depending upon the nature and object of the research inquiry.

Generalizations and Interpretation: If a hypothesis is tested and upheld several times, it may be possible for the researcher to arrive at a generalization, i.e., to build a theory. The real value of research lies in its ability to arrive at certain generalizations. If the researcher had no hypothesis to start with, then the researcher might seek to explain their findings based on some theory, it is known as interpretation. The process of interpretation may quite often trigger off new questions which in turn may lead to further research.

Preparation of the Report or the Thesis: Finally, the researcher has to prepare a report of what has been done. Writing of report must be done with great care keeping in view the following:

> The layout of the report should be as follows: the preliminary pages, the main text, and the end matter. In its preliminary pages, the report should carry the title and date followed by acknowledgments and foreword. Then there should be a table of contents followed by a list of tables and a list of graphs and charts, if any, given in the report.

The report should be written in a concise and objective style in simple language avoiding vague expressions such as it seems, there may be, and the like.

Charts and illustrations in the main report should be used only if they present the information more clearly and forcibly.

Calculated 'confidence limits' must be mentioned and the various constraints experienced in conducting research operations may as well be stated.

1.8 SUMMARY

This lesson started with the discussion by emphasizing that research is the application of science to the art of management. Research methodology minimizes the degree of uncertainty involved in management decisions. The research lays the structure for decision-making. Research is not synonymous with common sense. The characteristics of business research highlight the various aspects related to good research. The role of research in the important areas of management has been briefly covered. The areas include marketing, production, banking, materials, human resource development, and government. The research process involves eleven important steps. All these steps have been explained in detail with their key elements.

1.9 GLOSSARY

Business Research : Process of acquring detailed information of all the areas of Business and using that information for sales and profit.

1.10 LESSON END EXERCISE

Briefly define the meaning of research.

____

____

____

Highlight the role of business research.

____

____

____

Briefly explain the meaning and importance of each of the following in research.

Systematic : ____

Objectivity : ____

Relevance : ____

Discuss the various steps of the research process.

____

____

____

1.11 SUGGESTED READINGS

Boyd, Westfall, and Stasch, *Business Research Text and Cases.* All India Traveller Bookseller, New Delhi.

Brown, F.E. *Business Research, a structure for decision making.* Addison - Wesley Publishing Company.

Kothari, C.R. *Research Methodology - Methods and Techniques.* Wiley Eastern Ltd.

Stockton and Clark, *Introduction to Business and Economic Statistics.* D.B. Taraporevala Sons and Co. Private Limited, Bombay.

-000-

M.COM III SEM **UNIT- I**

MCOMC351 **LESSON - 2**

PROBLEM IDENTIFICATION AND USEFULNESS OF OBJECTIVES AND HYPOTHESES IN BUSINESS RESEARCH

STRUCTURE

2.1 INTRODUCTION

Educators can use problem identification to seek solutions to resolve problems in the school environment that occur at the district, school, class, and individual levels. Problem identification is part of the scientific method, as it serves as the first step in a systematic process to identify, evaluate a problem and explore potential solutions. You need to make sure you are identifying the true, underlying problem causing the public health issue and this is not always obvious. Objectives give you something to work towards and help to direct energy and effort. They stimulate the need to act. Research has found that where staff is involved in setting their objectives they are more motivated to achieve them. A hypothesis becomes an important aspect of any research. It helps to assume the probability of research failure and progress. It helps to provide a link to the underlying theory and specific research question. It helps in data analysis and measures the validity and reliability of the research. Different types of hypotheses that are used in the process of business research were also highlighted.

2.2 OBJECTIVES

After going through this lesson, you will be able to describe:-

The problem identification, how to formulate and define the research problem.

The usefulness of objectives in business research. The usefulness of hypothesis in business research. Different types of hypotheses.

Difference between null and alternate hypothesis.

2.3 PROBLEM IDENTIFICATION

The problem means a question or an issue to be examined. A research problem refers to some kind of problem that a researcher experiences or observes in the context of either a theoretical or practical situation. The researcher has to find out a suitable course of action by which the objective can be attained optimally in the context of the given environment. Thus, the selection of research problems has high value to society and the researcher must be able to identify those problems that need an urgent solution.

2.3.1 Characteristics of a Good Research Problem

To fulfill the need for a good research problem the following are the essential characteristics of an effective research problem. However, ensure to consider and cover most of these characteristics to enable researchers to look at examine, and understand the marketing research problem.

Clear and Unambiguous: The researcher must have a clear specific problem statement in the research. The complex research project may end up wasting time and money if these issues are not of high importance. The researcher must ensure that they are always concerned about the critical needs and make the right strategies.

Logical and Systematic: The problem should be logically or clearly stated in the research project, the problem likely is a weak one or probably a non-problem. To test this, consolidate the specific problem into a paragraph and ensure it makes sense and covers all the crucial points. Share the problem with others, and if this is not understood by them, consider a more systematic approach to specify the issue or problem.

Empirical: There is a difference between beliefs and facts. Keep fiction out of this. Research must be based on facts and not beliefs. Hypothetical events will not do the research any good. The researcher must consider the research findings true or accurate if not based on facts and evidence. It must suggest a meaningful and testable hypothesis.

Relation Between Variables: There should be a proper relationship among the values of the variable corresponding to the values of another variable, for each case in the sample/dataset. In other words, knowing the value of one variable, for a given case, helps you to predict the value of the other one. If the variables are perfectly related, then knowing the value of one variable tells exactly what the value of the other variable is. To measure the relationship among variables, what is the level of measurement of the variable is to be known. The level of measurement determines what kinds of mathematical operations can meaningfully be performed on the values of a variable.

Proper Questionnaire: The research must create multiple questions. These questions should be more specific to the research that highlights different components or aspects of the problem. These questions must assist in addressing the issue better, providing a robust framework for the research. Formulating such questions is a challenge that needs to be addressed correctly.

Verifiable: Research is based on facts and findings, and there must be multiple cases or findings to prove the research and it should be verifiable. Research that is based on limited data will do no good to test a theory or satisfy it. If data is unavailable to do research, it does not make sense to go ahead with the research.

2.3.2 Various aspects of a problem

For an effective formulation of the problem, following aspects are to be considered by the researcher :-

Definition of the problem: Before one takes up a problem for the study one needs to be defined it properly. The issues for inquiry are to be identified clearly and specified in detail. If anyexisting theoretical framework is tested, the particular theorem or theories must be identified.Similarly, if there are any assumptions made and terms used the meaning of them must be made clear. As far as possible the statement of the problem should not give any scope for ambiguity.

Scope of the problem: The researcher has to fix up the four walls of the study. There searcher must identify which of the aspects they are trying to prove.

Justification of the problem: Many a time research studies are put to the test of justification or relevance. In the scientific curiosity of the problems, the problem that needs an urgent solution must be given preference.

Feasibility of the problem: Although a problem needs urgent attention and is justifiable in several respects, one has to consider its feasibility of the same. Feasibility means the possibility of conducting the study successfully. The elements of time, data and cost are to be taken into consideration before a topic is selected for the study.

The originality of the problem: In social sciences, particularly in commerce and management, there is no systematic compilation of the works already done or on hand. Two people may be doing work more or less on a similar topic. In such situations, it is not advisable to continue work in the same manner. What is advisable is that each of them should try to focus on different aspects, so that they could enrich their field of knowledge with their respective studies.

2.3.3 Defining and Formulating a Research Problem.

Research is to be defined along with the bounds in which it is to be studied. Therefore defining a problem involves the task of laying down boundaries within which a researcher shall study the problem with a predetermined objective in view. Defining a research problem is acrucial part of a research study and must in no case be accomplished hurriedly.

Steps for Defining a Research Problem

Stating the problem in a general way: The researcher should state the problem in general terms, keeping in view either some practical concern or some scientific or intellectual interest. Often the guides put forth the problem in general terms and the researcher narrows down the problem and phrases the problem in operational terms. The problem stated generally may contain various ambiguities which must be resolved by proper thinking and rethinking the problem.

Understanding the nature of the problem: For understanding the nature of the problem in a better way, the researcher has to hold discussions with those who have knowledge of the problem.

Surveying the available literature: This is necessary because only through such a survey, a researcher can understand the relevant theories, reports, etc.

studies on related problems are useful for knowing the type of difficulties that may encounter in the present study.

Developing the ideas through discussions: A researcher must discuss his problem with his colleagues and those who have enough experience in the same area or in working on similar problems. People with experience can enlighten the researcher on various aspects of the study.

Rephrasing the research problem:- A researcher must rephrase the research problems into a working proposition. 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 a working hypothesis.

2.4 USEFULNESS OF OBJECTIVES IN BUSINESS RESEARCH

Research can simply be explained as a systematic search for knowledge. To successfully conduct research, a research topic/subject has to be decided. The research process involves gathering data on the research subject and analyzing it to reach a comprehensive conclusion. But the important question here is; who will decide what should be searched and how it will be decided? The answer is that the researcher has to define the objectives and decide what needs to be searched. The researcher has full control of the research subject. But how it will be decided what needs to be searched? It's tricky but involves simple logic. In our daily lives, if we know our destination, we can choose a path to reach there.

Similarly in research, if we know what we are looking for then we can decide how we can accomplish it. Thus, we can say that the intended results define the framing of the objectives. Here the objectives come into the scene and gain importance :-

Objectives give direction to the paper: The objectives in a study provide a clear direction. As soon as the researcher defines the objectives, the scope is defined. Based on the objectives, further processes are carried out including primary and secondary data collection, data analysis, drawing interpretation and conclusion of the research. All these research processes are dependent on the objectives and are carried out by it.

Objectives help to avoid any diversion from the topic: Once the researcher starts gathering data for the research, especially the secondary data, the researcher comes across abundant data related to the research subject. The researcher often gets blinded by the wide amount of data available on a variety of sources and frequently diverts from the core subject. The abundance of data brings in confusion as to which data to be utilized for present research and which data should be discarded. The objectives help the researcher to stick to the current research and avoid any diversion from the research topic.

Research objectives minimize wastage of resources: By preventing the researcher to deviate from the research topic, objectives minimize the wastage of the researcher's time, money and energy. Objectives help the researcher to concentrate on the current research. Wastage of the research resources is reduced with clearly-defined objectives and hence, the efficiency of the study is enhanced.

Objectives ease the understanding of the research by the target audience: All research goes in vain if the target audience i.e. the beneficiaries of the research fails to understand it. The objectives help the target audience to clearly understand the purpose of particular research and it, therefore, eases understanding. Thus, objectives also make research meaningful for the target audience.

2.5 USEFULNESS OF HYPOTHESIS IN BUSINESS RESEARCH

The hypothesis is a tentative statement showing the relationship between two or more variables, there liability and validity of which is to be tested and verified. It expresses the nature and degree of the relationship between variables. It also means a mere assumption or some supposition to be proved or disproved. But for a researcher hypothesis is a formal question that intends to be resolved.

Thus a hypothesis may be defined as a proposition or a set of propositions set forth as an explanation for the occurrence of some specified group of phenomena either asserted merely as a provisional conjecture to guide some investigation or accepted as highly probable in the light off acts. Quite often a research hypothesis is a predictive statement, capable of being tested by scientific methods, that relates an independent variable to some dependent variable.

2.5.1 Characteristics of Hypothesis

The hypothesis must possess the following characteristics:

The hypothesis should be clear and precise. If the hypothesis is not clear and precise, the inferences drawn on its basis cannot be taken as reliable.

The hypothesis should be capable of being tested. In a swamp of untestable hypotheses, many a time the research programs have bogged down. Some prior studies may be done by researchers to make the hypothesis a testable one. A hypothesis "is testable if other deductions can be made from it which, in turn, can be confirmed or disproved by observation."

The hypothesis should state the relationship between variables if it happens to be a relational hypothesis.

The hypothesis should be limited in scope and must be specific. A researcher must remember that narrower hypotheses are generally more testable and he should develop such hypotheses.

The hypothesis should be stated as far as possible in most simple terms so that the same is easily understandable by all concerned. But one must remember that simplicity of the hypothesis has nothing to do with its significance.

The hypothesis should be consistent with most known facts i.e., it must be consistent with a substantial body of facts. In other words, it should be one that judges accept as being the most likely.

The hypothesis should be amenable to testing within a reasonable time. One should not use even an excellent hypothesis if the same cannot be tested in a reasonable time for one cannot spend a lifetime collecting data to test it.

The hypothesis must explain the facts that gave rise to the need for explanation. This means that by using the hypothesis plus other known and accepted generalizations, one should be able to deduce the original problem condition. Thus hypothesis must explain what it claims to explain; it should have an empirical reference.

2.5.5 Different Types of Hypotheses

Simple Hypothesis: In this type of hypothesis all parameters of the distribution are specified. It predicts the relationship between two variables i.e. the dependent and the independent variable. A Simple hypothesis is also known as a composite hypothesis.

Complex Hypothesis: In the complex hypothesis relationship between two or more independent variables and two or more dependent variables was examined.

Research Hypothesis: A research hypothesis is a specific, clear prediction about the possible outcome of a scientific research study based on specific factors of the population. It is also known as the working hypothesis.

Null Hypothesis: A null hypothesis is a general statement that states no relationship between two variables or two phenomena. It is usually **denoted by H_0.**

Alternative Hypothesis: An alternative hypothesis is a statement that states some statistical significance between two phenomena. It is usually **denoted by H_1 or H_A.**

Logical Hypothesis: A logical hypothesis is a planned explanation holding limited evidence.

Statistical Hypothesis: A statistical hypothesis, sometimes called confirmatory data analysis, is an assumption about a population parameter.

2.5.3 Difference Between Null Hypothesis And Alternate Hypothes

S. No.	Null Hypothesis	Alternative Hypothesis
	A null hypothesis represents the hypothesis that there is "no relationship" or "no association" or "no difference" between two variables.	An alternative hypothesis is the opposite of the null hypothesis where we can find some statistical importance or relationship between two variables.
	In a null hypothesis, the researcher tries to invalidate or reject the hypothesis.	In an alternative hypothesis, the researcher wants to show or prove some relationship between variables.
	It is an assumption that specifies a possible truth to an event where an effect is absent.	It is an assumption that describes an alternative truth where there is some effect or some difference.
	A n ull hypothesis is a statement that signifies no change , no effect and no differences between variables .	An a lternative hypothesis is a statement that signifies some change, some effect and some differences between variables.
5	If the null hypothesis is true, any discrepancy between observed data and the hypothesis is only due to chance.	If the alternative hypothesis is true, the observed discrepancy between the observed data and the null hypothesis is not due to chance.
	A null hypothesis is denoted as H_0.	An alternative hypothesis is denoted as H_1 or H_A.
	Example : There is no life on the moon. This need not to be proved because it is the universal truth. $H_0:\mu=0$	Example: There is life on the moon. This need to be proved. $H_A:\mu\neq 0$

Table 2.1 Difference Between Null and Alternate Hypothesis

2.6 SUMMARY

This lesson started with a brief introduction to problem identification where you learned about the characteristics of a good research problem with various aspects. How to define and formulate the business research problem. Why objectives are useful in defining the research problem. The role of hypothesis with its types was also highlighted. Lastly, some points of difference between alternate and null hypotheses were discussed.

2.7 GLOSSARY

Problem Identification: Problem identification requires the use of an appropriate measure or assessment tool to determine whether a problem exists.

Hypothesis: A hypothesis in a scientific context, is a testable statement about the relationship between two or more variables or a proposed explanation for some observe phenomenon.

Nul Hypothesis : A null hypothesis is a type of statistical hypothesis that proposes statistical significance exists in a set of given observations.

Alternate Hypothesis: The alternative hypothesis is a statement used in a statistical inference experiment. An alternative theory is a statement that a researcher is testing.

2.8 LESSON END EXERCISE

What are the characteristics of a good research problem?

____

____

____

Briefly define the steps involved in defining the business research problem.

____

____

____

Why is hypothesis an important part of research?

____

____

____

Highlight the difference between null and alternative hypotheses.

____

____

____

2.9 SUGGESTED READINGS

Boyd, Westfall, and Stasch, *Business Research Text and Cases.* All India Traveller Bookseller, New Delhi.

Brown, F.E. *Business Research, a structure for decision making.* Addison - Wesley Publishing Company.

Cooper, D.R.,Schildler, P.S. *Business Research Methods.* Tata Mc Graw Hill, New Delhi.

Kothari, C.R. *Research Methodology - Methods and Techniques.* Wiley Eastern Ltd.

DATA COLLECTION METHODS FOR BUSINESS RESEARCH - OBSERVATION, INTERVIEWS, QUESTIONNAIRE AND SCHEDULES

STRUCTURE

3.1 INTRODUCTION

The increasingly complex nature of business and government has focused attention on the uses of research methodology in solving managerial problems. The credibility of the results derived from the application of such methodology is dependent upon the up-to-date information about the various pertinent characters included in the analysis. To illustrate, the demand for disc records dropped dramatically after cassettes entered the market commercially. This information must be taken into consideration for formulating a marketing strategy by a dealer selling musical products. The information expressed in the appropriate quantitative form is known as data. The necessity and usefulness of information gathering or data collection cannot be overemphasized in government policies. The government must be aware of the actual scenario of the acceptance of family planning before it can formulate any policy on this matter. The components of this scenario are provided by appropriate data to be collected from various families. In industrial disputes regarding wages, the cost of living index, a data-based indicator of inflation is often accepted as a guideline for arbitration.

3.2 OBJECTIVES

After going through this lesson, you will be able to describe:-

- Discuss the necessity and usefulness of data collection;
- Distinguish between primary and secondary data;
- Describe different methods of data collection;
- Describe the concept of observation, interview, questionnaire and schedule with its types.

3.3 DATA COLLECTION FOR BUSINESS RESEARCH

Data collection is defined as the procedure of collecting, measuring and analyzing accurate insights for research using standard validated techniques. A researcher can evaluate their hypothesis based on collected data. In most cases, data collection is the primary and most important step for research, irrespective of the field of research.

The approach to data collection is different for different fields of study, depending on the required information. The most critical objective of data collection is ensuring that information-rich and reliable data is collected for statistical analysis so that data-driven decisions can be made for research.

Data represents information collected in the form of numbers and text. Data Collection is generally done after the experiment or observation. Data collection is helpful in planning and estimating. Data collection is either qualitative or quantitative. Quantitative data is expressed in numbers and graphs and is analyzed through statistical methods whereas qualitative data is expressed in words and analyzed through interpretations and categorizations.

3.3.1 Importance of data collection

There are a bunch of underlying reasons for collecting data, especially for a researcher. Walking you through them, here are a few reasons.

The integrity of the Research: A key reason for collecting data, be it through quantitative or qualitative methods is to ensure that the integrity of the research question is indeed maintained.

Reduce the likelihood of errors: The correct use of appropriate data collection of methods reduces the likelihood of errors consistent with the results.

Decision Making: To minimize the risk of errors in decision-making, accurate data must be collected so that the researcher doesn't make uninformed decisions.

Save Cost and Time: Data collection saves the researcher time and funds that would otherwise be misspent without a deeper understanding of the topic or subject matter.

To support a need for a new idea, change, and/or innovation: To prove the need for a change in the norm or the introduction of new information that will be widely accepted, it is important to collect data as evidence to support these claims.

3.4 METHODS OF DATA COLLECTION FOR BUSINESS RESEARCH

The task of data collection begins after a research problem has been defined and the research plan chalked out. When deciding about the method of data collection to be used for the study, the researcher should keep in mind two types of data.

Primary Data

Secondary Data

Primary Data:-When an investigator collects data himself with a definite plan or design in their way, then the data is known as primary data. Generally, the results derived from the primary data are accurate as the researcher gathers the information. But, one of the disadvantages of primary data collection is the expenses associated with it. Primary data research is very time-consuming and expensive.

Secondary Data:-Data that the investigator does not initially collect but instead obtains from published or unpublished sources are secondary data. Secondary data is collected by an individual or an institution for some purpose and is used by someone else in another context. It is worth noting that although secondary data is cheaper to obtain, it raises concerns about accuracy. As the data is second-hand, one cannot fully rely on the information to be authentic.

3.4.1 Observation Method

Observation is the systematic viewing of a specific phenomenon in its proper setting for the specific purpose of gathering data for a particular study.In this method of data collection, the investigator collects the requisite information personally through observation. As the investigator is solely responsible for the Collection of data by this method, his training, skill and knowledge play an important role in the quality of primary data. A slight variation of this procedure is an indirect oral investigation where the data are collected through indirect sources. Persons who are likely to have information about the problem are interrogated and based on their answers, primary data become available. Most of the commissions of enquiry or committees appointed by the government collect primary data by this method. The accuracy of the primary data collected by this method depends largely upon the type of persons interviewed and hence these persons have to be selected very carefully. For example, to study the conditions of students residing in a university, the investigator meets the students in their hostels and collects necessary data after a personal study.

Types of Observation Methods

Following are the types of observation methods :

Structured Observation Method:-This is a systematic observation method where data is collected as per a pre-defined schedule. The specific variable is used in this method for data collection.

Unstructured Observation Method:-The unstructured observation method is conducted in a free and open manner without using any pre-determined objectives, schedules or variables.

Direct Observation Method:-Direct observation refers to the situation when the observer remains physically present and personally monitors what takes place.This approach is very flexible because it allows the observer to react to and report subtle aspects of events as they occur.

Indirect Observation Method:-Indirect observation occurs when the recording is done by mechanical, photographic, videotape, cameras, or other electronic means.

Controlled Observation Method:-The controlled observation is carried out in a closed space. It is the researcher who has the authority to decide the place and the time where and when the observation will take place. He also decides who the participants will be and in what circumstances will he use the standardized process.

Naturalistic Observation Method:-Social scientists and psychologists generally use the naturalistic observation method. The process involves observing and studying the spontaneous behavior of the participants in open or natural surroundings. The role of the researcher is to find and record whatever he can see and observe in the natural habitat.

Participant Observation Method:-The participant observation method is often considered a variant of the naturalistic observation method because it has some similarities with it. The point of difference is that the researcher is not a distant observer anymore because he has joined the participants and become a part of their group. He does this to get a more in-depth and greater insight into their lives.

Advantages of the Observation Method

Directness:-The main advantage of observation is its directness. We can collect data at the time they occur. The observer does not have to ask people about their behavior and reports from others. The researcher can simply watch as individuals act and speak. While the survey respondents may have a hazy or lapsed memory about events that occurred in the distant past, the observer is studying events as they occur.\

Natural Environment:- Observation is neither as restrictive nor as artificial as other techniques that are artificiality introduced into the research environment, data collected in an observation study describes the observed phenomena as they occur in their natural setting.

Longitudinal Analysis:-Since the observation is possible to be conducted in a natural setting, the observer can conduct his or her study over a much longer period than with either the survey or experiment.

Non-Verbal Behavior:-Observation is superior to survey research, experimentation, or document study for collecting data on non-verbal behavior. Some studies focus on individuals who are unable to give verbal reports or articulate themselves meaningfully.

Disadvantages of the Observation Method

Lack of Control:-Despite the advantage achieved from the natural environment, the observation study, however, has little control over extraneous variables that may affect the data.The presence of a stranger (the observer) and the error involved in human observation and the recording of data, which may remain out of control of the observer, are likely to bias the observations to a great extent.

Difficulties in Quantification:-Measurement in observational studies generally takes the form of observer's un-quantified perceptions rather than the quantitative measures often used in survey and experimental studies.

Smallness in Sample Size:-Because observational studies are generally conducted in-depth, with data that are often subjective and difficult to quantify, the sample size is usually kept at a minimum. Also, the in-depth nature of the observation studies generally requires that they are conducted over an extended period, than the survey method or experiments. This feature tends to limit the size of the sample.

No Opportunity to Learn about the past :- In an observational study, there is no way to know the past. It is also difficult to gather information on such topics as intentions, opinions, attitudes, or preferences.

3.4.2 Interviews

An interview is one of the popular methods of research data collection. The term interview can be dissected into two terms 'inter' and 'view'. The essence of the interview is that one mind tries to read the other. The interviewer tries to assess the interviewed in terms of the aspects studied or issues analyzed. It is generally a qualitative research technique that involves asking open-ended questions to converse with respondents and collect elicit data about a subject. The interviewer in most cases is the subject matter expert who intends to understand respondent opinions in a well-planned and executed series of questions and answers.The interview method of collecting data involves the presentation of oral-verbal stimuli and replies in terms of oral-verbal responses.

Purpose and Importance of Interview

The main purpose of the interview as a tool of data collectionis to gather data extensively and intensively. According to Pauline, the objectives of the interview may be an exchange of ideas and experiences, eliciting information about a very wide range of data in which the interviewee may wish to rehearse his past, define his present and canvas his future possibilities. Thus, in brief, the objectives of the interviewee are two-fold:

To exchange ideas and experiences

and To elicit information.

The importance of the interview may be known through these points:

It is the method best suited for the assessment of personal qualities.

It has definite values for the diagnosis of emotional problems and treatments.

It is one of the major bases upon which counseling procedures are carried

out. It provides information to supplement other methods of collecting data.

It may be used, in addition to observation, to verify information obtained through correspondence methods.

Types of Interviews

There are different types of interviews used in the research data collection.

Structured Interview:-Structured interviews are defined as research tools that are extremely rigid in their operations and allow very little or no scope of prompting the participants to obtain and analyze results. It is thus also known as a standardized interview and is significantly quantitative in its approach. Questions in this interview are pre-decided according to the required detail of information. It is typically formal and organized that may include several interviewers, commonly referred to as a panel interview an interviewer who has a more structured style will usually begin the interview. These types of interviews are excessively used in survey research to maintain uniformity throughout all the interview sessions. There can be closed-ended as well as open-ended according to the type of target population. Closed-ended questions can be included to understand user preferences from a collection of answer options whereas open-ended can be included to gain details about a particular section in the interview.

Semi-Structured Interview:-Semi-structured interviews offer a considerable amount of freedom to the researcher to probe the respondents along with maintaining a basic interview structure. Even if it is a guided conversation between researchers and interviewees, appreciable flexibility is offered to the researchers. A researcher can be assured that multiple interview rounds will not be required in this type of research interview. It does maintain some structure but it also provides the researcher with the ability to probe the participant for additional details. Keeping the structure in mind, the researcher can follow any idea or take creative advantage of the entire interview. Additional respondent probing is always necessary to garner information for a research study. The best application of a semi-structured interview is when the researcher doesn't have time to conduct research and requires detailed information about the topic.

Unstructured Interview:-Unstructured interviews are usually described as conversations held with a purpose in mind to gather data about the research study. They are also known as in-depth interviews. These interviews have the least number of questions as they lean more towards a normal conversation but with an underlying subject. It is a relatively formless interview style that researchers use to establish rapport and comfort with the participant, and is extremely helpful when researchers are discussing sensitive topics. The main objective of most researchers using unstructured interviews is to build a bond with the respondents due to which there are high chances that the respondents will be 100% truthful with their answers. There are no guidelines for the researchers to follow so, they can ethically approach the participants to gain as much information as they possibly can for their research topic.

There are also some other types of interviews, a few of them are described below:

Focus-Group Interviews:-A focus-group interview is one in which a group of people jointly participate in an unstructured-indirect interview. The group, usually consisting of 8 to 12 people, is generally selected purposively to include persons who have a common background or similar buying or use experience that relates to the problem to be researched. The interviewer, moderator, as he or she is more often called, attempts to focus the discussion on the problem areas in a relaxed, nondirected manner. The objective is to foster involvement and interaction among the group members during the interview will lead to spontaneous discussion and the disclosure of attitudes, opinions, and information on present or prospective buying and use behavior.

Focused Interviews: This is a semi-structured interview where the investigator attempts to focus the discussion on the actual effects of a given experience to which the respondents have been exposed. It takes place with the respondents known to have been involved in a particular experience. The focused interview permits the interviewer to obtain details of personal reactions, specific emotions and the like. The merits of using this type of interview are that, it's free from the inflexibility of formal methods, yet gives the interview a set form and insured adequate coverage of all the relevant topics. The respondent is asked for certain information, yet he has plenty of opportunities to present his views. The interviewer is also free to choose the sequence of questions and determine the extent of probing.

The Third-Person Technique: The simplest way of obtaining information through indirect questioning of a respondent is to ask for the view of a neighbor, an (unnamed) associate, or some other person whose views on the subject at hand might reasonably be known. This permits the respondent to project his views with no feeling of social pressure to give an "acceptable" answer.

The Personal Interview: As the name implies, the personal interview consists of an interviewer asking questions of one or more respondents in a face-to-face situation. The interviewer's role is to get in touch with the respondents, ask the desired questions, and record the answers obtained. The recording of the information obtained may be done either during or after the interview. In either case, it is a part of the interviewer's responsibility to ensure that the content of the answers is unambiguous and that it has been recorded correctly.

The Telephone Interview: Telephone interviews are sometimes used instead of personal interviews, especially when the information must be collected quickly and inexpensively and the amount of information required is limited. The telephone interview is well suited to such research problems as determining "coincidental" viewing of television or listening to radio programs.

Advantages of the Interview Method

Personal interviews, compared especially to questionnaires usually yield a high percentage of returns.

The interview method can be made to yield an almost perfect sample of the general population because practically everyone can be reached by and can respond to this approach. It will be remembered that the questionnaire approach is severely limited by the fact that only literate persons can be covered by it.

The information secured through interviews is likely to be more correct compared to that secured through other techniques. The interviewer who is present on the spot can clear up the seemingly inaccurate or irrelevant answers by explaining the questions to the informant. If the informant deliberately falsifies replies, the interviewer can effectively check them and use special devices to verify the replies.

Scoring and test devices can be used, the interviewer acting as an experimenter. At the same time, visual stimuli to which the informant may react can be presented.

The use of the interview method ensures a greater number of usable returns compared to other methods. Returned visits to complete items on the schedule or to correct mistakes can usually be made without annoying the informant.

The interviewer can usually control which person or persons will answer the questions. This is not possible in the mailed questionnaire approach. If so desired and warranted group discussions may also be held.

A personal interview may take long enough to allow the informant to become oriented to the topic under investigation. Thus, recall of relevant information is facilitated. The informant can be made to devote more time if, as is the case, the interviewer is present on the spot to elicit and record the information.

The interview is a more appropriate technique for revealing information about complex, emotionally-laden subjects or for probing the sentiments underlying an expressed opinion.

Disadvantages of the Interview Method

In terms of cost, energy and time, the interview approach poses a heavy demand. The transportation cost and the time required to cover addresses in a large area as also the possibility of non-availability may make the interview method uneconomical and often inoperable.

The efficacy of interviews depends on the thorough training and skill of interviewers as also on rigorous supervision over them. Failing this, the data recorded may be inaccurate and incomplete.

The presence of the interviewer on the spot may overstimulate the respondent, sometimes even to the extent that he may give imaginary information just to make it interesting. He may tell things about which he may not himself be very sure.

In the interview method, the organization required for selecting, training and supervising a field staff is more complex.

It is the usual experience that costs per interview are higher when field investigators are employed. This is especially so when the area to be covered is widely spread out.

Interviewing may also introduce systematic errors. For example, if the interviews are conducted at their homes during the day, a majority of informants will be housewives. Now if the information is to be obtained from the male members, most of the fieldwork will have to be done in the evening or on holidays. If this is the case, only a few hours can be used per week for interviews.

Many actions human beings carry out are not easily verbalized, but easily observed. Through observation, a social process may be followed as it develops. Verbal techniques such as interviews may give valuable reports, but post hoc, unless one is dealing with rather unusual respondents capable of acting and being interviewed at the same time.

3.4.3 Questionnaire

This method of data collection is quite popular, particularly in the case of big inquiries. It is being adopted by private individuals, research workers, private and public organizations and even by governments. In this method, a questionnaire is sent to the persons concerned with a request to answer the questions and return the questionnaire. A questionnaire consists of several questions printed or typed in a definite order on a form or set of forms. The questionnaire is mailed to respondents who are expected to read and understand the questions and write down their replies in the space meant for the questionnaire itself. The respondents have to answer the questions on their own. A questionnaire is the most popular and widely used tool for collecting primary data. It suits any kind of research problem. It is not used only in the marketing field, but also in all types of social research projects. It consists of a formulated series of questions that contains a set of questions systematically and deliberately prepared to investigate the problem. It is a list of various types of questions related to a specific area or problem, expressed in some logical patterns and order, which can be used for data collection purposes.

Types of Questionnaire

The questionnaire can be classified based on several criteria as stated below:

Based on the Structure and Disguise there are four types of the questionnaire:

Structured Undisguised: This type of questionnaire involves structured and undisguised questions. The response is limited to certain options. A structured means that answers to the questions are predetermined. Respondents have to select the answer from the given list of answers. Undisguised means questions are open-ended. They are asked directly. Respondents can know what the researcher wants to know.

Unstructured Undisguised: Unstructured means free questions are asked. Their response is not limited to certain answers only. They have full freedom to answer the question. In short, answers to the question are not decided in advance.

Structured Disguised: Structured means the answers to the questions to be asked are determined in advance. Disguised means the indirect way of asking questions. Customers do not know the exact purpose/intention of a question but can answer it easily.

Unstructured Disguised: Here, the response is not fixed. Respondents have full freedom to answer the question. Disguised means something hidden.

Based on the purpose of the questionnaire, there are three types of the questionnaire:

Questionnaire for Personal Interview: This questionnaire is prepared to administer for the personal interview. It may involve more questions and indirect questions which require explanation or clarification.

Questionnaire for Telephone Survey: This questionnaire is prepared to collect information via telephone. Such a questionnaire involves a limited number of short and simple questions.

Questionnaire for Mail Survey: This questionnaire is meant for the mail survey. This questionnaire is sent to respondents with a request to return it dully filled. It also involves short and simple questions. However, it consists of more questions.

Based on administration, there can be two types of the questionnaire:

Interviewer-administered Questionnaire: When this questionnaire is to be administered, the presence of both interviewer and respondents is essential. Here, questions are asked to respondents one by one. Their response is recorded either in the same questionnaire or in a separate form.

Self-administered Questionnaire: It is given to the respondent by the interviewer to fill in her/his answers. This is possible even in absence of an interviewer. For a mail survey, this type of questionnaire is used.

Based on the type of questions, there can be two types of the questionnaire:

- **Simple Questionnaire:** This questionnaire involves a certain number of questions of the same type. It may involve only dichotomous, multiple-choice, or otherwise. It has limited utility.
- **Multiple Questionnaires:** It involves a variety of questions. Such a questionnaire consists of certain questions of different categories. It is a popular questionnaire.

Advantages of Questionnaire

Following are the advantages of questionnaire :

Economical: It is an economical way of accumulating information. It is economical both for the sender and for the respondent in time, effort and cost. The cost of conducting the study with the help of the questionnaire method is very low. In the questionnaire, the researcher has to spend on paper printing and postage only. There is no need to visit every respondent personally. So it does not require a high cost for the conduct of the research.

Wide Coverage: It is probably the best method to collect information, compared to the other methods like interviews or observation when the sample population is spread over a large territory. It permits nationwide or even international coverage. The questionnaire makes it possible to contact many people who could not otherwise be reached. It can cover a large group at the same time. Goode and Hatt say that when the researcher has to cover a group of respondents who are widely scattered, the researcher can use the questionnaire to minimize the cost.

Rapidity: Replies may be received very quickly in the questionnaire method. In this case, there is no need to visit the respondent personally or continue the study over a long period. Therefore in comparison with other methods, the mailed questionnaire is the quickest method.

Suitable in Special Type of Response: The information about certain personal, secret matters can be best obtained through the questionnaire method.

Repetitive Information: Compared to other methods like schedule, interview or observation, the questionnaire method is regarded as more useful and cheap, where the repetitive information has to be collected at regular intervals.

An Easier Method: A questionnaire is comparatively an easier method to plan, construct and administer. It does not require much technical skill or knowledge.

Uniformity: It helps in focusing the respondent's attention on all the significant items. As it is administered, in a written form, its standardized instructions for recording responses ensure some uniformity. The questionnaire does not permit many variations.

Useful Preliminary Tool: The questionnaire may be used as a preliminary tool for conducting a depth study later on by any other method.

Greater Validity: The questionnaire has some unique merits as regards the validity of the information. In methods like interview and observation, the reliability of responses depends on the way the investigator has recorded them. Here they may present biased or prejudiced information of their own. But in the questionnaire method, the responses given by the subjects are available in their language and version. Therefore, it cannot be wrongly interpreted by the researcher.

Anonymity: The questionnaire ensures anonymity to its respondents. The respondents have greater confidence that they will not be identified by anybody for giving a particular view or opinion. They feel more comfortable and free to express their view in this method.

Flexibility: The questionnaire is no doubt the most flexible tool for collecting both quantitative and qualitative information.

Disadvantages of Questionnaire

Limited Response: One of the major limitations of the questionnaire is that it can be applied only to those respondents who have a considerable amount of education. It can neither be used for illiterate nor for semi-literate persons. The questionnaire quite often fails to cover very busy and pre-occupied persons among the respondents, lazy and indifferent type of persons, the type of respondents who need to conceal a lot about themselves, the easy-going and shirkers among the respondents, the persons who have an unreasonable contempt for research and reform and the persons who unnecessarily doubt the research worker’s intentions, sincerity, devotion and commitment.

Lack of Personal Contact: As in the case of the questionnaire the researcher does not go to the field, and is not able to establish a proper personal relationship with the respondents. If the respondent fails to understand some of the technical terms or he has any doubt, there is nobody to clarify these technical terms or doubts. Without proper personal contact, it is very difficult to motivate the respondent to fill up the questionnaire.

Poor Response: In the case of the mailed questionnaire method, the proportion of returns is usually low. The factors which are likely to affect the returns are the layout of the questionnaire, its size, the organization conducting the research work, the nature of an appeal, the kind of respondents chosen for research, inducement for the response, etc.

Unreliability: The information collected through thequestionnaire cannot be said to be very much reliable or valid. If the subject misinterprets a question or gives an incomplete or indefinite response very little can be done to connect such a response. As against this, in an interview, there is always the possibility of rephrasing questions for further clarification. The investigator here is not in a position to observe the gestures and expressions of the respondents. He cannot cross-check the inconsistencies or misrepresentation of the replies. So in the questionnaire method, the reliability of responses is very low.

Illegibility: Illegible handwriting of the respondent sometimes creates many difficulties for the researcher to understand the responses. Sometimes the respondents erase and overwrite too much. These create many difficulties in reading the answers.

Incomplete Entries: Often most of the respondents fill up the questionnaire form very poorly. They sometimes leave out many questions altogether or fill in such a way that, it becomes very difficult on the part of the investigator to follow those responses. Other than this, there may be the problem of language, use of abbreviations and ambiguous terms, etc. All these make a questionnaire an incomplete one.

Possibility of Manipulated Entries: In the case of the interview the investigator directly interacts with the respondents personally' and intensively in a face-to-face situation. He can judge a respondent, his attitude, and understanding of the research topic and, if necessary, can ask some cross-questions to correct various errors. So usually the respondent cannot manipulate his answer. But in the questionnaire, it is very difficult to detect the errors of the respondents. Here the investigator does not have any facility to check the validity and reliability of the information. In the absence of the researcher, the respondents may supply manipulated information.

Useless in Depth-Studies: In the questionnaire method, it is not possible on the part of the researcher to conduct an intensive or in-depth study of the feelings, reactions and sentiments of the respondents. All these require a healthy interaction of the researcher with the respondents. But in the questionnaire method, the investigator is not present in the field, so nothing can be done to establish rapport with the respondent. Due to this lack of interaction with the respondent, the researcher cannot go into the details of the respondent's life. So through the questionnaire method, one cannot conduct an in-depth study.

Improper Response: The respondents who return the questionnaires may not constitute a representative section of the entire group. Only mere responsible, research-minded, or those in favor of the issue may prefer to respond. Some of the important sections of the group may remain silent. This vitiates the conclusions and findings.

Lack of Rapport with the Subject: Many people would not like to share any important information unless and until they are impressed with the rationale of the study and the personality of the investigator. The questionnaire does not provide any opportunity for the investigator to establish rapport with the subject and this cannot attract the respondent for a better response.

Not Suitable for Delicate Issues: Some of the research areas are so delicate, sensitive, intricate and confidential that it becomes difficult to frame questions on them. It is impossible to put down certain delicate issues in writing.

3.4.4 Schedules

The scheduling methodis one more important method for the study of social problems. This method is in many respects close to the questionnaire method but the major difference between the two is that whereas in the questionnaire method there is none to assist the respondent in filling in the questionnaires. In this method, there is an investigator who assists the informants and gives them necessary clarifications as and when required. Two methods, in many respects, are different in so far as the collection of data is concerned. This method is used for obtaining answers to a set of questions from the respondents or informants. The researcher or the interviewer generally fills it in himself, sits with the informant face to face, and fills up the information supplied by him on the prescribed schedule. A schedule is like a questionnaire, which contains a set of questions.Both questionnaires and schedules are very similar, but they also differ in some respects. These questions are required to be replied to by the respondent with the help of an investigator. The schedule is the name usually applied to a set of questions that are asked and filled in by the investigator in a face-to-face situation with another person.A questionnaire is sent to the respondents by mail, whereas a schedule is used directly in interviews.

Contents of Schedules

The schedule is divided into three parts:

Introductory part: This part of the schedule includes (a) the introductory information about the problem under investigation and the respondent such as the name, serial number,etc of the survey (b) general information about the respondent like address, age, sex, education, income etc. (c) the date, place and time of interview.

The main schedule part: It consists of titles, columns, questions and blank tables that are meant for securinginformation from respondents in respect of the problem under investigation.

Instructions part: In this part, the directions are to be given to the researcherregarding the method of interview.

Types of Schedules

Schedules are of different types, though all the schedules aim to collect data.

Observation Schedule: This is a type of schedule in which questions are put on a specific topic about which the investigator wants to collect data and information. In this case, the respondent can be an individual or group of individuals, and the schedule is filled under specific conditions. Many times, this method is adopted to verify the information already collected.

Rating Schedule: In social research rating schedules are used when information is to be collected about attitudes, opinions, preferences, inhibitions and other like elements and their value is to be assessed and the value of each is required to be measured. These prove very useful when factors that are responsible for measuring a phenomenon are to be measured. Different scales of measurement are to be constructed for evaluation.

Document Schedule: In this schedule whole study is based on certain schedules e.g. studies that deal with the writing of history etc. With the help of these documents, certain questions are asked about the life history of a person, and based on replies received efforts are made to construct life history.

Evaluation Schedule: The nature and complexity of the institution decide the size of the schedule. With the help of this schedule, it becomes both easy as well as possible to study both traditional as well as immediate problems of an institution. These schedules are used to gather information about some institutions or agencies. They help us to study their immediate problems.

Interview Schedule: The interview schedule is used for testing as well as collecting data as well as for collecting supplementary data. The information takes the schedule with him and interviews the respondent and fills in the forms. Usually in this method, the interviewer asks certain standardized questions.

Advantages of Schedule Method

By using this method data can be easily collected by conducting an online surveyvia web, mobiles, emails, google forms, etc. wherea hundred or more respondents can be accommodated fastly.

This method is less expensive when compared to other methods of data collection.

It is easy to analyze and present with different data visualization types and a wide range of data types can be collected such as attitudes, opinions, values, etc.

Enumerators explain the aims and objects of the investigation and also remove the difficulties.

The researcher will solve the non-response issue as the enumerators go to get the information directly.

This method of data collection is very useful in extensive inquiries and can lead to fair reliability.

In extensive studies, it is very useful and can get more reliable data.

Disadvantages of Schedules Method

Many peoplegive wrong answers in online surveys just to receive a promised reward which is known as survey fraud.

Many questions might be left unanswered and participants may not stay fully engaged to the end.

Without someone to explain, participants may have different interpretations of your questions.

This method of data collection cannot fully capture emotions and feelings.

The enumerators should be trained.Enumerators should be intelligent and must possess the capacity for cross-examination to find out the truth.

Train enumerators was a very expensive process and takes a lot of time.

Even in the presence of the researcher, the respondents may not respond to some personal questions.

The reliability of the collected data depended on how the data collection was done.

3.5 SUMMARY

The pattern of business and industry in the present-day environment has become quite complex and involved due to a variety of reasons. Any meaningful decision to be made in this context must be objective and fact-based. This is achieved by collecting and analyzing appropriate data. Data may broadly be divided into two categories, namely primary data and secondary data. The primary data are those which are collected for the first time by the organization which is using them. The secondary data, on the other hand, are those which, have already been collected by some other agency but also can be used by the organization under consideration. Primary data maybe collected by observation, schedules, questionnairesor by interviews.

3.6 GLOSSARY

Data Collection: The process of gathering and analyzing accurate data from various sources to find answers to research problems, trends and probabilities, etc., to evaluate possible outcomes is Known as Data Collection.

Observation: It is an act or instance of noticing or perceiving in the Natural sciences and the acquisition of information from a primary source.

Interview: An interview is generally a qualitative research technique that involves asking open-ended questions to converse with respondents and collect elicit data about a subject.

Schedule: The schedule is the tool or instillment used to collect data from the respondents while the interview is conducted. The schedule contains questions

3.7 LESSON END EXERCISES

Describe the various methods of collecting primary data and comment on their relative advantages and disadvantages.

____

____

____

Differentiate between primary and secondary data.

____

____

____

Examine the merits and limitations of the observation method in collecting data. Illustrate your answer with suitable examples.

____

____

____

What are the guiding considerations in the construction of the questionnaire? Explain.

____

____

____

Briefly define :

Schedules : ____________________________________

Interview : ____________________________________

Questionnaire : __________________________________

3.8 SUGGESTED READINGS

Kothari, C.R., *Research Methodology Methods and Techniques*. Wiley Eastern Limited: New Delhi.

Sadhu, A.N. and A. Singh., *Research Methodology in Social Science*s. Sterling Publishers Private Limited: New Delhi.

William G. Z. *Business Research Methods*. Thomson, India.

DESIGNING OF QUESTIONNAIRE, CHOICE BETWEEN PRIMARY AND SECONDARY DATA

STRUCTURE

4.1 INTRODUCTION

No survey can achieve success without a well-designed questionnaire. Unfortunately, questionnaire design has no theoretical base to guide the marketing researcher in developing a flawless questionnaire. All the researcher has to guide him/her is a lengthy list of dos and don’ts born out of the experience of other researchers past and present. Hence, questionnaire design is more of an art than a science. If you want to find out something about a person, you would usually just ask them. If you want to ask a few questions to a group of people, maybe you would get together as a group and give them a questionnaire.

4.2 OBJECTIVES

After going through this unit, you should be able to describe:-

To design a questionnaire for the collection of primary data;

To understand the attributes of a well-designed questionnaire;

To adopt a framework for developing questionnaires;

To choose between primary and secondary data;

To examine the reliability, suitability and adequacy of primary and secondary data.

4.3 DESIGNINGA QUESTIONNAIRE

The questionnaire is a systematic, data collection technique that consists of a series of questions required to be answered by the respondents to identify their attitude, experience, and behavior towards the subject of research.Questionnaires are a set of written questions designed to gather standardized information about the opinions, preferences, experiences, intentions, and behavior of individuals, and can be devised for a scientific study. Traditionally, they have been said to contrast with surveys in the sense that they do not collect mass data for further analysis, however, the terms are largely used interchangeably these days.

The questionnaire is the heart of the primary data collection technique. Hence; its drafting requires the utmost skill. The questions must be clear, simple and to the point. They must be well organized from the point of view of the respondent and be formulated in such a manner as to provide the data in so far as possible in the desired form. This is especially true of a mail questionnaire which essentially has to speak for itself. If it is not clear; not only the replies may be vague and of little value but many potential respondents may not bother returning the questionnaire at all.Table 1 depicts the different types of questions considered while designing the questionnaire.

4.3.1 Questionnarie Desgin Process

The following steps are involved in the questionnaire design process:

Specify the Information Needed: The first and foremost step in designing the questionnaire is to specify the information needed from the respondents such that the objective of the survey is fulfilled. The researcher must completely review the components of the problem, particularly the hypothesis, research questions, and the information needed.

Define the Target Respondent: At the very outset, the researcher must identify the target respondent from whom the information is to be collected. The questions must be designed keeping in mind the type of respondents under study. Such as, the questions that are appropriate for servicemen might not be appropriate for a businessman. The less diversified respondent group shall be selected because the more diversified the group is, the more difficult it will be to design a single questionnaire that is appropriate for the entire group.

Specify the type of Interviewing Method: The next step is to identify how**therespondentsarereached**. In personal interviews, the respondent is presented with a questionnaire and interacts face-to-face with the interviewer. Thus, lengthy, complex and varied questions can be asked using the personal interview method. In telephone interviews, the respondent is required to give answers to the questions over the telephone. Here the respondent cannot see the questionnaire and hence this method restricts the use of small, simple and precise questions.

Determine the Content of Individual Questions: Once the information needed is specified and the interviewing methods are determined, the next step is to decide the content of the question. The researcher must decide on what should be included in the question such that it contributes to the information needed or serve some specific purpose. In some situations, indirect questions which are not directly related to the information needed may be asked. It is useful to ask neutral questions at the beginning of a questionnaire with the intent to establish the respondent's involvement and rapport. This is mainly done when the subject of a questionnaire is sensitive or controversial. The researcher must try to avoid the use of**double-barreled questions**.

Overcome Respondent's Inability and Unwillingness to Answer: The researcher should not presume that the respondent can provide accurate responses to all the questions. He must attempt to overcome the respondent's inability to answer. The questions must be designed in a simple and easy language such that it is easily understood by each respondent. In situations, where the respondent is not at all informed about the topic of interest, then the researcher may ask the **filter questions**, an initial question asked in the questionnaire to identify the prospective respondents to ensure that they fulfill the requirements of the sample.

Decide on the Question Structure: The researcher must decide on the structure of questions to be included in the questionnaire. The question can be structured or unstructured. The **unstructured questions are the open-ended questions**that are answered by the respondents in their own words. These questions are also called **free-response** or **free-answer questions**. Whereas, **structured questions are called closed-ended questions**that pre-specify the response alternatives. These questions could be a multiple-choice question, dichotomous (yes or no), or a scale.

Determine the Question Wording: The desired question content and structure must be translated into **words that are easily understood** by the respondents. At this step, the researcher must translate the questions into easy words such that the information received from the respondents is similar to what was intended.

Determine the Order of Questions: At this step, the researcher must decide the **sequence in which the questions are to be asked**. The opening questions are crucial in establishing the respondent's involvement and rapport, and therefore, these questions must be interesting, non-threatening and easy. Usually, **open-ended questions** which ask respondents for their opinions are considered good opening questions, because people like to express their opinions.

Identify the Form and Layout: The **format, positioning and spacing** of questions have a significant effect on the results. The layout of a questionnaire is specifically important for self-administered questionnaires. The questionnaires must be divided into several parts, and each part shall be numbered accurately to clearly define the branches of a question.

Reproduction of Questionnaire: Here, we talk about the**appearance of the questionnaire,** i.e. the quality of the paper on which the questionnaire is either written or printed. In case, the questionnaire is reproduced on a poor-quality paper; then the respondent might feel the research is unimportant due to which the quality of response gets adversely affected. Thus, it is recommended to reproduce the questionnaire on a good-quality paper having a professional appearance. In case, the questionnaire has several pages, then it should be presented in the form of a booklet rather than the sheets clipped or stapled together.

Pretesting: Pretesting means **testing the questionnaires on a few selected respondents** or a small sample of actual respondents to improve the questionnaire by identifying and eliminating the potential problems. All the aspects of the questionnaire must be tested such as question content, structure, wording, sequence, form and layout, instructions, and question difficulty. The researcher must ensure that the respondents in the pretest should be similar to those who are to be finally surveyed.

4.3.2 Main Aspects of a Questionnaire

The questionnaire is considered the heart of a surveyoperation. Hence it should be very carefully constructed. If it is not properly set up, then the surveyis bound to fail. This fact requires us to study the main aspects of a questionnaire viz., the generalform, question sequence and question formulation and wording.

General form: So far as the general form of a questionnaire is concerned, it can either be a structured or unstructured questionnaire. Structured questionnaires are those questionnaires in which there are definite, concrete and pre-determined questions. The questions are presented with the same wording and in the same order to all respondents. The resort is taken to this sort of standardization to ensure that all respondents reply to the same set of questions. The form of the question may be either closed (i.e., of the type 'yes' or 'no') or open (i.e., inviting free response) but should be stated in advance and not constructed during questioning. When these characteristics are not present in a questionnaire, it can be termed an unstructured or non-structured questionnaire.

Question sequence: To make the questionnaire effective and to ensure the quality of the replies received, a researcher should pay attention to the question sequence in preparing the questionnaire. A proper sequence of questions reduces considerably the chances of individual questions being misunderstood. The questionsequence must be clear and smoothly moving, meaning thereby that the relation of one question to another should be readily apparent to the respondent, with questions that are easiest to answer being put in the beginning. The first few questions are particularly important because they are likely to influence the attitude of the respondent and in seeking his desired cooperation. The opening questions should be such as to arouse human interest.

Question formulation and wording: About this aspect of the questionnaire, the researcher should note that each question must be very clear for any sort of misunderstanding can do irreparable harm to a survey. The question should also be impartial in order not to give a biased picture of the true state of affairs. Questions should be constructed with a view to their forming a logical part of a well-thought-out tabulation plan. In general, all questions should meet the following standards (a) should be easily understood; (b) should be simple i.e., should convey only one thought at a time; (c) should be concrete and should conform as much as possible to the respondent's way of thinking.

4.3.3 Essentials of a Good Questionnaire

To be successful, the questionnaire should be comparatively short and simple i.e., the size of thequestionnaire should be kept to the minimum.

Questions should proceed in logical sequence moving from easy to more difficult

1a				
Closed-ended questions Did you experience cough and cold in the past 6 months?				
? Yes		? No		
Think of the last time you had a cough and cold. Did you consult a doctor?				
? Yes		? No		
Open-ended questions				
What is/are the main reason(s) for you to consult a doctor for your cough and cold? (You can list down more than one reason)				
____				
1b. Filtering				
Think of the last time you had a cough and cold. Did you consult a doctor?				
? Yes		? No		
1c. Sensitive questions				
It is common for people to ask for antibiotics when they have a cough and cold. Did you request an antibiotic from the doctor for your cough and cold?				
? Yes		? No		
1d. Likert scale				
We should always consult a doctor for cough and cold. (Please circle the answer)				
1	2	3	4	5
Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1e. "Double-barreled" questions				
Did you experience sore throat and fever during your cough and cold?				
? Yes		? No		
Should be: Did you experience a sore throat during your cough and cold? Did you experience fever during your cough and cold?				
Or: Did you experience the following symptoms during your cough and cold? (You can tick more than one box)				
? Sore throat	? Cough		? Fever	
1f. Ambiguous Questions				
How often do you visit the doctor for cough and cold?				
? Not at all	? Not very often	? Sometimes	? Quite often	? Very often
Change to: How many times did you visit the doctor for cough and cold in the past 1 year?				
_________ times in the past 1 year				

Table 4.1 : Sample Questionnaire

questions.

Personal and intimate questions should be left to the end. Technical terms and vagueexpressions capable of different interpretations should be avoided in a questionnaire.

Questions may be dichotomous (yes or noanswers), multiple choice (alternative answerslisted), or open-ended. The latter type of question is often difficult to analyze and henceshould be avoided in a questionnaire to the extent possible.

Thereshould be some control questions in the questionnaire which indicate the reliability ofthe respondent.

4.4 CHOICE BETWEEN PRIMARYAND SECONDARY DATA

Primary and Secondary Date

The primary dataare those which are collected afresh and for the first time, and thus happen to be original. Such data are published by authorities who themselves are responsible for their collection. Whereas the secondary data,on the other hand, are those which have already been collected by some other agency and which have already been processed. Generally speaking, secondary data is information that has previously been collected by some organization to satisfy its own need but is being used by the department under reference for an entirely different reason.

Difference between Primary data and Secondary data

Primary data is original whereas secondary data is not original

The collection of data is expensive in the case of primary data whereas the collection of secondary data isless expensive.

Primary data is in the shape of raw materials whereas secondary data is the shape of finishedproducts.

Primary data is adequate and suitable whereas secondary data need not be ample and apposite.

Editing of Primary Data

Editinginvolves reviewing the data collected by investigators to ensure maximum accuracy and unambiguity. It should be done as soon as possible after the data have been collected. If the size of the data is relatively small, only one person should edit all the data for the entire study. The different steps of editing are indicated below:

Checking legibility: Obviously, the data must be legible to be used. If a response is not presented clearly, the concerned investigator should be asked to rewrite it.

Checking completeness: An omitted entry on a fully structured questionnaire may mean that no attempt was made to collect data from the respondent or that the investigator simply did not record the data. If the investigator did not record the data, prompt editing and questioning of the investigator may provide the missing item. If an entry is missing because of the first possible cause, there is not much that can be done, except to make another attempt to get the missing data. This requires knowing why the entry is missing.

Checking consistency: The editor should examine each questionnaire to check inconsistency or inaccuracy if any, in the statement. The income and expenditure figures may be unduly inconsistent. The age and the date of birth may disagree. The area of an agricultural plot may be unduly large. The concerned investigators should be asked to make the necessary corrections. If there is any repetitive response pattern in the reports of individual investigators they may represent investigator bias or perhaps attempted dishonesty.

Collection of Secondary Data

Secondary dataare data that have been collected and analyzed by some other agency. The sources of secondary data could be:

Various publications of Central, State and local governments: The important official publications are Statistical Abstract, India-Annual; Monthly Abstract of Statistics (both published by Central Statistical Organization); Indian Agricultural Statistics (Annual) (Published by Ministry of Food and Agriculture); Index

Number of Wholesale Prices in India (Weekly) (Published by Ministry of Commerce and Industry); Reserve Bank of India Bulletin (Monthly) (Published by Reserve Bank of India).

Various publications of foreign governments or international bodies: The important publications are publications of international bodies like UNO, FAO, WHO, UNESCO, ILO, Statistical Year Book (Published by the Statistical Office of the United Nations), Yearbook of Labour Statistics (Published by ILO, Geneva). The secondary data provided by such publications are authentic, but along with other things, one must be especially careful about the units in respect of currency, weight, etc. which greatly vary from one country to another.

Journals of trade, commerce, economics, and engineering published by responsible trade associations, Chambers of Commerce provide secondary data in respect of some important items. Some examples of this kind of publication are "Annual Report of the Chief Inspector of Mines in India" (issued annually by the office of the Chief Inspector of Mines, Dhanbad) and "Indian Textile Bulletin (issued monthly by the Textile Commissioner, Bombay).

The other sources of secondary data are books, magazines and newspapers, reports prepared by various universities, historical documents, diaries, letters, unpublished biographies and autobiographies.

Scrutiny of Secondary Data

Primary data are to be scrutinized after the questionnaires are completed by the interviewers. Likewise, the secondary data are to be scrutinized before they are compiled from the source. The scrutiny should be made to assess the suitability, reliability, adequacy and accuracyof the data to be compiled and to be used for the proposed study.

Suitability: The compiler should satisfy himself that the data contained in the publication will be suitable for his study. In particular, the conformity of the definitions, units of measurement and time frame should be checked. For example, one US gallon is different from one British gallon.

Reliability: The reliability of the secondary data can be ascertained from the collecting agency, mode of collection and the period of collection. For instance, secondary data collected by a voluntary agency with unskilled investigators are unlikely to be reliable.

Adequacy: The source of data may be suitable and reliable but the data may not be, adequate for the proposed inquiry. The original data may cover a bigger or narrower geographical region or the data may not cover suitable periods. For instance, the per capita income of Pakistan before 1971 is inadequate for reference during the subsequent periods as it became separated into two different countries with-considerable variation in the standard of living.

Accuracy: The user must be satisfied with the accuracy of the secondary data. The process of collecting raw data, the reproduction of processed data in the publication, and the degree of accuracy desired and achieved should also be satisfactory and acceptable to the researcher.

4.5 SUMMARY

Questionnaires may be used for data collection by interviewers. They may also be mailed to prospective respondents. The drafting of a good questionnaire requires the utmost skill. The process of interviewing also requires a great deal of tact, patience and competence to establish rapport with the respondent. As we have already seen,

there are a lot of differences in the methods of collecting Primary and Secondary data. Primary data which is to be collected originally involves an entire scheme of the plan starting with the definitions of various terms used, units to be employed, type of inquiry to be conducted, the extent of accuracy aimed at etc. For the collection of secondary data, a mere compilation of the existing data would be sufficient. A proper choice between the type of data needed for any particular statistical investigation is to be made after taking into consideration the nature, objective and scope of the inquiry; the time and the finances at the disposal of the agency; the degree of precision aimed at and the status of the agency.

4.6 GLOSSARY

Questionnaire: A questionnaire is a research tool featuring a series of questions used to collect useful information from respondents. These instruments include either written or oral questions and comprise an interview-style format.

Primary Data: These are the data that are collected for the first time by an investigator for a specific purpose. Primary data are 'pure' in the sense that no statistical operations have been performed on them and they are original.

Secondary Data: Secondary data refers to data that is collected by someone other than the primary user.

4.7 LESSON END EXERCISES

Define the main steps involved in designing a questionnaire.

____

____

____

Describe the various methods of collecting primary data and comment on their relative advantages and disadvantages.

____

____

____

Define secondary data. State their chief sources and point out the dangers involved in their use and the precautions necessary to use them. Illustrate with examples.

INTRODUCTION TO BUSINESS RESEARCH

UNIT-I **Lesson No. 5**

SCALING TECHNIQUES IN BUSINESS RESEARCH AND PRECAUTIONS IN BUSINESS RESEARCH

STRUCTURE

5.1 INTRODUCTION

As we discussed earlier, the data consists of quantitative variables like price, income, sales, etc., and qualitative variables like knowledge, performance, character, etc. The qualitative information must be converted into a numerical form for further analysis. This is possible through measurement and scaling techniques. A common feature of survey-based research is to have respondents' feelings, attitudes, opinions, etc. in some measurable form. For example, a bank manager may be interested in knowing the opinion of the

Customers about the services provided by the bank. As the researcher,he may be interested in knowing the attitude of the people towards the services provided by them. In this chapter, we will discuss the issues related to measurement, different levels of measurement scales, various types of scaling techniques and also the selection of an appropriate scaling technique. It is the common feature of marketing research is the attempt to have respondents communicate their feelings, attitudes, opinions, and evaluations in some measurable form. To this end, marketing researchers have developed a range of scales. Each of these has unique properties. What is important for the marketing analyst to realize is that they have widely differing measurement properties. Some scales are at very best, limited in their mathematical properties to the extent that they can only establish an association between variables. Other scales have more extensive mathematical properties and some, hold out the possibility of establishing cause and effect relationships between variables.

5.2 OBJECTIVES

After going through this unit, you should be able to describe:-

To explain the concepts of measurement and scaling;

An understanding of the four levels of measurement that can be taken by researchers;

To classify and discuss different scaling techniques;

The ability to distinguish between comparative and non-comparative measurement scales;

To learn about the precautions considered while conducting the research.

5.3 SCALING TECHNIQUES IN BUSINESS RESEARCH

Scaling is the procedure of measuring and assigning the objects to the numbers according to the specified rules. In other words, the process of locating the measured objects on the continuum, a continuous sequence of numbers to which the objects are assigned is called scaling. This technique is considered an extension of measurement. Measurement is defined as the process of assigning numbers or symbols to the characteristics of the object as per the specified rules. Here, the researcher assigns numbers, not to the object, but to its characteristics such as perceptions, attitudes, preferences, and other relevant traits. In other words, the scaling technique is a method of placing respondents in continuation of gradual change in the pre-assigned values, symbols, or numbers based on the features of a particular object as per the defined rules. All the scaling techniques are based on four pillars, i.e., order, description, distance and origin. The marketing research is highly dependable upon the scaling techniques, without which no market analysis can be performed.

5.3.1 Types of Scaling Techniques

There are many types of scales and scaling techniques. We have to choose the most appropriate technique to suit the research being done. The statements contained therein have to be prepared in such a way that the answers received can easily be converted into numerical values. The researchers have identified many scaling techniques; today, we will discuss some of themost common scales used by business organizations, researchers, economists, experts, etc.These techniques can be classified as primary scaling techniques and other scalingtechniques.

Primary Scaling Techniques

Scaling is an important technique and the type of scale used in taking measurements directly impinges on the statistical techniques which can legitimately be used in the analysis. Typically, there are four levels of measurement scales or methods of assigning numbers in the category of primary scaling technique as represented in figure 5.1: Nominal scale, Ordinal scale, Interval scale, and Ratio scale. These major four scales used in statistics for market research are described as follows:

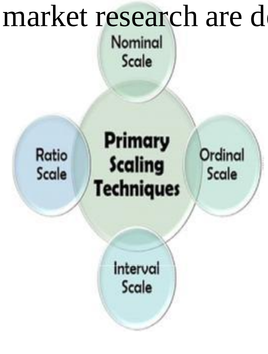


Figure 5.1: Primary Scaling Techniques

Nominal Scale: Nominal scales are adopted for non-quantitative labeling variables that are unique and different from one another. It is the crudest of measurement scales, that classifies individuals, companies, products, brands, or other entities into categories where no order is implied. That is why it is also referred to as a categorical scale. It is a system of classification and does not place the entity along a continuum. It involves a simple count of the frequency of the cases assigned to the various categories, and if desired numbers can be nominally assigned to label each category. The nominal scale simply allows the categorization of responses into several mutually exclusive categories. There are no relationships between the categories, implying that there is no ranking or order. The typical applications of the nominal scale are in the classification of responses by social class, "like" or "dislike", "yes" or ''no", and so on. The statistical operation possible for nominally scaled data is counting only.

An example of a nominal scale is to classify the respondent's income into three groups: the highest income as group 1. The middle income as group 2, and the low-income as group

The assigned numbers have no arithmetic properties and act only as labels. The only statistical operation that can be performed on nominal scales is a frequency count. We cannot determine an average except mode.

Ordinal Scale: Ordinal scales involve the ranking of individuals, attitudes, or items along the continuum of the characteristic being scaled. This scale functions on the concept of the relative position of the objects or labels based on the individual's choice or preference. It allows the respondents to rank some alternatives by some common variable. The main characteristic of the ordinal scale is that the categories have a logical or ordered relationship. This type of scale permits the measurement of degrees of difference, (that is, 'more' or 'less') but not the specific amount of differences (that is, how much 'more' or 'less'). This scale is very common in marketing, satisfaction and attitudinal research.By using this technique it is feasible for a user of the product to rank the brands from the best to the worst. However, the amount of difference between the ranks cannot be found. All of the information a nominal scale would have given is available from an ordinal scale. In addition, positional statistics such as the median, quartile and percentile can be determined.

For example, while doing online shopping, every product has a customer review section where the buyersrate the listed product according to their buying experience, product features, quality, usage,etc.The ratings so provided are as follows:5 Star – Excellent, 4 Star – Good, 3 Star – Average, 2 Star – Poor and 1 Star – Worst.

Interval Scale: An interval scale is also called a cardinal scale in which the numerical labeling with the same difference among the consecutive measurement units. With the help of this scaling technique, researchers can obtain a better comparison between the objects. The deficiencies of the nominal and the ordinal scales are taken care of in the interval scale. The scale has an arbitrary zero point with numbers placed at equally appearing intervals. This of course has implications for the type of data manipulation and analysis we can carry out on data collected in this form. Many statistical operations including addition, subtraction, and computation of the mean can be done on interval scaled data. It is the only scale with interval-scaled data that researchers can justify the use of the arithmetic mean as the measure of average.

For example, A survey was conducted by an automobile company to know the number of vehiclesowned by the people living in a particular area who can be its prospective customers in the future. It adopted the interval scaling technique for the purpose and provided the units as 1, 2,3, 4, 5, 6 to select from.In the scale mentioned above, every unit has the same difference, i.e., 1, whether it is between2 and 3 or between 4 and 5.Temperature is interval scaled, being measured either in Centigrade or Fahrenheit. We cannot speak of 50°F being twice as hot as 25°F since the corresponding temperatures on the centigrade scale, 10°C and -3.9°C, are not in the ratio of 2:1.

Ratio Scale: The ratio scaling techniqueis one of the most superior measurement techniques. Similar to an intervalscale, a ratio scale is an abstract number system. It allows measurement at proper intervals,order, categorization and distance, with an added property of originating from a fixed zeropoint. Here, the comparison can be made in terms of the acquired ratio.The highest level of measurement is a ratio scale. This has the properties of an interval scale together with a fixed origin or zero point. Ratio scales permit the researcher to compare both differences in scores and the relative magnitude of scores.

However, for most behavioural research, interval scales are typically the highest form of measurement. Most statistical data analysis procedures do not distinguish between the interval and ratio properties of the measurement scales and it is sufficient to say that all the statistical operations that can be performed on an interval scale can also be performed on ratio scales.

Examples of variables that are ratio scaled include weights, lengths and times. For instance, the difference between 5 and 10 minutes is the same as that between 10 and 15 minutes, and 10 minutes is twice as long as 5 minutes.

Other Scaling Techniques

Scaling of objects can be used for a comparative study between more than one object (products, services, brands, events, etc.). Or can be individually carried out to understand the consumer's behaviour and response towards a particular object. This type of scale used in marketing research falls into two broad categories: comparative and noncomparative. In comparative scaling, the respondent is asked to compare one brand or product against another. With noncomparative scaling, respondents need only evaluate a single product or brand. Their evaluation is independent of the other product and/or brands that the marketing researcher is studying. Noncomparative scaling is frequently referred to as monadic scaling and this is the more widely used type of scale in commercial marketing research studies. Following are the two categories represented in figure 5.2 with other scaling techniques placed based on their comparability.

Comparative Scales

For comparing two or more variables, a comparative scale is used by the respondents. The comparative scales can further be divided into the following four types of scaling techniques: Paired Comparison Scale, Rank Order Scale, Constant Sum Scale, and Q-sort Scale. Following are the different types of comparative scaling techniques:

Paired Comparison: It is sometimes the case that marketing researchers wish to find out which are the most important factors in determining the demand for a product. Conversely, they may wish to know which are the most important factors acting to prevent the widespread adoption of a product. A paired comparison symbolizes two variables from which the respondent needs to select one. This technique is mainly used at the time of product testing, to facilitate the consumers with a comparative analysis of the two major products in the market. The data obtained are ordinal.To compare more than two objects firstly, the first object is compared with the second one and then the superior one (i.e., one with a higher percentage) is compared with the third one.

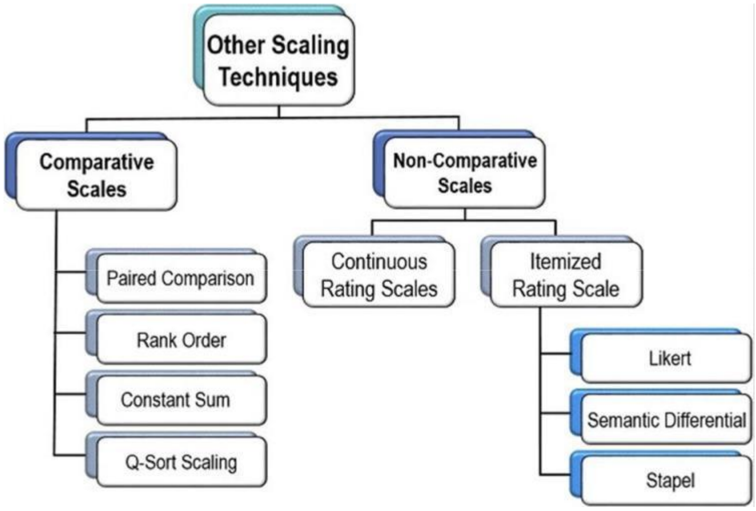


Figure 5.2 : Other Scaling Techniques

For example, A market survey was conducted to find out consumers' preferences for the network service provider brands, A and B. The outcome of the survey was as follows:

Brand 'A' = 57%

Brand 'B' = 43%

Thus, it is visible that the consumers prefer brand 'A', over brand 'B'.

Rank Order Scale: This is another type of comparative scaling technique in which respondents are presented with several items simultaneously and asked to rank them in the order of priority. This is an ordinal scale that describes the favoured and unfavoured objects but does not reveal the distance between the objects. In rank order scaling the respondent needs to rank or arrange the given objects according to their preference. Data generated using this technique are employed with conjoint analysis because of the discriminatory potential of the scaling, stimulating the consumers to discriminate one brand from the other. Under the assumptions of transitivity, rank order can be converted to equivalent paired comparison data, and vice versa.

For example, A soap manufacturing company conducted a rank order scaling to find out theorderly preference of the consumers. It asked the respondents to rank the following brands inthe sequence of their choice:

SOAP BRANDS	RANK
Brand V	4
Brand X	2
Brand Y	1
Brand Z	3

Table 5.1: Rank Order Scale

The above scaling shows that soap of brand Y is the most preferred brand, followed by brand X, then brand Z and the least preferred one is brand V.

Constant Sum Scaling: This technique allows the respondents to allocate a constant sum of units, such as points, rupees or among a set of stimulus objects to some criterion. The attributes are scaled by counting the points assigned to each one by all the respondents and dividing the number of respondents. This predominantly uses ordinal because of its comparative nature and the resulting lack of generalisability. Constant sum scaling has an advantage in that it allows for discrimination among stimulus objects without requiring toomuch time. Its advantage involves the allocation of more or fewer units than those specified.This technique where a continual sum of units like dollars, points, chits, chips, etc. isgiven to the features, attributes and importance of a particular product or service by therespondents.The technique involves askingthe respondents to assign 10 points to attributes of a sports utility vehicle. If the attribute isunimportant, then the respondents would want to enter zero.

For example, The respondents belonging to 3 different segments were asked to allocate 50 points to the following attributes of any product ‘P’:

ATTRIBUTES	SEGMENT 1	SEGMENT 2	SEGMENT 3
Finish	11	8	9
Skin Friendly	11	12	12

Fragrance	7	11	8
Packaging	9	8	10
Price	12	11	11

Table 5.2 : Constant Sum Scale

From the above constant sum scaling analysis, we can see that:Segment 1 considers product 'P' due to its competitive price as a major factor. But segment 2 and segment 3 prefer the product because it is skin-friendly.

Q-Sort Scaling: This is a comparative scale that uses a rank order procedure to sort objects based on similarity to some criterion. The important characteristic of this methodology is that it is more important to make comparisons among different responses of a respondent than the responses between different respondents. In this method, the respondent is given statements in a large number for describing the characteristics of a product or a large number of brands of a product.This technique is more suitable in the case where the number of objects is not less than 60 and more than 140,the most appropriate of all ranging between 60 to 90 responses that result in a normal distribution. This method is faster and less tedious than paired comparison measures. It also forces the subject to conform to quotas at each point of scale to yield a quasi-normal distribution. The utility of Qsort in marketing research is to derive clusters of individuals who display similar preferences, thus representing unique market segments.

For example, you may wish to determine the preference from among a large number of magazines. The following format shown in the table below may be given to a respondent to obtain the preferences.

Instructions: The bag given to you contain pictures from 90 magazines. Please choose 10 magazines you 'prefer most', 20 magazines you 'like', 30 magazines to which you are 'neutral (neither like nor dislike)', 20 magazines you 'dislike', and 10 magazines you 'prefer least'. Please list the sorted magazine names in the respective columns of the form provided to you.

Format:

Prefer Most	Like	Neutral	Dislike	Prefer Least
10	20	30	20	10

Table 5.3: Q-Sort Scalling

Here, in the above table 10, 20, 30, 20, and 10 are the number of magazine names specified in each category.

Noncomparative Scales

A non-comparative scale is used to analyze the performance of an individual product orobject on different parameters. These scales or otherwise called nomadic scales because only one object is evaluated at a time. Researchers use this scale allowing respondents to employ whatever rating standard seems appropriate to them and not specified by the researcher. The respondents do not compare the object being rated either to another object or to some specified standard set by the researcher. Non-comparative techniques use continuous and itemized rating scales. In such scales, each object is scaled independently of the other objects in the stimulus set, the resulting data is generally assumed to be interval or ratio scale.

Continuous Rating Scale: In a continuous rating scale, the respondentsrate the objects by placing a mark at the appropriate position on a continuousline that runs from one extreme of the criterion variable to the other.This is also otherwise called a graphic rating scale. This is a type of scale that offersrespondents a form of the continuum (such as a line) on which to provide a rating of an object.Researchers develop a continuous rating scale allowing the respondents to indicate their ratingby placing a mark at the appropriate point on a line that runs from one end of the criterionvariable to the other or a set of predetermined response categories. Here the respondents neednot select marks already set by the researcher. Several variations are possible. The line may be vertical or horizontal; it may beunmarked or marked; if marked, the divisions may be few or as many as in the thermometerscale; the scale points may be in the form of numbers or brief descriptions.

For example, A mattress manufacturing company used a continuous rating scale to find outthe level of customer satisfaction for its new comfy bedding. The response can be taken in thefollowing different ways.

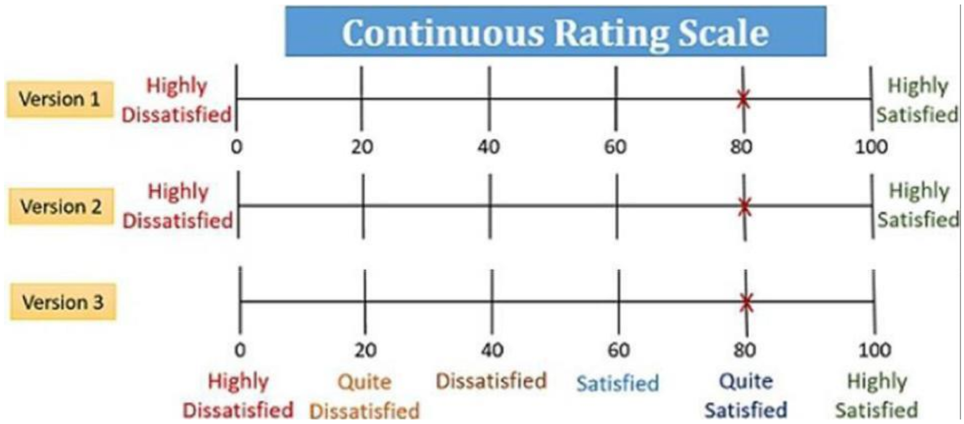


Figure 5.3: Continous Rating Scale

The above diagram shows a non-comparative analysis of one particular product, i.e. comfybedding. Thus, making it very clear that the customers are quite satisfied with the product andits features.

Itemized Rating Scales: Itemized scale is another essential technique under the non-comparative scales. It emphasizeschoosing a particular category among the various given categories by the respondents.Each class is briefly defined by the researchers to facilitate such selection.With an itemized scale, respondents are provided with a scale having numbers and/or brief descriptions associated with each category and are asked to select one of the limited numbers of categories, ordered in terms of scale position, that best describes the product, brand, company or product attribute being studied.Itemized rating scales are widely used in marketing research.

This scale is similar to the graphic scale in that the individuals make their judgments independently, without the benefit of direct comparison. The respondents are provided with ascale that has a number or brief description associated with each category. This scale allowsthe respondents to choose from a more limited number of categories, usually five to seven,although 10 or more are rarely used. The categories are ordered in terms of scale positionand the respondents are required to select the specified category that best describes the objectbeing rated. The categories are given a verbal description, although this is not necessary. These scales are widely used in research and nowadays, more complex types suchas multi-item rating scales are used. There are few variants among itemized rating scales.They are Likert, Semantic differential and Stapel scales.

Itemized Graphic Scale ***Itemized Verbel Scale*** ***Itemized Numeric Scale***

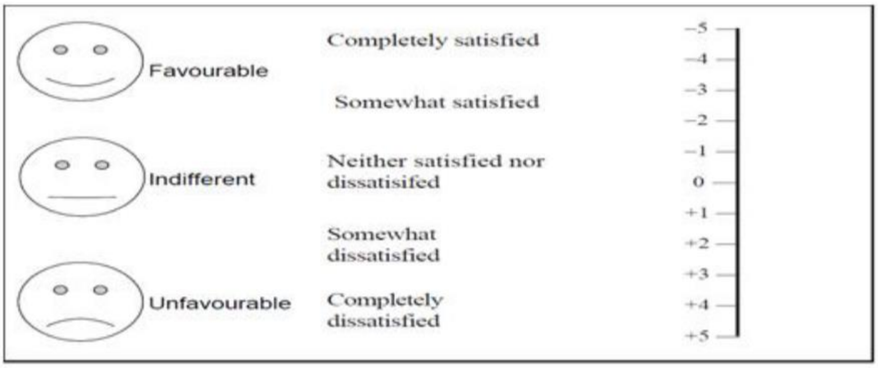


Figure 5.4 : Itemized Rating Scales

Likert scales: A Likert scale is what is termed a summated instrument scale. This means that the items making up a Liken scale are summed to produce a total score. A Likert scale is a composite of itemized scales. In business research, the Likert scale, developed by RensisLikert, is extremely popular for measuring attitudes, because, the method issimple to administer. With the Likert scale, the respondents indicate their attitudes by checking how strongly they agree or disagree with carefully wordedstatements that range from very positive to very negative towards the attitudinalobject. Respondents generally choose from five alternatives (say stronglyagree, agree, neither agree nor disagree, disagree, strongly disagree).

In the Likert scale, the researcher provides some statements and asks therespondents to mark their level of agreement or disagreement over these statements byselecting any one of the options from the five given alternatives.Likert scales are treated as yielding Interval data by the majority of marketing researchers.

For example, A shoe manufacturing company adopted the Likert scale technique for its newsports shoe range named Z sports shoes. The purpose is to know the agreement ordisagreement of the respondents.For this, the researcher asked the respondents to circle a number representing the mostsuitable answer according to them, in the following representation:

Statement	**Strongly Agree**	**Agree**	**Neither**	**Disagree**	**Strongly Disagree**
Z sports shoes are very lightweight	1	2	3	4	5
Z sports shoes are extremely comfortable	1	2	3	4	5
Z sports shoes look too trendy	1	2	3	4	5
I will recommend Z sports shoes to friends, family and colleagues	1	2	3	4	5
The price of Z shoes is economical	1	2	3	4	5

Table 5.4 : The Likert Scale

Statement	**Strongly Agree**	**Agree**	**Neither**	**Disagree**	**Strongly Disagree**
If the price of raw materials fell firms would reduce the price of their food products.	1	2	3	4	5
Without government regulation, the firms would exploit the consumer.	1	2	3	4	5
Most food companies are so concerned about making profits they do not care about quality.	1	2	3	4	5
The food industry spends a great deal of money making sure that its manufacturing is hygienic.	1	2	3	4	5
Food companies should charge the same price for their products throughout the country	1	2	3	4	5

Table 5.5 : The Likert Scale

The above examples will help the company to understand what the customers think about itsproducts. Also, whether there is any need for improvement or not.

Semantic Differential Scale: Semantic differential scale is a popular scaling technique next to the Likert scale. In this scale, therespondents associate their responses with bipolar labels that have semantic meaning. Therespondent's rate objects on several itemized, seven-point rating scales bounded at eachend by one of two bipolar adjectives such as "Excellent" and "Very bad". The respondentsindicate their response by choosing the one that best describes their choice.This scale is used for a variety ofpurposes. It can be used to find whether a respondent has a positive or negativeattitude towards an object. It has been widely used in comparing brands,products and company images. It has also been used to develop advertising andpromotion strategies and in a new product development study.The points are marked either from - 3 to +3 through a zero or from 1 to 7. The middle value maybe treated as a neutral position. The value zero in the first type is the neutral point and 4 inthe second type is the neutral point. The resulting data are commonly analyzed through profileanalysis. In such analysis, the means or median values on each rating scale are calculated andcompared by plotting or statistical analysis. This would help the researcher to determine theoverall differences and similarities among the objects.

For example, A well-known brand for watches, carried out semantic differential scaling tounderstand the customer's attitude towards its product. The pictorial representation of thistechnique is as follows:

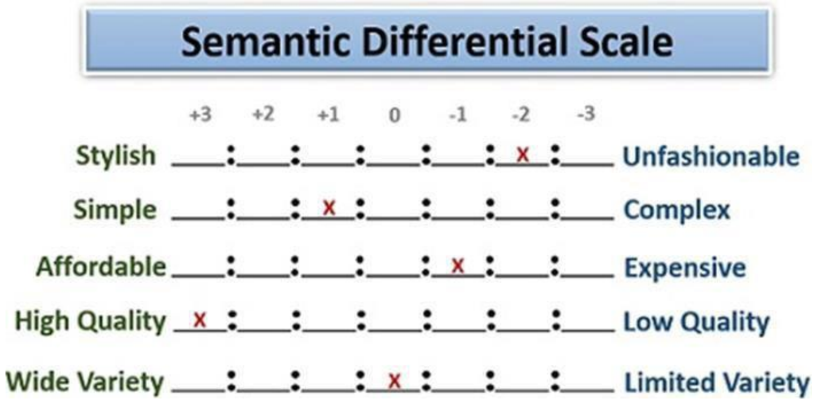


Figure 5.5 : Semantic Differential Scale

From the above diagram, we can analyze that the customer finds the product of superiorquality; however, the brand needs to focus more on the styling of its watches.

Stapel Scale: The Stapel scale was originally developed to measure thedirection and intensity of an attitude simultaneously. Modern versions of theStapel scale place a single adjective as a substitute for the Semantic differentialwhen it is difficult to create pairs of bipolar adjectives. This scale is named after Jan Stapel, who developed it. This is a unipolar rating scale withingeneral 10 categories number from -5 to +5 without a neutral point (zero). This scale isusually presented vertically and respondents choose their response based on how accuratelyor inaccurately each item describes the object by selecting an appropriate numerical responsecategory. A higher number indicates a more accurate description of the object and a lowernumber indicates a lower description of the object.The modified Stapelscale places a single adjective in the center of an even number of numericalvalues (say, +3, +2, +1, 0, –1, –2, –3). This scale measures how close to or howdistant from the adjective a given stimulus is perceived to be.

For example, A tours and travel company asked the respondent to rank their holiday packagein terms of value for money and user-friendly interface as follows:

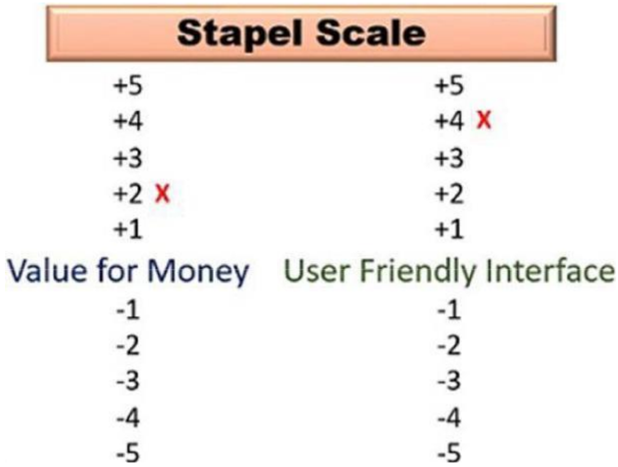


Figure 5.6 : Stapel Scale

With the help of the above scale, we can say that the company needs to improve its packagein terms of value for money. However, the decisive point is that the interface is quite userfriendlyfor the customers.

5.4 PRECAUTIONS IN BUSINESS RESEARCH

A researchreport forms a channel of communication between the research findings given by the researcher on the one hand and the readers of the report on the other hand and effective communication of research findings to the readers is the basic task of the research report. The research report must be prepared by keeping the following precautions in view:

Length of the report: While determining the length of the report one should keep in view the fact that it should be long enough to cover the subject but short enough to maintain interest. Report writing should not be a means to learning more and more about less and less.

Interesting: A research report should not if this can be availed, be dull, it should be such as to sustain the reader's interest.

Use of abstract terminology and Jargon: Abstract terminology and technical Jargon should be avoided in the research report. The report should be able to convey the matter as simply as possible. This, in other words, means that report should be written in an objective style in simple language, avoiding, an expression such as, "it seems", "There may be" and the like.

Presentation of the findings: Readers are often interested in acquiring a quick knowledge of the main findings and as such the report must provide ready availability of the findings. for this purpose, charts, graphs and statistical tables may be used for the various results in the main report in addition to the summary of important findings.

Presentation of the report: The layout of the report should be well thought out and must be appropriate and according tothe objective of the research problems.

Writing of the Report: The reports should be free from grammatical mistakes and must be prepared strictly following the techniques of composition footnotes, documentation, proper punctuation and use of abbreviations in footnotes.

Logical presentation of the report: The report must present the logical analysis of the subject matter. it must reflect a structure wherein the different pieces of analysis relating to the research problem fit well.

Originality in writing report: A research report should show originality and should necessarily be an attempt to solve some intellectual problem. it must contribute to the solution of a problem and must add to the store of knowledge.

Plan for future research and implications: Towards the end, the report must also state the policy implications relating to the problem under consideration. It is usually considered desirable if the report makes a forecast of the probable future of the subject concerned and indicates the kinds of research still needs to be done in that particular field.

Appendices: Appendices should be enlisted in respect of all the technical data in the report.

Bibliography:A bibliography of sources consulted is a must for a good report and must necessarily be given.

Index: The index is also considered an essential part of a good report and as such must be prepared and appended at the end.

Appearance: Report must be attractive, neat and clean, whether typed or printed.

Stating confidence limits: Calculated confidence limits must be mentioned and the various constraints experienced in conducting the research study may also be stated in the report.

Introduction: The objective of the study, the nature of the problem, the methods employed and the analysis techniques adopted must all be clearly stated at the beginning of the report in the form of an introduction.

5.5 SUMMARY

Scaling techniques provide a clear picture of the product life cycle and the market acceptability of the products offered. It facilitates product development and benchmarking through rigorous market research. There are four levels of measurement: nominal, ordinal, interval and ratio. These constitute a hierarchy where the lowest scale of measurement, nominal, has far fewer mathematical properties than those further up this hierarchy of scales. Nominal scales yield data on categories; ordinal scales give sequences; interval scales begin to reveal the magnitude between points on the scale and ratio scales explain both order and the absolute distance between any two points on the scale. The other measurement scales, commonly used in marketing research, can be divided into two groups; comparative and non-comparative scales. Comparative scales involve the respondent in signaling where there is a difference between two or more producers, services, brands or other stimuli. Examples of such scales include; paired comparison, dollar metric, unity-sum-gain and line marking scales. Non-comparative scales, described in the textbook, are; continuous rating scales, itemized rating scales, semantic differential scales and Likert scales.Last but not least the steps defined were considered precautionary steps in business research.

5.6 GLOSSARY

Scaling: Scaling is the branch of measurement that involves the construction of an instrument that associates qualitative constructs with quantitative metric units.

Primary Scaling: A scale is the measuring stick by which study subjects are evaluated.

Comparative Scaling: It measures how respondents interact with items and then infers differences between items from responses to obtain scale values.

Non-Comparative: Noncomparative scaling techniques are a set of scaling methods in which each item is scaled independently of each of the others.

5.7 LESSON END EXERCISE

What is the more common name given to ordinal scales?

____

____

____

Briefly define the types of primary scaling techniques.

____

____

____

How many types of scaling techniques come in the category of other scaling techniques?

____

____

____

Explain what is meant by a semantic differential scale.

Briefly define the non-comparative scaling techniques.

5.8 SUGGESTED READINGS

Harper, W. B., Ralph, W., Stanley, F. S. Marketing Research: Text & Cases. Homewood, Irwin.

Kothari, C.R., Research Methodology Methods and Techniques. Wiley Eastern Limited: New Delhi.

Sekaran, U., Bougie, R. Research Methods for Business: A skill Building Approach. Wiley, India.

William G. Z. Business Research Methods. Thomson, India.

UNIT-II **Lesson No. 6**

CONCEPTAND NATURE OFRESEARCH DESIGN

STRUCTURE

6.1 INTRODUCTION

A research design relates to the criteria that are employed when evaluating business research. In focusing on the different kinds of research design, we are paying attention to the different frameworks for the collection and analysis of data. It is, therefore, a framework for the generation of evidence that is suited both to a certain set of criteria and to the research question in which the investigator is interested. A good research design should reduce the biasness while should maximize the reliability of the data being collected and analyzed. A good research design should provide the opportunity as per the various aspects of the research problem. It should minimize the experimental error and should provide maximum information. Hence, it can be concluded the selection of research design relies upon the research problem and the nature of the research.

6.2 OBJECTIVES

After going through this unit, you should be able to describe:-

Concept of research design

Need and features of good research design

Concepts related to research design

Nature of research design

6.3 CONCEPT OF RESEARCH DESIGNS

The formidable problem that follows the task of defining the research problem is thepreparation of the design of the research project, popularly known as the “research design”. Decisionsregarding what, where, when, how muchand by what means concerning an inquiry or a research studyconstitutea research design. It is the arrangement of conditions for the collection andanalysis ofdata in a manner that aims to combine relevance to the research purpose with economy inprocedure.

The research design is the conceptual structure within which research is conducted; itconstitutes the blueprint for the collection, measurement and analysis of data. As such the designincludes an outline of what the researcher will do from writing the hypothesis and its operationalimplications to the final analysis of data. Keeping in view the above-stated design decisions; one may split the overall research designinto the following parts:

the sampling design which deals with the method of selecting items to be observedfor the given study;

the observational design which relates to the conditions under which theobservations are to be made;

the statistical design which concerns with the question of how many items are tobe observed and how the information and data gathered are to be analyzed; and

the operational design which deals with the techniques by which the proceduresspecified in the sampling, statistical and observational designs can be carried out.

In brief, the research design must, at least, contains

a clear statement of the research problem;

procedures and techniques to be used for gathering

information; the population to be studied, and

methods to be used in the processing and analyzing of data.

6.3.1 Need for reserch Designs

Research design is needed because it facilitates the smooth sailing of the various research operations,thereby making research as efficient as possible yielding maximal information with minimal expenditureof effort, time and money. Just as for better, economical and attractive construction of a house, weneed a blueprint well thought out and prepared byan expert architect, similarly we need a research design or a plan in advance of data collection andanalysis for our research project. Research design stands for the planning of the methods to beadopted for collecting the relevant data and the techniques to be used in their analysis, keeping inview the objective of the research and the availability of staff, time and money. Preparation of theresearch design should be done with great care as any error in it may upset the entire project.

Research design has a great bearing on the reliability of the results arrived at and as suchconstitutes the firm foundation of the entire edifice of the research work.Even then the need for a well-thought-out research design is at times not realized by many. Theimportance that this problem deserves is not given to it. As a result, many types of research do not servethe purpose for which they are undertaken. They may even give misleading conclusions.Thoughtlessness in designing the research project may result in rendering the research exercisefutile. It is, therefore, imperative that an efficient and appropriate design must be prepared beforestarting research operations. The design helps the researcher to organize his ideas in a form wherebyit will be possible for him to look for flaws and inadequacies. Such a design can even be given toothers for their comments and critical evaluation. In the absence of such a course of action, it will bedifficult for the critic to provide a comprehensive review of the proposed study.

6.3.2 Characteristics of Good Reserarch Designs

Proper research design sets your study up for success. Successful research studies provide insights that are accurate and unbiased. A survey is to be created that meets all of the main characteristics of a design. There are four key characteristics:

Neutrality: When you set up your study, you may have to make assumptions about the data you expect to collect. The results projected in the research should be free from bias and neutral. Understand opinions about the final evaluated scores and conclusions from multiple individuals and consider those who agree with the derived results.

Reliability: With regularly conducted research, the researcher involved expects similar results every time. Your design should indicate how to form research questions to ensure the standard of results. You'll only be able to reach the expected results if your design is reliable.

Validity: There are multiple measuring tools available. However, the only correct measuring tools are those which help a researcher in gauging results according to the objective of the research. The questionnaire developed from this design will then be valid.

Generalization: The outcome of your design should apply to a population and not just a restricted sample. A generalized design implies that your survey can be conducted on any part of a population with similar accuracy.

6.3.3 Important Concepts relating to Research Designs

Before describing the different research designs, it will be appropriate to explain the variousconcepts relating to designs so that these may be better and easily understood.

Dependent and independent variables: A concept that can take on differentquantitative values is called a variable. As such the concepts like weight, height and income are allexamples ofvariables. Qualitative phenomena (or the attributes) are also quantified based on thepresenceor absence of the concerning attributes. Phenomena that can take on quantitativelydifferent values even in decimal points are called 'continuous variables'. But all variables are notcontinuous. If they can only be expressed in integer values, they are non-continuous variables or instatistical language 'discrete variables'. Age is an example of a continuous variable, but the number ofchildrenis an example of a non-continuous variable. If one variable depends upon or is a consequenceof the other variable, it is termed a dependent variable and the variable that is antecedent to thedependent variable is termed an independent variable.

Extraneous variable: Independent variables that are not related to the purpose of thestudy, butmay affect the dependent variable are termed extraneous variables. Suppose theresearcher wants to test the hypothesis that there is a relationship between children's gains in socialstudies achievement and their self-concepts. In this case, self-concept is an independent variable andsocial studies achievement is a dependent variable. Intelligence may as well affect social studiesachievement, but since it is not related to the purpose of the study undertaken by the researcher, it willbe termed as an extraneous variable. Whatever effect is noticed on the dependent variable as a result ofextraneous variables is technically described as an 'experimental error'.

Control: One important characteristic of a good research design is to minimize theinfluence or effect of extraneous variables. The technical term 'control' is used when we design thestudy to minimize the effects of extraneous independent variables. In experimental research, theterm 'control' is used to refer to restrain experimental conditions.

Confounded relationship: When the dependent variable is not free from the influence ofextraneous variables, the relationship between the dependent and independent variables is said to beconfounded by extraneous variables.

Research hypothesis: When a prediction or a hypothesized relationship is to be tested byscientific methods, it is termed a research hypothesis. The research hypothesis is a predictivestatement thatrelates an independent variable to a dependent variable. Usually, a research hypothesismust contain, at least, one independent and one dependent variable. Predictive statements which arenot to be objectively verified or the relationships that are assumed but not to be tested are not termedresearch hypotheses.

Experimental and non-experimental hypothesis-testing research: When the purpose ofresearch is to test a research hypothesis, it is termed hypothesis-testing research. It can be theexperimental design or the non-experimental design. Research in which the independent variable ismanipulated is termed 'experimental hypothesis-testing research' and research in which anindependent variable is not manipulated is called 'non-experimental hypothesis-testing research'.

Experimental and control groups: In experimental hypothesis-testing research when agroup is exposed to usual conditions, it is termed a 'control group', but when the group has been exposedtosome novel or special condition, it is termed an 'experimental group'.

Treatments: The different conditions under which experimental and control groups are putare usually referred to as 'treatments'. In the illustration taken above, the two treatments are the usualstudies program and the special studies program. Similarly, if we want to determine through anexperiment the comparative impact of three varieties of fertilizers on the yield of wheat, in that case, the three varieties of fertilizers will be treated as three treatments.

Experiment: The process of examining the truth of a statistical hypothesis, relating tosome research problem, is known as an experiment.

Experimental units: The pre-determined plots or the blocks, where different treatmentsare used, are known as experimental units. Such experimental units must be selected verycarefully.

6.3.4 Factors Affecting Research Designs

Various factors that affect research design are as follows :

Research Questions: Research questions perform an important role in selecting the method to carry out research. There are various forms of research designs that include their methods for collecting data. For example, a survey can be conducted for the respondents to ask them descriptive or interconnected questions while a case study or a field survey can be used to identify the firm's decision-making process.

Time and Budget Limits: Researchers are bound by a restricted amount of time and budget to complete the research study. The researcher can select experimental or descriptive research when the time and budget constraints we favorable to him for the detailed study. otherwise, an exploratory research design can be adopted when the time is limited.

Research Objective: Every research is carried out to obtain the results which help to achieve some objectives. This research objective influences the selection of research design. The researcher should adopt the research design which is suitable for the research objective and also provides the best solution to the problem along with a valuable result.

Research Problem: Selection of the research design is greatly affected by the type of research problem. For example, the researcher selects an experimental research design to find out the cause-and-effect relationship of the research problem. Similarly, if the research problem includes an in-depth study, then the researcher generally adopts the experimental research design method.

Personal Experiences: Selection of research design also depends upon the personal experience of researchers. For example,the researcher who has expertise in the statistical analysis would be liable to select the quantitative research designs. While those researchers who are specialists in theoretical facets of research will be forced to select qualitative research design.

Target Audience: The type of target audience plays a very important role in the selection of research design. The researcher must consider the target audience for which the research is carriedout. Audiences may either be the general public, business professionals, or the government. For example,if the research is proposed for the general public, then the researcher should select a qualitative research design. Similarly, a quantitative research design would be appropriate for the researcher to introduce the report in front of the business experts.

6.4 NATURE OFRESEARCH DESIGNS

A good design is often characterized by adjectives like flexible, appropriate, efficient, economical and so on. Generally, the design which minimizes bias and maximizes the reliability of the data collected and analyzed is considered a good design. The design which gives the smallest experimental error is supposed to be the best design in many investigations. Similarly, a design that yields maximal information and provides an opportunity for considering many different aspects of a problem is considered the most appropriate and efficient design in respect of many research problems. Thus, the question of good design is related to the purpose or objective of the research problem and also to the nature of the problem to be studied. A design may be quite suitable in one case but may be found wanting in one respect or the other in the context of some other research problem. One single design cannot serve the purpose of all types of research problems.

A research design appropriate for a particular research problem usually involves the consideration of the following factors:

the means of obtaining information;

the availability and skills of the researcher and his staff,

if any; the objective of the problem to be studied;

the nature of the problem to be studied; and

the availability of time and money for the research work.

If the research study happens to be an exploratory or a formative one, wherein the major emphasis is on the discovery of ideas and insights, the research design most appropriate must be flexible enough to permit the consideration of many different aspects of a phenomenon. But when the purpose of a study is an accurate description of a situation or of an association between variables (or in what are called the descriptive studies), accuracy becomes a major consideration and a research design that minimizes bias and maximizes the reliability of the evidence collected is considered a good design.

6.5 PROCESS OF RESEARCH DESIGNS

The stages in the process of research design are interactive and often occur at the same time. Designing of the research study follows the given process. Steps in research design :

Figure 6.1: Steps in Research Design Process

Step 1: Defining Research Problem: The definition of a research problem is the foremost and important part of a research design process. Defining the research problem includes supplying the information that is required by the management. Without defining the research problem appropriately, the researcher can’t conclude accurate, results. While defining the research problem, the researchers first analyze the problems or opportunities in management, then analyze the situation. The purpose of clarifying the research problem is to make sure that the area of concern for research is properly reflected and management decision is correctly described. After situation analysis, they develop a model for research which helps in the next step which is a specification of information.

Step 2: Assess the Value of Information: When a research problem is approached, it is usually based on some information. These data are obtained from past experiences as well as other sources. Based on this information, some preliminary judgments are made regarding the research problem. There is always a need for additional information which is available without additional cost and delay but waiting and paying for the valuable information is quite difficult.For example, a car manufacturing industry may be concerned about the decrease in the sale of a particular model. A researcher will look for solutions by analyzing various aspects.For this, the researcher has to continuously collect a lot of information and needs to evaluate them by understanding its value and filtering out useless information.

Step 3: Select the Approach for Data Collection: For any type of research, a researcher needs data. Once, it is identified which kind of information is required for conducting the research, the researchers proceed towards collecting the data. The data can be collected using secondary or primary sources.Secondary data is the previously collected information for some other purpose, while the primary data is collected by the researcher, especially for the research problem.

Step 4: Select the Measurement Technique: After collecting data, the measurement technique for the collected data is selected. The major measurement techniques used in research are as follows:

Questionnaire: The questionnaire is a formal structure that contains questions to collect information from the respondents regarding their attitude, beliefs, behavior, knowledge, etc.

Attitude Scales: Attitude scales are used to extract the beliefs and feelings of the respondents regarding an object or issue.

Observation: It is the monitoring of behaviors and psychological changes of the respondents. It is widely used in research.

Projective Techniques and Depth Interview: Sometimes direct questions are not sufficient to get true responses from the individuals, that is why. different approaches like depth interviews and projective techniques are used. These techniques allow the respondents to give their responses without any fear. The researcher neither disagrees nor advises on these techniques.

Step 5: Sample Selection: Once, the measurement technique has been selected, the next step is selecting the sample to conduct the research. The researchers in this stage select a sample out of the total population instead of considering the population as a whole. Samples can be selected by using two techniques, i.e., random sampling techniques, and non-random sampling techniques.

Step 6: Selecting Model of Analysis: Researchers select the model of analysis or technique of data analysis, before collecting data. After this, researchers evaluate the techniques using hypothetical values to ensure that the measurement technique would provide the desired outcome regarding the research problem.

Step 7: Evaluate the Ethics of Research: While conducting research, it becomes very much necessary for the researcher to follow ethical practices. The research which is conducted ethically draws the interests of the general public, respondents, clients and other research professionals. Hence, it becomes the duty of the researcher to evaluate the practices in research, to avoid any biasness on behalf of the observer and researcher as well.

Step 8: Estimate Time and Financial Requirements: This step is one of the most important steps in designing research. Here, researchers use different methods like the Critical Path Method (CPM) and Programme Evaluation Review Technique (PERT) to design the plan as well as control the process and determine the resources required.A flowchart of these activities along with their approximate time is prepared for visual assessment of the research process. With the help of this chart, the researcher can find out the sequence of activities to be taken.

Step 9: Prepare the Research Proposal: The final step in the process of research design is preparing the research proposal. A research proposal or the research design is prepared for the operation and control of research. An effective research proposal is prepared before the actual conduction of the research.

6.6 SUMMARY

The research design is to provide a plan of study that permits an accurate assessment of cause and effect relationships between independent and dependent variables. The classic controlled experiment is an ideal example of good research design. Factors that jeopardize the evaluation of the effect of experimental treatment (internal validity) and the generalizations derived from it (external validity) are identified. Sources of variance can be controlled by eliminating a variable, randomization, matching, or including a variable as part of the design. A research project should be so designed that (1) it answers the questions being investigated,

extraneous factors are controlled, and (3) the degree of generalization that can be made is valid.

6.7 GLOSSARY

Research Design: Research design is a blueprint of a scientific stud, It includes research methodolotlies. tools, and techniques to conduct the research. It helps to iderinfs and address the problem that may rise during the process of research and analy'sis.

Dependent Variable: The variable that depends on other factors **that are** measured. These variables are expected to change as a result of experimental manipulation of the independent variable or variables.

Independent Variable: It is a variable that stands alone and isn't changed by the other variables you are trying to measure.

Extraneous Variable: An extraneous variable is any variable that is not investigated and that can potentially affect the dependent variable of your research study.

6.8 LESSON END EXERCISE

What is research design?

____

____

What are the characteristics of research design?

Explain the need for research design.

Briefly highlight the concepts related to research design.

Define the steps involved in the research design process.

6.9 SUGGESTED READINGS

Boyd, Westfall, and Stasch, Business Research Text and Cases. All India Traveller Bookseller, New Delhi.

Brown, F.E. Business Research, a structure for decision making. Addison - Wesley Publishing Company.

Stockton and Clark, Introduction to Business and Economic Statistics. D.B. Taraporevala Sons and Co. Private Limited, Bombay.

Srivastava, T. N., Rego, S., Business Research Methodology.TataMcGraw Hill, New Delhi.

UNIT-II **Lesson No. 7**

CLASSIFICATION OFRESEARCH DESIGN

STRUCTURE

7.1 INTRODUCTION

By the term 'research', we can understand that it's a collection of data that includes critical information by considering research methodologies. In other words, it is a compilation of information or data explored by setting a hypothesis and consequently coming up with substantive findings in an organized way. Research can be done on an academic as well as on a scientific basis as well. To carry out research in an effective manner various methods of research design are discussed. Using the correct form of research will also ensure that you are measuring and observing the correct elements and by way of a frequency best suited to your research issue. Moreover, when you employ the proper type of research, it is far less likely to come upon errors and gaps that require answers. Thus, there is less of a need to start again or switch to a different type of research. All of these areas of importance would be impossible to fulfill if you do not become familiar with them and are not able to tell them apart.

7.2 OBJECTIVES

After going through this unit, you should be able to describe:-

Various types of research design

Use of the descriptive method for research design

Use of the exploratory method for research design

Use of the case study method for research design

7.3 CLASSIFICATION OFRESEARCH DESIGNS

A Research Design is simply a structural framework of various research methods as well as techniques that are utilized by a researcher. It helps a researcher to pursue their journey into the unknown but with a systematic approach by their side. The way an **engineer** or**architect**frames a design for a structure, likewise the researcher picks the design from various approaches to check which type of research to be carried out. There are several ways to distinguish research design.Depending on the method and purpose of design, we can categorize research design into the following types:

7.3.1 Descriptive Method

In Descriptive Research Design, the researcherdescribes the situation or case in depth in their research materials. This type of research design is purely on a theoretical basis where the individual collects data, analyses, prepares and then presents it understandably. It is the most generalized form of research design. To explore one or more variables, a descriptive design might employ a wide range of research approaches. Unlike in experimental research, the researcher does not control or change any of the variables in a descriptive research design; instead, he or she just observes and measures them. In other words, while qualitative research may also be utilized for descriptive reasons, a descriptive method of research design is typically regarded as a sort of quantitative research. To guarantee that the results are legitimate and dependable, the study design should be properly constructed.

In this hypothesis-based design methodology, the researcher primarily describes the subject matter central to the research. Descriptive research design applies to natural observations, case studies, and surveys. This method involves data collection, data analysis, and presentation. It allows the researcher to put forth the problem to persuade others to comprehend the necessity of the research.

Here are some examples of the descriptive design of research type:

How has the Delhi housing market changed over the past 20 years?

Do customers of company A prefer product C or product D?

What are the main genetic, behavioural and morphological differences between Indian wildcows and hybridcows?

How prevalent is COVID-19 in the population of a specific area?

What does this study explain?

The subject is being observed in a completely natural and unchanged natural environment. True experiments, whilst giving analyzable data, often adversely influence the normal behavior of the subject.

Descriptive research is often used as a pre-cursor to more quantitatively research designs, the general overview giving some valuable pointers as to what variables are worth testing quantitatively.

If the limitations are understood, they can be a useful tool in developing a more focused study.

Descriptive studies can yield rich data that lead to important recommendations.

The approach collects a large amount of data for detailed analysis.

What this study doesn't explain?

The results from a descriptive research can not be used to discover a definitive answer or to disprove a hypothesis.

Because descriptive designs often utilize observational methods [as opposed to quantitative methods], the results cannot be replicated.

The descriptive function of research is heavily dependent on instrumentation for measurement and observation.

Advantages of Descriptive Method

Large scale and quick: Descriptive data can be gathered from a large number of respondents using the right mechanism that maximizes reach along with increasing promptness in distribution and collection.

Cohesive mechanism: This method of research collects both qualitative and quantitative data and can be used as a stand-alone data collection method.

Natural environment: The possibility to observe the phenomenon in a completely natural and unchanged natural environment.

Comprehensive: The opportunity to integrate the qualitative and quantitative methods of data collection. Accordingly, research findings can be comprehensive.

Less time-consuming: It consumes less time when compared with other quantitative experiments.

Easy to handle: Effective to analyze non-quantified topics and issues. Practical use of research findings for decision-making.

Disadvantages of Descriptive Method

Cause and effect relationship cannot be established: It is difficult to highlight cause and effect relationships using correlational studies.

Restrictive: Case studies, which are a type of descriptive study may limit your participants to peculiar individuals which are a good fit for that case.

Doesn't uncover a cause and effect relationship: This method does not allow the identification of a cause behind a phenomenon.

Lack of Statistical tools: Descriptive studies cannot test or verify the research problem statistically and research results may reflect a certain level of bias due to the absence of statistical tests.

Not repeatable: The majority of descriptive studies are not 'repeatable' due to their observational nature.

7.3.2 Exploratory Method

An exploratory design is conducted about a research problem when there are few or no earlier studies to refer to. The focus is on gaining insights and familiarity for later investigation or undertaken when problems are in a preliminary stage of the investigation. This is utilized by researchers to explore, expand and explain theories and innovative ideas. This design method is applied to find the missing pieces of a puzzle or obtain clarity on vague aspects of a certain topic.

The primary objective of exploratory research is to provide insights into, and an understanding of, the problem confronting the researcher. It is used in cases when the researcher defines the problem more precisely, identifies relevant courses of action, or gains additional insights before an approach can be developed. The information needed is only loosely defined at this stage and the research process that is adopted is flexible and unstructured.

The goals of exploratory research are intended to produce the following possible insights:

Familiarity with basic details, settings and concerns.

Well-grounded picture of the situation being developed.

Generation of new ideas and assumptions, development of tentative theories or hypotheses.

Determination about whether a study is feasible in the future.

Issues get refined for more systematic investigation and formulation of new research questions.

For example, it may consist of personal interviews with industry experts. The sample, selected to generate maximum insights, is small and non-representative. The primary data are qualitative arid and are analyzed accordingly. Given these characteristics of the research process, the findings of exploratory research should be regarded as tentative or as input to further research.

What does this study explain?

Design is a useful approach for gaining background information on a particular topic.

Exploratory research is flexible and can address research questions of all types (what, why, how).

Provides an opportunity to define new terms and clarify existing concepts.

Exploratory research is often used to generate formal hypotheses and develop more precise research problems.

Exploratory studies help establish research priorities.

What this study doesn't explain?

Exploratory research generally utilizes small sample sizes and, thus, findings are typically not generalizable to the population at large.

The exploratory nature of the research inhibits an ability to make definitive conclusions about the findings.

The research process underpinning exploratory studies is flexible but often unstructured, leading to only tentative results that have limited value in decision-making.

Design lacks rigorous standards applied to methods of data gathering and analysis because one of the areas for exploration could be to determine what method or methodologies could best fit the research problem.

Methods of Exploratory Research

A researcher may choose from four general categories of exploratory research methods: (1) Experience Surveys, (2) Secondary Data Analysis, (3) Case Studies/Analysis, (4) Focus Group and (5) Projective Techniques. Each category provides various alternative ways of gathering information.

Experience Surveys: In attempting to understand the problems at hand, managers may discuss issues and ideas with top executives and knowledgeable managers who have had personal experience in the field. This constitutes an informal experience survey.

Secondary Data Analysis: Another economical and quick source of background information is trade literature in the public library. Searching through such material is exploratory research using secondary data analysis. Basic theoretical research is rarely conducted without extensive reviews of the literature in the field or similar research. Using secondary data may be equally important in applied research.

Case Studies/Analysis: The purpose of the case study method is to obtain information from one or a few situations that are similar to the researcher's problem situation.

Focus Group: Focus groups are small groups of people brought together and guided by a moderator throughout the unstructured, spontaneous discussion to gain information relevant to the research objective.

Projective Techniques: Projective techniques seek to explore hidden consumer motives for buying goods & services by asking participants to project themselves into a situation and then respond to a specific question regarding that situation. Sentence completion, cartoon test, word association, etc. are such types of techniques.

Advantages of Exploratory Research

The following set of advantages of exploratory research advocate for its use:

Exploratory research offers a great amount of researcher discretion. The lack of structure enables the researcher to direct the progression of the research processes and in that sense, it offers a greater degree of flexibility and freedom.

Exploratory research is the economical way in which the process can be

conducted. Exploratory research uses a relatively smaller group of people for defining and understanding the research issue.

Exploratory research when done properly can lay a strong foundation for any study that is carried around the same issue in the future. Exploratory research that is properly carried out helps in determining research design, sampling methodology, and data collection. This also comes with a sense of responsibility for the researcher to try and inspect the issue in-depth and concentrate on authentic reporting of results.

Analyzing the feasibility and viability of the research issue is another pro of exploratory research. No organization wants to invest time, effort, and resources in an area that is incapable of making value addition to the overall functioning. By carrying out an early study, exploratory research gauges the future importance that the research topic holds and accordingly directs organizational attitude.

Exploratory research formulates a greater understanding of a previously unresearched topic and satisfies the researcher uncovers facts and brings new issues to light. In doing so, it helps refine future research questions. It also helps decide the best approach to reach the objective.

Disadvantages of Exploratory Research

Exploratory research comes with its own set of cons that can act as roadblocks that impede a seamless data collection experience which lays the groundwork for future probes as well:

Exploratory research brings up tentative results and so is inconclusive. The focus of such research is to grasp and formulate a better understanding of the issue at hand. These research insights cannot be relied upon for effective decision-making.

Another con of **exploratory research is its qualitative data and subsequent analysis.** It is difficult to derive accurate insights that can be summarized objectively. The variability in qualitative data itself evaluates data collected, a difficult and cumbersome process.

The small sample used for exploratory research increases the risk of the sample responses being non-representative of the target audience. Smaller groups of people as samples, however useful for a quick study, can hinder a cohesive understanding which not only deteriorates the current quality of research but also adversely impacts future research carried out along similar lines.

Data, when gathered through secondary resources, can supply obsolete information which may not generate any significant contribution to the understanding of an issue in the current scenario. Outdated information is neither actionable nor supportive in offering any sort of clarity under dynamic market conditions.

7.3.3 Case Study Method

A case study is an investigation into an individual circumstance. The investigation may be of a single person, business, event, or group. The investigation involves collecting in-depth data about the individual entity through the use of several collection methods. Interviews and observation are two of the most common forms of data collection used. This method was originally developed in the field of clinical medicine. It has expanded since to other industries to examine key results, either positive or negative, that were received through a specific set of decisions. This allows for the topic to be researched with great detail, allowing others to glean knowledge from the information presented. For example, if an anthropologist were to live amongst a remote tribe, whilst their observations might produce no quantitative data, they are still useful to science.

A case study is a detailed study of a specific subject, such as a person, group, place, event, organization, or phenomenon. Case studies are commonly used in social, educational, clinical, and business research.A case study is an in-depth study of a particular research problem rather than a sweeping statistical survey. It is often used to narrow down a very broad field of research into one or a few easily researchable examples. The case study research design is also useful for testing whether a specific theory and model applies to phenomena in the real world. It is a useful design when not much is known about a phenomenon.

A case study research design usually involves qualitative methods, but quantitative methods are sometimes also used. Case studies are good for describing, comparing, evaluating and understanding different aspects of a research problem. Case studies are often a good choice in a thesis or dissertation. They keep your project focused and manageable when you don't have the time or resources to do large-scale research.

What does this study explain?

Approach excels at bringing us to an understanding of a complex issue through detailed contextual analysis of a limited number of events or conditions and their relationships.

A researcher using a case study design can apply a variety of methodologies and rely on a variety of sources to investigate a research problem.

Design can extend experience or add strength to what is already known through previous research.

Social scientists, in particular, make wide use of this research design to examine contemporary real-life situations and provide the basis for the application of concepts and theories and the extension of methods.

The design can provide detailed descriptions of specific and rare cases.

What this study doesn't explain?

A single or small number of cases offers little basis for establishing reliability or generalizing the findings to a wider population of people, places, or things.

The intense exposure to the study of the case may bias a researcher's interpretation of the findings.

The design does not facilitate the assessment of cause and effect relationships.

Vital information may be missing, making the case hard to interpret.

The case may not be representative or typical of the larger problem being investigated.

If the criteria for selecting a case is because it represents a very unusual or unique phenomenon or problem for study, then your interpretation of the findings can only apply to that particular case.

Types of Cases

Following on from the issue of external validity, it is useful to consider a distinction between different types of case that is sometimes made by writers. distinguishes five types.

The Critical Case: Here the researcher has a specified hypothesis, and a case is chosen because it will allow a better understanding of the circumstances in which the hypothesis will and will not hold.

The Unique Case: The unique or extreme case is the common focus in clinical studies.

The Revelatory Case: The basis for the revelatory case exists when an investigator has an opportunity to observe and analyze a phenomenon previously inaccessible to scientific investigation. While the idea of the revelatory case is interesting, it seems unnecessary to restrict it solely to situations in which something has not previously been studied.

The Representative/Typical Case: This type seeks to explore a case that exemplifies an everyday situation or form of organization.

The Longitudinal Case: This type of case is concerned with how a situation changes over time.

Advantages of the Case Study Method

It turns client observations into useable data. Case studies offer verifiable data from direct observations of the individual entity involved. These observations provide information about input processes. It can show the path taken which led to specific results being generated. Those observations make it possible for others, in similar circumstances, to potentially replicate the results discovered by the case study method.

It turns opinion into fact. Case studies provide facts to study because you're looking at data that was generated in real-time. It is a way for researchers to turn their opinions into information that can be verified as fact because there is a proven path to positive or negative development. Singling out a specific incident also provides in-depth details about the path of development, which gives it extra credibility to the outside observer.

It is relevant to all parties involved. Case studies that are chosen well will be relevant to everyone who is participating in the process. Because there is such a high level of relevance involved, researchers can stay actively engaged in the data collection process. Participants can further their knowledge growth because there is interest in the outcome of the case study. Most importantly, the case study method essentially forces people to decideon the question being studied, then defend their position through the use of facts.

It uses different research methodologies. The case study method involves more than just interviews and direct observation. Case histories from a records database can be used with this method. Questionnaires can be distributed to participants in the entity being studied. Individuals who have kept diaries and journals about the entity being studied can be included. Even certain experimental tasks, such as a memory test, can be part of this research process.

It can be done remotely. Researchers do not need to be present at a specific location or facility to utilize the case study method. Research can be obtained over the phone, through email, and other forms of remote communication. Even interviews can be conducted over the phone. That means this method is good for exploratory and formative research, even if it must be completed from a remote location.

It is inexpensive. Compared to other methods of research, the case study method is rather inexpensive. The costs associated with this method involve accessing data, which can often be done for free. Even when there are in-person interviews or other on-site duties involved, the costs of reviewing the data are minimal.

It is very accessible to readers. The case study method puts data into a usable format for those who read the data and note its outcome. Although there may be perspectives of the researcher included in the outcome, the goal of this method is to help the reader be able to identify specific concepts to which they also relate. That allows them to discover unusual features within the data, examine outliers that may be present, or draw conclusions from their own experiences.

Disadvantages of the Case Study Method

It can influence factors within the data. Every person has their own unconscious bias. Although the case study method is designed to limit the influence of this bias by collecting fact-based data, it is the collector of the data who gets to define what is a "fact" and what is not. That means the real-time data being collected may be based on the results the researcher wants to see from the entity instead. By controlling how facts are collected, research can control the results this method generates.

It takes longer to analyze the data. The information collection process through the case study method takes much longer to collect than other research options. That is because there is an enormous amount of data that must be sifted through. It's not just the researchers who can influence the outcome of this type of research method. Participants can also influence outcomes by giving inaccurate or incomplete answers to questions they are asked. Researchers must verify the information presented to ensure its accuracy, and that takes time to complete.

It can be an inefficient process. Case study methods require the participation of the individuals or entities involved for it to be a successful process. That means the skills of the researcher will help to determine the quality of information that is being received. Some participants may be quiet, and unwilling to answer even basic questions about what is being studied. Others may be overly talkative, exploring tangents that have nothing to do with the case study at all. If researchers are unsure of how to manage this process, then incomplete data is often collected.

It requires a small sample size to be effective. The case study method requires a small sample size for it to yield an effective amount of data to be analyzed. If there are different demographics involved with the entity, or there are different needs that must be examined, then the case study method becomes very inefficient.

It is a labor-intensive method of data collection. The case study method requires researchers to have a high level of language skills to be successful with data collection. Researchers must be personally involved in every aspect of collecting the data as well. From reviewing files or entries personally to conducting personal interviews, the concepts and themes of this process are heavily reliant on the amount of work each researcher is willing to put into things.

7.4 SUMMARY

This chapter briefed on the classification of research design methods like descriptive method, exploratory method and case study method. The chapter also highlights the advantages and disadvantages of the various research designwere also discussed.

7.5 GLOSSARY

Desciptive Method : Descriptive research aims to accuretely and systematically describe a population. situation or phenomenon.

Exploratory Method: Exploratory research is a methodological approach that investigates research questions that have not previously been studied in depth. Exploratory research is often qualitative and primary.

Case Study Method: A case study is a research approach that is used to generate **an in** depth, multi-faceted understanding of a complex issue in its real-life context.

7.6 LESSON END EXERCISE

Briefly define the descriptive method.

____

____

____

Briefly define the case study method.

____

____

____

How many types of cases are used in the case study method?

____

____

Briefly define the exploratory method and its types.

7.7 SUGGESTED READINGS

Boyd, Westfall, and Stasch, Business Research Text and Cases. All India Traveller Bookseller, New Delhi.

Sekaran, U., Bougie, R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Kothari, C.R., Research Methodology Methods and Techniques, Wiley Eastern Limited: New Delhi.

William G. Z. Business Research Methods. Thomson, India.

UNIT-II **Lesson No. 6**

QUALITATIVEVS QUANTITATIVE RESEARCH, CROSS-SECTIONAL STUDY, LONGITUDINALSTUDYAND ERRORSAFFECTING RESEARCH DESIGN

STRUCTURE

8.1 INTRODUCTION

Keeping its dynamics into consideration, the research design is categorized into two different perspectives, i.e. Quantitative Research Design and Qualitative Research Design. A researcher should have a clear understanding of how their project can be implemented in the research design. The suggestion that such criteria are mainly relevant to quantitative research is examined, along with the proposition that an alternative set of criteria should be employed in qualitative research. This alternative set of criteria, which is concerned with the issue of trustworthiness, is outlined briefly. The cross-sectional study is the most frequently used descriptive design in marketing research. Cross-sectional designs involve the collection of information from any given sample of population elements only once. They may be either single cross-sectional or multiple cross-sectional. In longitudinal designs, a fixed sample (or samples) of population elements is measured repeatedly. A longitudinal design differs from a cross-sectional design in that the sample or samples remain the same over time. In other words, the same respondents are studied over time.Any research project is subject to errors, so a research designer must do everything to minimize them. Understanding the possible errors that can taint the accuracy of the information in your study is key to avoiding and correcting sampling errors and non-sampling errors.

8.2 OBJECTIVES

After going through this unit, you should be able to describe:-

The difference between Quantitative and Qualitative research design

Cross-sectional study design, the most common form of which is social survey research

Longitudinal study design, such as the panel study and the cohort

study Sampling and non-sampling errors

8.3 QUALITATIVEVS QUANTITATIVE RESEARCH

Qualitative research is a research strategy that indicates the relationship between theory and research and usually emphasizes how theories were generated. As a research strategy, qualitative research is inductivist, constructionist, and interpretivist, but qualitative researchers always don't subscribe to all three of these methods whereas Quantitative research is an inquiry into a social problem, explaining phenomena by gathering numerical data that are analyzed using mathematically based methods in particular statistics. In this type of research, the researcher primarily uses a post-positivist approach to develop knowledge when quantitative research is selected, employs strategies of inquiry such as experiments and surveys, and collects data on predetermined instruments that yield statistical data.

8.3.1 Qualitative Research Design

This type of research is quite contrary to quantitative research design. It is explanatory and always seeks answers to "What's" and "How's". It mainly focuses on why a specific theory exists and what would be the respondent's answer to it. This allows a researcher to conclude with proper findings. Case studies are mainly used in Qualitative Research Design to understand various social complexities. This type of data includes textual information that presents insights and explanations. Unstructured questions which allow the respondents to elaborate their opinions without any restrictions are included under this sub-type.

Advantages of Qualitative Research Design

Comprehensive: A respondent has the liberty to dive deep into explanations and follow-ups that allow researchers to identify positives and fallouts.

Flexible and genuine: Given that the onus of answering lies completely with the respondent as there are no fixed answer choices, the number of respondents does not matter as long as the ones included in the research are representative of the target group's sentiment and sampling bias can be avoided.

Disadvantages of Qualitative Research Design

Difficult to analyze: Analysing qualitative information is a meticulous activity that requires careful attention to detail while listing key takeaways from individual responses

Highly subjective: Qualitative responses vary in terms of length and type of responses. Listing one standard summary that provides an umbrella to each response is tedious, particularly when there is a large sample involved

8.3.2 Quantitative Research Design

In Quantitative Research Design, a researcher examines the various variables while including numbers as well as statistics in a project to analyze its findings. The use of graphics, figures, and pie charts is the main form of data collection measurement and meta-analysis (it is information about the data by the data).This includes pre-planned questions which provide numeric answers to questions. This method usually explains the "what" in research.

Advantages of Quantitative Research Design

Easily summarized: Numerical data is easy to analyze using data and statistical analysis tools that provide a holistic summary of respondent answers.

Objective in nature: The questions in this method have limited answer options making it easy to establish standardized answers that have low room for variability.

Disadvantages of Quantitative Research Design

Lacks insights: Quantitative method does not allow the respondent to elaborate on the reasoning behind the choice of their answers. The restrictive nature of such a study makes it necessary for the researcher to complement their study by using open-ended comments that give a peek into the respondent's mindset.

Requires expertise: Although there are available tools to put quantitative data into perspective, the leg work of choosing and applying those specific tools requires a certain level of skill so that the analysis can be properly executed.

8.3.3 Difference between Quantitative and Qualitative Research Design

Quantitative and qualitative research use different research methods to collect and analyze data, and they allow you to answer different kinds of research questions.

Quantitative Research	Qualitative Research
Quantitative research is used if the researcher wants to confirm or test something. It focuses on testing theories and hypotheses.	Qualitative research is used if the researcherwants to understand something. It focuses on exploring ideas and formulating a theory or hypothesis
It is analyzed through mathematical andstatistical analysis	It is analyzed by summarizing, categorizing and interpreting
This design is mainly expressed in numbers, graphs and tables	This design is mainly expressed in words
It requires many respondents	It requires few respondents
It comprises Closed-ended (multiple choice) questions	It comprisesOpen-ended questions
Key terms: testing, measurement, objectivity, replicability	**Key terms:** understanding, context, complexity, subjectivity

Table 8.1 : Difference Between Quantitative and Qualitative Research Design

8.3.4 Data Collection and Data Analysis Methods

Quantitative Data Collection Methods

Surveys: List of closed or multiple-choice questions that are distributed to a sample (online, in person, or over the phone).

Experiments: Situation in which variables are controlled and manipulated to establish cause-and-effect relationships.

Observations: Observing subjects in a natural environment where variables can't be controlled.

Analyzing Quantitative Data

Quantitative data is based on numbers. Simple math or more advanced statistical analysis is used to discover commonalities or patterns in the data. The results are often reported in graphs and tables.Applications such as Excel, SPSS, or R can be used to calculate things like:

Average scores

The number of times a particular answer was given

The correlation or causation between two or more variables

The reliability and validity of the results.

Qualitative Data Collection Methods

Interviews: Asking open-ended questions verbally to respondents.

Focus groups: Discussion among a group of people about a topic to gather opinions that can be used for further research.

Ethnography: Participating in a community or organization for an extended period to closely observe culture and behavior.

Literature review: Survey of published works by other authors.

Analyzing Qualitative Data

Qualitative data is more difficult to analyze than quantitative data. It consists of text, images or videos instead of numbers.Some common approaches to analyzing qualitative data include:

Qualitative content analysis: Tracking the occurrence, position and meaning of words or phrases

Thematic analysis: Closely examining the data to identify the main themes and patterns

Discourse analysis: Studying how communication works in social contexts.

8.4 CROSS-SECTIONALSTUDYAND LONGITUDINALSTUDY

Conclusive Research Designprovides information on the evaluation of alternative courses of action and selecting onefrom among a number available to the researcher. Conclusiveresearch is again classified as:

Descriptive research, and

II. Causal research.

Descriptive Research: It is simple to understand as the name itself suggests that it involvesdescribing something. The majority of research studies are descriptive studies.

As research studies involve investigatingthe customers/consumers, the collection of data includes interrogating the respondents in themarket and data available from secondary data sources. However, it cannot be concludedthat descriptive studies should be simply a fact-gathering process. A descriptive study dealswith the respondents in the market and hence, extreme caution has to be exercised indeveloping this study. Much planning should be done, objectives should be clear thanexploratory studies.

Descriptive studies are again classified into two types:

Cross-SectionalStudy

Longitudinal Study

A Cross-Sectional study is the most predominantly and frequently used descriptiveresearch design in marketing. It involves a sample of elements from the population ofinterest. The sample elements are measured on many characteristics. There are twotypes of cross-sectional studies:Field studies andSurveys.

Field studies and surveys may be no different but the same. However, forpractical reasons, they are classified into two categories cross-sectional research. Thefundamental difference lies in the depth of what this research cover. While the surveyhas a larger scope, the field study has greater depth. Survey attempts to be representativeof some known universe and field study is less concerned with the generation of largerepresentative samples and is more concerned with the in-depth study of a few typicalsituations.A cross-sectional study is defined as an observational study where data is collected as a whole to study a population at a single point in time to examine the relationship between variables of interest.In an observational study, a researcher records information about the participants without changing anything or manipulating the natural environment in which they exist. The cross-sectional design may be either single or multiple cross-sectional designs dependingon the number of samples drawn from a population. In a single cross-sectional design,only one sample of respondents is drawn whereas in multiple cross-sectional designs,there are two or more samples of respondents. A type of multiple cross-sectionaldesign of special interest is Cohort analysis.Cohort analysis consists of a series of surveys conducted at appropriate time intervals,where the cohort serves as the basic unit of analysis. A cohort is a group ofrespondents who experience the same event within the same time interval.

A Longitudinal Study, like a cross-sectional study, is also an observational study, in which data is gathered from the same sample repeatedly over an extended period. This study relies on panel data and panel methods. It involves fixing a panelconsisting of a fixed sample of subjects that are measured repeatedly. The panel membersare those who have agreed to provide information at specific intervals over an extendedperiod. Panel data is analytical and possess advantages for the information collected inthe study. They are also considered to be more accurate than cross-sectional data becausepanel data better handle the problem associated with the errors that arise in reporting pastbehaviour and the errors that arise because of the necessary interaction betweeninterviewer and respondent.The longitudinal study can last from a few years to even decades depending on what kind of information needs to be obtained. The benefit of conducting a longitudinal study is that researchers can make notes of the changes, make observations and detect any changes in the characteristics of their participants. One of the important aspects here is that longitudinal study extends beyond a single frame in time. As a result, they can establish a proper sequence of the events that occurred.

For example, data obtained from panels formed to provide information on marketshares are based on an extended period, but also allow the researcher to examinechanges in market share over time. New members may be included in the panel when there is a dropout of the existing members or to maintain representativeness.

Causal Research: It is used to obtain evidence of cause-and-effect relationships with isotherwise known as the independent-dependent relationship or the predictive relationships.This is an important type of research useful for marketers as this allows marketers to basetheir decision on assumed causal relationships. Causal research is done in the followingsituations:

To identify which variables are the cause and which are the effect. In statistical terms,causal variables are called independent variables and effectual variables are calleddependent variables.

To determine the nature of the relationship between the causal variables and the effect tobe predicted.Causal research requires a strong degree of planning on the design as its success depends onthe structure of the design.

8.4.1 Difference between Cross-Sectional and Longitudinal Study

Cross-sectional and longitudinal studies both are types of observational studies, where the participants are observed in their natural environment. There are no alterations or changes in the environment in which the participants exist.Despite this marked similarity, there are distinctive differences between both these forms of study. Let us analyze the differences between cross-sectional studies and longitudinal studies.

Cross-Sectional Study	Longitudinal Study
Cross-sectional studies are quick to conduct.	Longitudinal studies may vary from a few years to even decades.
A cross-sectional study is conducted at a given point in time.	A longitudinal study requires a researcher to revisit participants of the study at proper intervals.
A cross-sectional study is conducted with different samples.	A longitudinal study is conducted with the same sample over the years.
Cross-sectional studies cannot pin down the cause-and-effect relationship.	The longitudinal study can justify the cause-and-effect relationship.
Multiple variables can be studied at a single point in time.	Only one variable is considered to conduct the study.
A cross-sectional study is comparatively cheaper.	Since the study goes on for years longitudinal study tends to get expensive.

Table 8.2: Difference between Cross-Sectional and Longitudinal Study

8.5 ERRORSAFFECTINGRESEARCH DESIGN

Designing a research project takes time, skill and knowledge. It is the process with a clear goal and methods, that likely comes out with skewed data or an inaccurate picture of what the researcher was trying to accomplish. Errors occur so often that it is a statistical practice to include a margin of error in the final results. A margin of error is the amount allowed in case of a miscalculation to represent the difference between the sample and the actual population. So it’s important to use the proper methodology in the research process, and it’s equally important to avoid making critical mistakes that could produce inaccurate results. The following are the common errors accrued in the research process and how toavoid making them get the best data possible.There are two major types of errors in research design: sampling error and non-sampling error.

8.5.1 Sampling Errors

A sampling error is a statistical error that occurs when an analyst does not select a sample that represents the entire population of data. As a result, the results found in the sample do not represent the results that would be obtained from the entire population. Sampling is an analysis performed by selecting several observations from a larger population. The method of selection can produce both sampling errors. A sampling error is a deviation in the sampled value versus the true population value. Sampling errors occur because the sample is not representative of the population or is biased in some way. Even randomized samples will_have some degree of sampling error because a sample is only an approximation of the population from which it is drawn.

Types of Sampling Errors

There are different categories of sampling errors:

Population-Specific Error: A population-specific error occurs when a researcher doesn’t understand who to survey. This type of error occurs when the researcher selects an inappropriate population or universe from which to obtain data.

For example, frequently survey housewives because they are easy to contact. Also, it is assumed housewives decide what is to be purchased and do the actual purchasing for a household. however, in this situation, there often is a population specification error. increasingly, husbands may purchase a significant share of the packaged goods and have significant influence over what is bought.

Selection Error: Selection error is the sampling error that occurs when a sample is selected by a nonprobability method. Selection error occurs when the survey is self-selected, or when only those participants who are interested in the survey responses to the questions. Researchers can attempt to overcome selection errors by finding ways to encourage participation. For example, interviewers conducting a mall intercept study have a natural tendency to select those respondents who are the most accessible and agreeable. Such samples often comprise friends and associates who are rarely representative of the desired population. Selection error often reflects people who are easily reached, are better dressed, have better-kept homes, or are more pleasant. These types of samples rarely represent the desired population. having clear, written procedures that specify how to select respondents can help to reduce selection error.

Sample Frame Error: A sample frame error occurs when a sample is selected from the wrong population data. a sampling frame supposedly represents all the members of the population. it is usually a listing of the respondents you want to sample. For example, the sample frame for a study at a shopping mall includes all shoppers in the mall during the time of data collection. in years past, a commonly used frame for consumer research was the telephone directory. over time, this frame has increasingly introduced errors because many elements of the population (households, singles, students) are no longer included in the directory. there are also unlisted phone numbers, move-ins, and cell phones to consider.

Non-response Error: A non-response error occurs when a useful response is not obtained from the surveys because researchers were unable to contact potential respondents. Non-response error occurs when respondents and non-respondents are too different. Respondents should accurately represent the population that is to be sampled.if non-respondents are not equally distributed across the population, you will not have an accurate sample. there are two ways in which survey non-response can occur:(a) non-contact (the inability to contact all members of the sample frame); and(b) refusal (non-response to some or all items on the measurement instrument).For example, in telephone surveys, non-respondents are not available because they are not at home for the initial call or call-backs, they have moved, or they are away from home during the period of the survey.

Minimizing Sampling Errors

The prevalence of sampling errors can be reduced by increasing the sample size. As the sample size increases, the sample gets closer to the actual population, which decreases the potential for deviations from the actual population. Steps can also be taken to ensure that the sample adequately represents the entire population. Researchers might attempt to reduce sampling errors by replicating their studies. This could be accomplished by taking the same measurements repeatedly, using more than one subject or multiple groups, or by undertaking multiple studies. Of the two types of errors, sampling error is easier to identify. The biggest techniques for reducing sampling error are:

Increase the sample size:A larger sample size leads to a more precise result because the study gets closer to the actual population size.

Divide the population into groups:Instead of a random sample, test groups according to their size in the population. For example, if people of a certain demographic make up 35% of the population, make sure 35% of the study is made up of this variable.

Know your population: The error of population specification is when a research team selects an inappropriate population from which to obtain data. Know who buys your product uses it, works with you, and so forth. With basic socio-economic information, it is possible to reach a consistent sample of the population. In cases like marketing research, studies often relate to one specific population like Facebook users, Baby Boomers, or even homeowners.

8.5.2 Non-Sampling Errors

A non-sampling error is a statistical term that refers to an error that results during data collection, causing the data to differ from the true values. A non-sampling error differs from a sampling error.

It encompasses all types of errors, mostly caused by human error, such as questionnaire wording, data entry errors, and biased decisions. Non-sampling errors cover all other discrepancies, including those that arise from a poor sampling technique. These errors may be present in both samples and censuses in which an entire population is surveyed.

Non-sampling error can arise when either a sample or an entire population (census) is taken. It falls under two categories:

Random Errors: Random errors are errors that cannot be accounted for and just happen. In statistical studies, it is believed that each random error offsets each other, generally speaking, so they are of little to no concern.

Systematic Errors: Systematic errors affect the sample of the study and, as a result, will often create useless data. A systematic error is consistent and repeatable, so the study's creators must take great care to mitigate such an error**.**

Types of Non-Sampling Errors

Non-sampling errors can occur from several aspects of a study. The most common non-sampling errors include errors in data entry, biased questions and decision-making, non-responses, false information, and inappropriate analysis.There are several types of non-sampling errors, including:

Non-Response Error: A non-response error is caused by the differences between the people who choose to participate compared to the people who do not participate in a given survey. In other words, it exists when people are given the option to participate but choose not to; therefore, their survey results are not incorporated into the data.Estimates obtained after nonresponse has been observed and imputation has been used to deal with this nonresponse are usually not equivalent to the estimates that would have been obtained had all the desired values been observed without error. The difference between these two types of estimates is called the nonresponse error. There are two types of non-response errors: total and partial.**Total non-response error** occurs when all or almost all data for a sampling unit are missing. This can happen if the respondent is unavailable or temporarily absent, the respondent is unable to participate or refuses to participate in the survey, or if the dwelling is vacant. **Partial non-response error** occurs when respondents provide incomplete information. For certain people, some questions may be difficult to understand, they may refuse or forget to answer a question. Poorly designed questionnaires or poor interviewing techniques can also be reasons which result in partial nonresponse error.

Measurement Error: Measurement error, also called response error, is the difference between measured values and true values. It consists of bias and variance, and it results when data are incorrectly requested, provided, received or recorded. It also refers to all errors relating to the measurement of each sampling unit, as opposed to errors relating to how they were selected. The error often arises when there are confusing questions, low-quality data due to sampling fatigue and low-quality measurement tools. Measurement error is the difference between the measurements you obtain and the truth. This error comes up at many points throughout the research process, from the development of your survey to analyzing your findings. measurement error can be introduced by the interviewer, the questionnaire, or the respondent. Examples of measurement errors from the interviewer and questionnaire might include faulty wording of questions; bias in representative graphics materials; unintentional interviewer modification of the question's wording; interviewer misinterpretation or misrecording of the response. on the respondent side, measurement error includes the way a respondent interprets the question, and the respondent giving incorrect information.

Interviewer Error: Interviewer error occurs when the interviewer makes an error when recording a response. In qualitative research, an interviewer may lead a respondent to answer a certain way. In quantitative research, an interviewer may ask the question differently, which leads to a different result.

Adjustment Error: An adjustment error describes a situation where the analysis of the data adjusts it so that it is not entirely accurate. Forms of adjustment error include errors with weighting the data, data cleaning, and imputation.

Processing Error: A processing error arises when there is a problem with processing the data that causes an error of some kind.It includes all data processing activities after collection and before estimation, such as errors in data capture, coding, editing and tabulation of the data as well as in the assignment of survey weights. An example will be if the data were entered incorrectly or if the data file is corrupt.

Coverage Error: This **error** consists of omissions (undercover), erroneous inclusions, duplications and misclassifications (overcoverage) of units in the survey frame. Since it affects every estimate produced by the survey, they are one of the most important types of error. In the case of a census, it may be the main source of error. Coverage error can have both spatial and temporal dimensions and may cause bias in the estimates. The effect can vary for different subgroups of the population. This error tends to be systematic and is usually due to under coverage, which is why it's important to reduce it as much as possible.

Minimizing Non-Sampling Errors

Reducing non-sampling error is not as easily achieved as reducing sampling error. With sampling error, we can reduce the risk of error by simply increasing the sample size. It will not work for non-sampling error, which is often very difficult to detect and eliminate. To effectively reduce non-sampling errors, careful consideration must be taken by those designing the study to ensure the validity of the results. As such, a researcher may design a mechanism into the study to reduce the error while subsequently not introducing another error.Non-sampling error is broad because the types of marketing studies conducted are various. The following are general techniques used to minimize non-sampling error, but remember, an in-person study has different factors than a survey or questionnaire.

Randomize selection to eliminate bias: Select participants based on a random factor, like choosing every fourth person on a list.

Train your team: If the study is conducted by a researcher, either uses the same researcher or be sure to train your team on the procedure. Training and experience are essential.

Perform an external record check: Human error occurs in entering data. Have an external source check your records and confirm their consistency with written results.

8.6 SUMMARY

Research design greatly depends on the nature of the research questions. Whenever a researcher decides to collect data by deploying surveys to their participants, what matters the most are the survey questions that are placed tactfully, to gather meaningful insights.In other words, knowing what kind of information a study should be able to collect is the first step in determining how to carry out the rest of the study. What steps need to be included and what can be given a pass. How to distinguish between qualitative and qualitative research design. A study related to the first approach cross-sectional study and the second approach longitudinal study. Finally, the chapter ends with how sampling and non-sampling errors are generated and how they are minimized.

8.7 GLOSSARY

Qualitative Research: Qualitative research is research using methods such as participant observation or case studies that result in a narrative, descriptive account of a setting or practice.

Quantitative Research: It is the method of research that relies on measuring variables using a numerical system, analyzing these measurements using any of a variety of statistical models, and reporting relationships and associations among the studied ariables.

Cross-Sectional Study: Cross-sectional study design is a type of observational study design. In a cross-sectional study, the investigator measures the outcome and the exposures in the study participants at the same time.

Longitudinal Study: In a longitudinal study subjects are followed over time with continuous or repeated monitoring of risk factors or health outcomes, or both.

8.8 LESSON END EXERCISE

What are sampling and non-sampling errors?

____

____

____

What is the difference between Quantitative and Qualitative Research Design?

____

____

____

What is the difference between Cross-Sectional and Longitudinal Study?

____

____

____

8.9 SUGGESTED READINGS

Srivastava, T. N., Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

Kothari, C.R., Research Methodology Methods and Techniques. Wiley Eastern Limited: New Delhi.

Sadhu, A.N. and A. Singh., Research Methodology in Social Sciences. Sterling Publishers Private Limited: New Delhi.

William G. Z. Business Research Methods. Thomson, India.

UNIT-II **Lesson No. 9**

RELATIONSHIPAMONGDESCRIPTIVE, EXPLORATORY& CAUSAL RESEARCHANDADVANTAGES & DISADVANTAGES OF RESEARCH DESIGN

STRUCTURE

9.1 INTRODUCTION

Research design is a tool that is used in carrying out marketing research. The design is supposed to give in detail the procedures that are supposed to be followed to solve problems that marketing researchers present. The major approaches used in research include exploratory, causal, and exploratory research designs.Information requirements, measures, and scales have to be spelled out clearly. The study also highlights the advantages and disadvantages of research design.

9.2 OBJECTIVES

After going through this unit, you should be able to describe:-

Relation between various types of research design.

How descriptive, exploratory and causal research designs are related to each other.

Advantages of research design.

Disadvantages of research design.

9.3 RELATIONSHIPAMONGDESCRIPTIVE, EXPLORATORY& CASUAL RESEARCH

Exploratory Research : is the research design that is used to investigate a research problem that is not clearly defined or understood. It provides researchers with a deeper understanding of a research problem and its context before further research can be carried out. Therefore, exploratory research acts as a groundwork to further research and is a useful tool when dealing with research problems that have not been properly investigated in the past. This research design is also referred to as interpretive research, and helps answer questions like “what”, “where”, and “how”. A key feature of the exploratory research design is that it is unstructured and therefore very flexible.

Exploratory research design is normally used in research whose purpose involves inquiry into new products that should be developed, how the product appeal will enhance its advertising, and how the existing services can be improved.Research questions in exploratory research design border on the alternative ways that exist that can be used for example to provide lunch for school children, the kind of benefits people stand to get from the product, and the nature of dissatisfaction the customer may be getting from the product in the case of marketing research. The hypothesis here has unknown constructs. The hypothesis also suspects that the major problem in the study could be personalization.

For Example,Let’s assume a researcher wants to study the effects of social media on a teenager’s attention span. Before going forth with the investigation itself, the researcher may choose to conduct surveys or interviews using open-ended questions. The responses will be collected from the target audience which, in this case, comprises those who fall between the ages of 13 to 19. The data collected will provide the researcher with meaningful insights that will help them frame a more specific and realistic research question that can be investigated effectively.

Descriptive Research : It is a design used to describe a phenomenon and its different characteristics. It is concerned with gaining a deeper understanding of what the phenomenon is rather than why or how it takes place. It, therefore, describes the subject of the research without addressing why it happens.Without a thorough understanding of a research problem, researchers cannot effectively answer it. This enforces the importance of descriptive research as it gives researchers a proper understanding of a research problem before they begin investigating it.

When using descriptive research, researchers do not manipulate any variables. Instead, the observational method is used to observe and measure different variables and identify any changes and correlations depicted in the data collected. This design is mainly used to describe a behavior or type of subject. It is neither intended to look for specific relationships between variables nor does it correlate with variables. Its major setback is that it cannot identify the cause because its setting is completely natural.Moreover, it has all the variables present. Descriptive research design is an asset to a researcher because a lot of information can be acquired through the description. For purposes of identification of variables and hypothetical constructs which can be subjected to further investigations using other means, descriptive research design comes in handy.

For Example, Let's take an example of a shoe company that is trying to conduct market research to understand the shoe purchasing trends in the city of India. Before delving into the investigation itself, they may want to first conduct descriptive research to understand which variables and statistics are relevant to their company and therefore which variables and statistics need to be investigated. The descriptive research conducted will provide the company with a deeper understanding of the research topic before the investigation can be commenced.

Causal Research : It is a type of conclusive research, which attempts to establish a cause-and-effect relationship between two or more variables. An easy replicate of the process, Causal Research is widely employed by several companiesto find the connection between their customers and the changing prices of their goods.Thus, this method of research can be used by companies to help craft favorable outcomes for themselves. Such assessment can help businesses navigate their future with fewer interruptions and also help them plan better for various situations. It assists in determining the impact of a change in process and existing methods. It is easy to narrow down the cause and effect relationship by making sure that both variables are not affected by any force other than each other. To maintain accuracy, other variables are assumed to be constant. It can help determine the exact impact an individual variable has on another.

For Example, A researcher is trying to study the effects of alcohol consumption on health. They select a sample group consisting of people who consume different amounts of alcohol, and then also observe different metrics that are indicators of health. This is an example of a causal research design as the researcher is investigating the cause-and-effect relationship between alcohol consumption and a person's health.

A COMPARISON: THE DIFFERENCES BETWEEN EXPLORATORY, DESCRIPTIVEAND CAUSALRESEARCH?

Now there is an understanding of these three research designs, the following are their differences.

S. No.	Exploratory Research	Descriptive Research	Causal Research
1.	Exploratory research design is to discover ideas and insights.	The descriptive research design involves describing market aspects and functions.	A causal research design tries to determine the cause-effect relationships in the research one is conducting.
2.	The exploratory design is known due to its flexibility, its adaptability, and the front end of total research design.	The descriptive research design is marked by the past formulation of specific hypotheses from observations and studies. This design is pre-planned.	The causal research approach can proceed through one or more independent variables. Other mediating variables that fall under this design can also be controlled.
3.	The exploratory research approach entails the use of surveys, case studies,	The descriptive research approach uses information from	The Causal research design strictly uses experiments.

	information from other studies, and qualitative analyses.	other studies, panels, analyses, and observations.	
4.	exploratory research is highly unstructured and provides a lot of flexibility as it is generally the first step in any research process and is therefore in the early stages of decision making.	Descriptive research is conducted after explorative research and its research design has more structure thanexploratory design but less structure than the causal design.	Causal research has a highly structured and rigid research design and is generally conducted in the later stages of decision making.
5.	In exploratory research, the key research statement is the research question itself.	In descriptive research, the key research statement is the research question itself.	In causal research, the key research statement is generally the research hypothesis.

Table 9.1: Comparison between Exploratory, Descriptive and Causal Research

9.3 ADVANTAGES OFRESEARCH DESIGN

This type of research always requires a lot of time, dedication and effort to get powerful insights, but it's a worthwhile investment because it will get us there:

Real and consistent data: Dataon which to base the designs. In this way, we will not be influenced by fashions or our own beliefs, but we will have evidence obtained in the analysis.

Empathized: Clientscan put themselves in their place and see things from their point of view. It is about getting to know your user so thoroughly that you know better and what is required for their needs.

Focused: It will be able to prioritize functionalities during the project. It often happens that during the creation process, sometimes, the uncertainty becomes more evident and it is in these moments that the research will help us put focus on the work.

Time: Research design helps the researcher to ensure the project research is on the proper schedule and consumes less time. It helps in the proper planning of the resources and their procurement at right time.

Systematic research: Research design helps the researcher toprepare the research design and to carry out research properly and systematically.

Proper documentation: Research design depicts better documentation of the various activities while the project work is going on.

Research satisfaction: Research design provides satisfaction and confidence, accompanied by a sense of success from the beginning of the work of the research project.

Products and services of value to the user: At the end of the whole process you will have created a surprising experience for the user that will adapt to their needs and solve their problems.

9.4 DISADVANTAGES OFRESEARCH DESIGN

The following are the disadvantages of research design:

Knowledge related to the data: The opportunities that are possible throughresearch become possible due to the industry-related expertise that researchers have. If an interviewer is unfamiliar with the subject involved, then they cannot ask the relevant questions that are necessary to get the results that are desired. Every research method requires some level of personal knowledge from those involved.

Identify unspoken data points: Researchers are trusted to connect all of the dots when they gather information through this process. They must read non-verbal cues, have an empathetic response, and understand the situation behind each action a person takes when going through this process. Missing this data can lead to incomplete results, false conclusions, or worse.

Requires repetitive research efforts: As the research focuses on the sample size to develop a rich data profile, the complexity of the questions involved becomes a potential drawback. When there is a critical decision to be made, then discovering the potential consequences on both sides of that equation can help to develop a better SWOT profile.

Replicate the results using this method: Theories become fact because third-party researchers can verify the information that the initial studies produce. When others can duplicate your work, it becomes useful for the betterment of that demographic.

Lengthy process: Research design becomes a lengthy process due to the sequence of steps to be followed to reach the goal.

Biased results: If the data was not collected according to a suitable sampling technique result always be inaccurate.

Costly: Primary research can be used to yield more data and survey a larger sample. This data can be organized into information and used to provide insights and knowledge from a subject thus, it becomes an expensive process.

9.5 SUMMARY

This chapter started with a brief introduction to the various types of research design like Exploratory, Descriptive and Causal Research. The chapter also highlights the relationship between these types of research designs. Last but not least the advantages and disadvantages of the research design were also briefed.

9.6 GLOSSARY

Descriptive Research: Descriptive research aims to accurately and systematically describe a population, situation or phenomenon. It can answer what, where. when and how questions. but not why questions. A descriptive research design can use a wide variety of research methods to investigate one or more variables.

Exploratory Research: Exploratory studies usually create scope for future research and future research may have a conclusive design.

9.7 LESSON END EXERCISE

Briefly define the relationship betweenExploratory, Descriptive and Causal Research.

____

____

____

Explain the advantages of research design.

____

____

____

Explain the disadvantages of research design.

____

____

____

9.8 SUGGESTED READINGS

Brown, F.E. Business Research, a structure for decision making, Addison - Wesley Publishing Company.

Cooper, D.R., Schildler, P.S. Business Research Methods. Tata Mc Graw Hill, New Delhi.

Srivastava, T. N., Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

Stockton and Clark, Introduction to Business and Economic Statistics. D.B. Taraporevala Sons and Co. Private Limited, Bombay.

UNIT-II **Lesson No. 10**

ONCEPTAND RELEVANCE OFSAMPLINGDESIGN IN BUSINESS RESEARCH

STRUCTURE

10.1 INTRODUCTION

When the researcherresearches a group of people, it's rarely possible to collect data from every person in that group. Instead, you select a sample. The sample is the group of individuals who will participate in the research. To draw valid conclusions from your results, you have to carefully decide how you will select a sample that is representative of the group as a whole. Sampling may be defined as the procedure in which a sample is selected from an individual or a group of people of a certain kind for research purposes. In sampling, the population is divided into several parts called sampling units

10.2 OBJECTIVES

After going through this unit, you should be able to describe:-

- Concept of sampling in business research
- Steps and characteristics of sampling
- Types of sampling
- Advantages & disadvantages of sampling

10.3 CONCEPT OF SAMPLING DESIGN IN BUSINESS RESEARCH

Sampling is a process used in statistical analysis in which a predetermined number of observations are taken from a larger population. The methodology used to sample from a larger population depends on the type of analysis being performed, but it may include simple random sampling or systematic sampling. **Sampling** helps a lot in research. It is one of the most important factors which determines the accuracy of your research/survey result. If anything goes wrong with your sample then it will be directly reflected in the final result. There are a lot of techniques that help us to gather samples depending upon the need and situation.

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 also leads to a procedure to tell the number of items to be included in the sample i.e., the size of the sample. Hence, sample design is determined before the collection of data. Among various types of sample design techniques, the researcher should choose that samples are reliable and appropriate for his research study.

Basic terms related to sampling

Population: Population is the collection of elements which has some of the other characteristics in common. The number of elements in the population is the size of the population.

Sample: The sample is the subset of the population.A sample is defined as a smaller set of data that a researcher chooses or selects from a larger population by using a pre-defined selection method.

Sampling: The process of selecting a sample is known as sampling. The number of elements in the sample is the sample size.

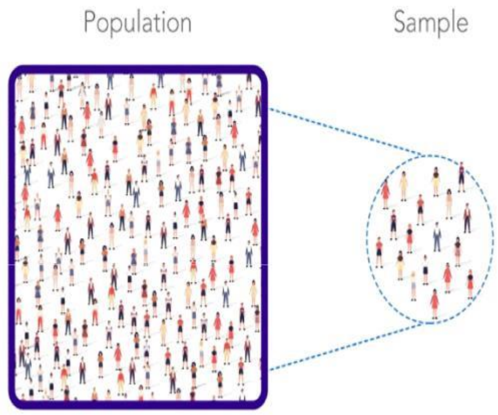


Figure 10.1: Process of Sampling

10.3.1 Steps in Sampling

There are various steps that the researcher should follow.

Type of universe: In the first step the researcher should clarify and should be an expert in the study of the universe. The universe may be finite or Infinite.

Sampling unit: A decision has to be taken concerning a sampling unit before selecting a sample. The sampling unit may be a geographical one such as state, district, village, etc., or a construction unit such as a house, flat, etc., or it may be a social unit such as family, club, school, etc., or it may be an individual.

Source list: The source list is known as the 'sampling frame' from which the sample is to be drawn. It consists of the names of all items of a universe. Such a list would be comprehensive, correct, reliable and appropriate and the source list should be representative of the population.

Size of sample: The size of a sample refers to the number of items to be selected from the universe to constitute a sample. The selection of sample size is a headache for the researcher. The size should not be too large or too small rather it should be optimum. An optimum sample fulfills the requirements of efficiency, representativeness, reliability, and flexibility. The parameters of interest in a research study must be kept in view while deciding the size of the sample. Cost factors i.e., budgetary conditions should also be taken into consideration.

Sampling procedure: In the final step of the sample design, a researcher must decide the type of sample that will be used and decide the techniques to be used in selecting the items for the sample.

10.3.2 Characteristics of Good Sample Design

The characteristics of a good sample are as follows:

Sample design must result in a truly representative sample,

Sample design must be such that it results in a small sampling error,

Sampling design must be viable in the context of funds available for the research study,

Sample design must be such that systematic bias can be controlled in a better way, and

The sample should be such that the results of the sample study can be applied, in general, to the universe with a reasonable level of confidence.

10.4 RELEVANCE OF SAMPLING DESIGN IN BUSINESS RESEARCH

Everyone who has ever worked on a research project knows that resources are limited; time, money and people never come in an unlimited supply. For that reason, most research projects aim to gather data from a sample of people, rather than from the entire population (the census is one of the few exceptions). This is because sampling allows researchers to:

Save Time: Contacting everyone in a population takes time. And, invariably, some people will not respond to the first effort at contacting them, meaning researchers have to invest more time in follow-up. Random sampling is much faster than surveying everyone in a population and obtaining a non-random sample is almost always faster than random sampling. Thus, sampling saves researchers lots of time.

Save Money: The number of people a researcher contacts is directly related to the cost of a study. Sampling saves money by allowing researchers to gather the same answers from a sample that they would receive from the population. Non-random sampling is significantly cheaper than random sampling because it lowers the cost associated with finding people and collecting data from them. Because all research is conducted on budget, saving money is important.

Collect Richer Data: Sometimes, the goal of the research is to collect a little bit of data from a lot of people (e.g., an opinion poll). At other times, the goal is to collect a lot of information from just a few people (e.g., a user study or ethnographic interview). Either way, sampling allows researchers to ask participants more questions and to gather richer data than does contacting everyone in a population.

Market Research: Expanding the customer base may mean finding new market nichers. A market niche is a group of individuals that have similar demographics such as age gender, income, geographic location, marital status and education levels. Sampling the population of that niche lets the business know whether this is a lucrative prospect and should be pursued or is lukewarm at best and can be put on the back burner.

New Product Development: One product development philosophy is to invent a new product and then find a market for it. The other philosophy is to listen to the market and develop a product that solves its problems or fills its needs. Both philosophies benefit from sampling the market the potential customers. For example, if you manufacture fruit drinks, you may find out that most of the market doesn't care one way or the other about using high - fructose corn syrup as an ingredient, but a smallar market segement does. that smaller segment is where to market your sugar-sweetened fruit drink.

Customer Satisfaction: A very small business most likely could ask all its customers about their level of satisfaction. However, that's not possible with a business that has several hundred customers a day, such as a retail website or restaurant. In that case, sampling the customers is the way to go. Sample at several different times and days to get the entire picture. Service businesses can use customer sampling as well. It may make sense to hire an outside firm to conduct the sampling and administer the survey, so the customers feel comfortable telling how they really feel, not what they think you want to hear.

Sampling Management: The population of the sample has to reflect the makeup of **the** general population it's being pulled from, otherwise the results aren't meaningful. The survey if one is being used shouldn't lead the participants to the answers the small-business person expects. The sampling should be completed within a short window, not over a long period of time, since the sample might change or an event may occur that changes their answers.

10.5 TYPES OF SAMPLING DESIGN

Sampling forms an integral part of the research design as this method derives the quantitative data and the qualitative data that can be collected as part of a research study. Sampling methods are characterized by two distinct approaches: probability sampling and non-probability sampling. The difference lies between the above two is whether the sample selection is based on randomization or not. With randomization, every element gets an equal chance to be picked up and to be part of the sample for the study.

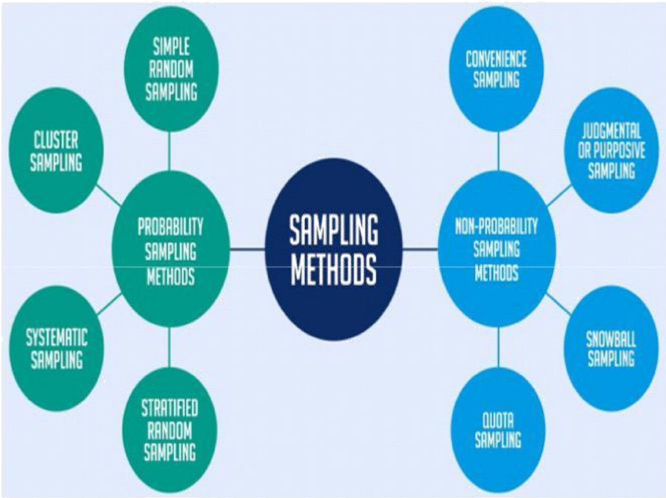


Figure 10.2: Type of Sampling

10.5.1 Probability Sampling

Probability sampling is also known as 'random sampling' or 'chance sampling'. Under this sampling design, every item of the universe has an equal chance of inclusion in the sample. It is, so to say, a lottery method in which individual units are picked up from the whole group not deliberately but by some mechanical process. Here it is blind chance alone that determines whether one item or the other is selected. The results obtained from probability or random sampling can be assured in terms of probability i.e., we can measure the errors of estimation or the significance of results obtained from a random sample, and this fact brings out the superiority of random sampling design over the deliberate sampling design. Random sampling ensures the law of Statistical Regularity which states that if on average the sample chosen is a random one, the sample will have the same composition and characteristics as the universe.

Probability sampling can be further classified into four distinct types of samples. They are:

Simple random sampling: It is the most straightforward way of selecting a sample is simpler and omsampling. In this method, each member has an equal chance of being a part of the study. The objects in this sample population are chosen purely on a random basis, and each member has the same probability of being selected. For example, if a university dean would like to collect feedback from students about their perception of the teachers and level of education, all 1000 students in the University could be a part of this sample. Any 100 students can be selected at random to be a part of this sample.

Cluster sampling: Cluster sampling is a type of sampling method where the respondent population is divided into equal clusters. Clusters are identified and included in a sample based on defining demographic parameters such as age, location, sex, etc. This makes it extremely easy for a survey creator to derive practical inferences from the feedback. For example, if the FDA wants to collect data about adverse side effects from drugs, they can divide the mainland US into distinctive clusters, like states. Research studies are then administered to respondents in these clusters. This type of generating a sample makes the data collection in-depth and provides easy to consume and act upon, insights.

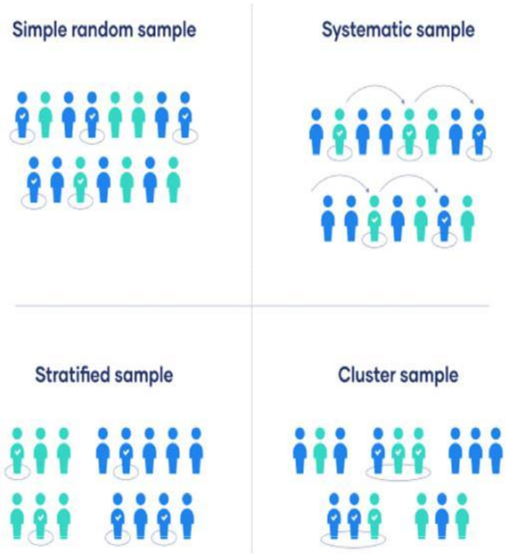


Figure 10.3:Types of Probability Sampling

Systematic Sampling: Systematic sampling is a sampling method where the researcher chooses respondents at equal intervals from a population. The approachto selecting the sample is to pick a starting point and then pick respondents at a pre-defined sample interval. For example, while selecting 1,000 volunteers for the Olympics from an application list of 10,000 people, each applicant is given a count of 1 to 10,000. Then starting from 1 and selecting each respondent with an interval of 10, a sample of 1,000 volunteers can be obtained.

Stratified Random Sampling: Stratified random samplingis a method of dividing the respondent population into distinctive but pre-defined parameters in the research design phase. In this method, the respondents don't overlap but collectively represent the whole population. For example, a researcher looking to analyze people from different socioeconomic backgrounds can distinguish respondents by their annual salaries. This forms smaller groups of people or samples, and then some objects from these samples can be used for the research study.

10.5.2 Non-Probability Sampling

Non-probability sampling is a sampling procedure that does not afford any basis for estimating the probability that each item in the population has of being included in the sample. Non-probability sampling is also known by different names such as deliberate sampling, purposive sampling and judgment sampling. In this type of sampling, items for the sample are selected deliberately by the researcher; his choice concerning the items remains supreme. In other words, under non-probability sampling, the organizers of the inquiry purposively choose the particular units of the universe for constituting a sample on the basis that the small mass that they so select out of a huge one will be typical or representative of the whole.

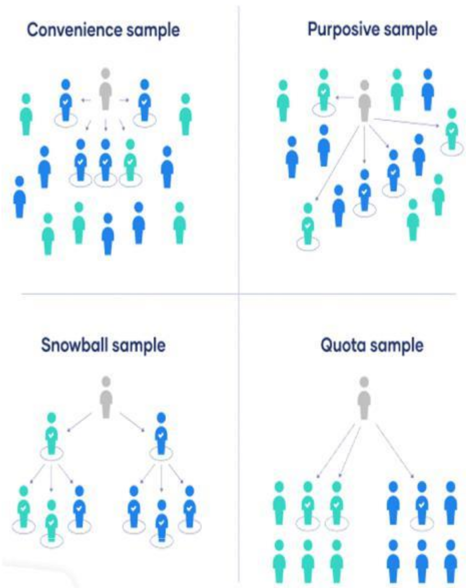


Figure 10.4 : Types of Non-Probability Sampling

Non- Probability sampling can be further classified into four distinct types of samples. They are:

Convenience sampling: Conveniencesampling, in easy terms, stands for the convenience of a researcher accessing a respondent. There is no scientific method of deriving this sample. Researchers have nearly no authority over selecting the sample elements, and it's purely based on proximity and not representativeness.This non-probability sampling method is used when there is time and cost limitations in collecting feedback. For example, researchers are conducting a mall-intercept survey to understand the probability of using a fragrance from a perfume manufacturer. In this sampling method, the sample respondents are chosen purely on their proximity to the survey desk and their willingness to participate in the research.

Judgemental/purposive sampling: The judgemental or purposive sampling method is a method of developing a sample purely on the basis and discretion of the researcher purely based on the nature of the study along with his/her understanding of the target audience. In this sampling method, people who only fit the research criteria and end objectives are selected, and the remaining are kept out.For example, if the research topic is understanding what University a student prefers for a master's, if the question asked is "Would you like to do your Master's?" anything other than a response, "Yes" to this question, everyone else is excluded from this study.

Snowball sampling: Snowball sampling or chain-referral sampling is defined as a non-probability sampling technique in which the samples have traits that are rare to find. This is a sampling technique, in which existing subjects provide referrals to recruit samples required for a research study.For example, while collecting feedback about a sensitive topic like AIDS, respondents aren't forthcoming with information. In this case, the researcher can recruit people with an understanding or knowledge of such people and collect information from them or ask them to collect information.

Quota sampling: Quota samplingis a method of collecting a sample where the researcher has the liberty to select a sample based on its strata. The primary characteristic of this method is that two people cannot exist under two different conditions. For example, when a shoe manufacturer would like to understand from Millenials their perception of the brand with other parameters like comfort, pricing, etc. It selects only females who are millennials for this study as the research objective is to collect feedback about women's shoes.

10.6 ADVANTAGES OFSAMPLINGDESIGN

Sampling ensures convenience, collection of intensive and exhaustive data, suitability in limited resources and better rapport. In addition to this, sampling has the following advantages also.

Low cost of sampling: If data were to be collected for the entire population, the cost will be quite high. A sample is a small proportion of a population. So, the cost will be lower if data is collected from a sample of the population which is a big advantage.

Less time-consuming in sampling: Use of sampling takes less time also. It consumes less time than the census technique. Tabulation, analysis, etc., take much less time in the case of a sample than in the case of a population.

Scope of sampling is high: The investigator is concerned with the generalization of data. To study a whole population to arrive at generalizations would be impractical.Some populations are so large that their characteristics could not be measured. Before the measurement has been completed, the population would have changed. But the process of sampling makes it possible to arrive at generalizations by studying the variables within a relatively small proportion of the population

Accuracy of data is high: Having drawn a sample and computed the desired descriptive statistics, it is possible to determine the stability of the obtained sample value. A sample represents the population from which it is drawn. It permits a high degree of accuracy due to a limited area of operations. Moreover, careful execution of fieldwork is possible. Ultimately, the results of sampling studies turn out to be sufficiently accurate.

Organization of convenience: Organizational problems involved in sampling are very few. Since the sample is of small size, vast facilities are not required. Sampling is therefore economical in respect of resources. The study of samples involves less space and equipment.

Intensive and exhaustive data: In sample studies, measurements or observations are made of a limited number. So, intensive and exhaustive data are collected.

Suitable in limited resources: The resources available within an organization may be limited. Studying the entire universe is not viable. The population can be satisfactorily covered through sampling. Where limited resources exist, the use of sampling is an appropriate strategy while conducting marketing research.

Better rapport: An effective research study requires a good rapport between the researcher and the respondents. When the population of the study is large, the problem of rapport arises. But manageable samples permit the researcher to establish adequate rapport with the respondents.

10.7 DISADVANTAGES OFSAMPLINGDESIGN

The reliability of the sample depends upon the appropriateness of the sampling method used. The purpose of sampling theory is to make sampling more efficient. But the real difficulties lie in the selection, estimation and administration of samples.Disadvantages of sampling may be discussed under the heads:

Chances of bias: The serious limitation of the sampling method is that it involves biased selection and thereby leads us to draw erroneous conclusions. Bias arises when the method of selection of sample employed is faulty. Relative small samples properly selected may be much more reliable than large samples poorly selected.

Difficulties in selecting a truly representative sample: Difficulties in selecting a truly representative sample produce reliable and accurate results only when they are representative of the whole group. Selection of a truly representative sample is difficult when the phenomena under study are complex. Selecting good samples is difficult.

Inadequate knowledge of the subject: The use of the sampling method requires adequate subject-specific knowledge ofthe sampling technique. Sampling involves statistical analysis and calculation of probable error. When the researcher lacks specialized knowledge in sampling, he may commit serious mistakes. Consequently, the results of the study will be misleading.

The changeability of units: When the units of the population are not homogeneous, the sampling technique will be unscientific. In sampling, though the number of cases is small, it is not always easy to stick to the, selected cases. The units of the sample may be widely dispersed.Some of the cases of the sample may not cooperate with the researcher and some others may be inaccessible. Because of these problems, all the cases may not be taken up. The selected cases may have to be replaced by other cases. The changeability of units stands in the way of the results of the study.

Impossibility of sampling: Deriving a representative sample is difficult when the universe is too small or too heterogeneous. In this case, a census study is the only alternative. Moreover, in studies requiring a very high standard of accuracy, the sampling method may be unsuitable. There will be chances of errors even if samples are drawn most carefully.

10.8 SUMMARY

A **population**is any well-defined set of units of analysis.A **sample** is any subset of units collected in some manner from a population. Once a sample has been collected, one can derive **sample statistics** that measure the characteristics of the sample to **estimate** the value of **population parameters** that describe the characteristics of a population.The chapter discusses four examples of **probability samples**defined as samples for which each **element** in the population has a known probability of inclusion in the sample. The particular population from which a sample is drawn is called a **sampling frame,** and it must be specified clearly.The chapter discusses four examples of nonprobability samplesdefined as samples for which each element in the population has an unknown probability of inclusion in the sample. These sampling techniques, while less representative, are used to collect data when probability samples are not feasible.

10.9 GLOSSARY

Sampling : Sampling is a process in statistical analysis where researchers take a predetermined number of observations from a larger population.

Probability Sampling: Probability sampling specifies to the researcher that each segment of a known population will be represented in the sample. Probability samples lend themselves to rigorous analysis to determine the likelihood and possibility of bias and error

Non-Probability Sampling: Non-probability sampling is a method of selecting unit from a population using a subjective (i.e. non-random) method. Since non-probability sampling does not require a complete survey frame, it is a fast, easy and inexpensive way of obtaining data.

10.10 LESSON END EXERCISE

What is Sampling?

____

____

____

How many types of sampling?

____

____

____

Briefly define probability sampling techniques

____

____

____

Briefly define non-probability sampling techniques.

____

____

____

10.11 SUGGESTED READINGS

Boyd, Westfall, and Stasch, Business Research Text and Cases. All India Traveller Bookseller, New Delhi.

Brown, F.E. Business Research, a structure for decision making. Addison - Wesley Publishing Company.

Cooper, D.R., Schildler, P.S. Business Research Methods. Tata Mc Graw Hill, New Delhi.

Kothari, C.R. Research Methodology - Methods and Techniques. Wiley Eastern Ltd.

Stockton and Clark, Introduction to Business and Economic Statistics. D.B. Taraporevala Sons and Co. Private Limited, Bombay.

UNIT-III **Lesson No. 11**

PROCESSING OFDATA: FEEDINGAND CODINGOFDATA; RELIABILITYAND VALIDITY OF DATA

STRUCTURE

11.1 INTRODUCTION

After the data have been collected, the researcher turns to the task of analysing them. The survey data collected from the field should be processed and analysed as indicated in the research plan. Data processing primarily involves editing, coding, classification and tabulation of data, so that it becomes amenable for data analysis. This is essential for a scientific study and for ensuring that we have all relevant data for making contemplated comparisons and analysis. The analysis of data requires a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences. The unwieldy data should necessarily be condensed into a few manageable groups and tables for further analysis. Thus, a researcher should classify the raw data into some purposeful and usable categories. Coding operation is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted. Editing is the procedure that improves the quality of the data for coding. With coding the stage is ready for tabulation. Tabulation is a part of the technical procedure wherein the classified data are put in the form of tables. The mechanical devices can be made use of at this juncture. A great deal of data, especially in large inquiries, is tabulated by computers. Computers not only save time but also make it possible to study a large number of variables affecting a problem simultaneously. Analysis work after tabulation is generally based on the computation of various percentages, coefficients, etc., by applying various well defined methods or techniques. Supporting or conflicting with original new hypotheses should be subjected to tests of significance to determine with what validity data can be said to indicate any conclusion(s).

Through data processing, companies can gain valuable insights, make informed business decisions and secure a competitive edge by narrowing down data to obtain the most accurate, dependable information. When you know the importance of data processing in quantitative research and can successfully implement these methods, you can develop better outcomes for your business. Therefore, this lesson concentrates on various aspects of data processing.

11.2 OBJECTIVES

After studying this lesson, you will be able to:

Know the concept of data processing.
Know the procedure of data feeding/entry.
Methods for processing data in quantitative studies.
Design the coding scheme for a questionnaire.
How to code data.
Know the concept and types of reliability of data.
Know the concept and types of validity of data.

11.3 PROCESSINGOFDATA

If you were actually doing a research study, you would by now have reached a stage where you have either extracted or collected the required information. The next step is what to do with this information. How do you find the answers to your research questions? How do you make sense of the information collected? How do you prove or disprove your hypothesis if you had one? How should the information be analysed to achieve the objectives of your study? To answer these questions you need to subject your data to a number of procedures that constitute the core of data processing. These procedures are the same whether your study is quantitative or qualitative, but what you do within each procedure is different. For both types of study you need to visualise how you are going to present your findings to your readership in light of its background and the purpose of the study. You need to decide what type of analysis would be appropriate for the readers of your report. It is in light of the purpose of your study and your impression about the level of understanding of your readership that you decide the type of analysis you should undertake. For example, there is no point in doing a sophisticated statistical analysis if your readers are not familiar with statistical procedures. In quantitative research the main emphasis in data analysis is to decide how you are going to analyse information obtained in response to each question that you asked of your respondents. In qualitative research the focus is on what should be the basis of analysis of the information obtained; that is, is it contents, discourse, narrative or event analysis?

11.3.1 Feeding of Data

Data feeding also known as data entry is the process of transcribing information into an electronic medium such as a computer or other electronic device. It can either be performed manually or automatically by using a machine or computer. Most data feeding tasks are time consuming in nature; however data entry is considered a basic, necessary task for most organisations. Data feeding is considered a non–core process for most organisations and is usually performed on data forms such as spreadsheets, handwritten or scanned documents, audio or video. Addition, modification and deletion are the three modes of operation in data entry.

Data feeding jobs do not require any special qualifications, knowledge or talent, and only require accuracy and fast turnaround. As such, data entry jobs are frequently outsourced in order to lower costs. Computers are also used in automated data entry, as they are highly accurate and can be programmed to fetch and transcribe data into the required medium. Accurately keyed data is the base upon which the organisation can perform analyses and make plans. Manual data entry often requires good concentration and focus over a long duration of time, and this can prove physically and mentally challenging for data entry workers.

Although, data entry is often thought of as a time-consuming process, but there are steps you can take to make the process more efficient. Below are some general steps to help you get started:

Assign an ID number to each form or survey to be entered, and write the number at the top of each survey. These can be numbers such as 1, 2, 3, 4, but each number should be used only once, even across different batches of surveys. This will make it much easier to go back and re-enter data if you realize you have made a mistake.

In some instances, like surveys you collect on a recurring basis from the same people (such as teachers), you may wish to continue to assign the same ID for each respondent. You will need to maintain a master list that you can reference for assigning and tracking these ID numbers in the future.

Schedule a large enough block of time to enter an entire batch of data at once. The time needed for this will vary depending on the length of the survey and the number of participants, but it is best to enter all of the information at the same time if possible. This will minimize the chance that you enter the same survey twice, or forget to enter any remaining surveys.

Before you begin entering the data, take time to go through each completed survey and identify questionable responses. See the section on making data entry decisions to identify some common unexpected responses that survey respondents make, and take note of the tips for working with difficult or confusing surveys. By taking time up front to identify potential problem areas, you can make consistent decisions about what you plan to enter in each situation, and you will save time once you begin entering the data.

Data entry sheet, where all of the information from the surveys will be entered from left to right. Please note that open ended questions may not be entered in these databases. This is the first worksheet (see the first tab at the bottom of the spreadsheet as shown in the figure below).

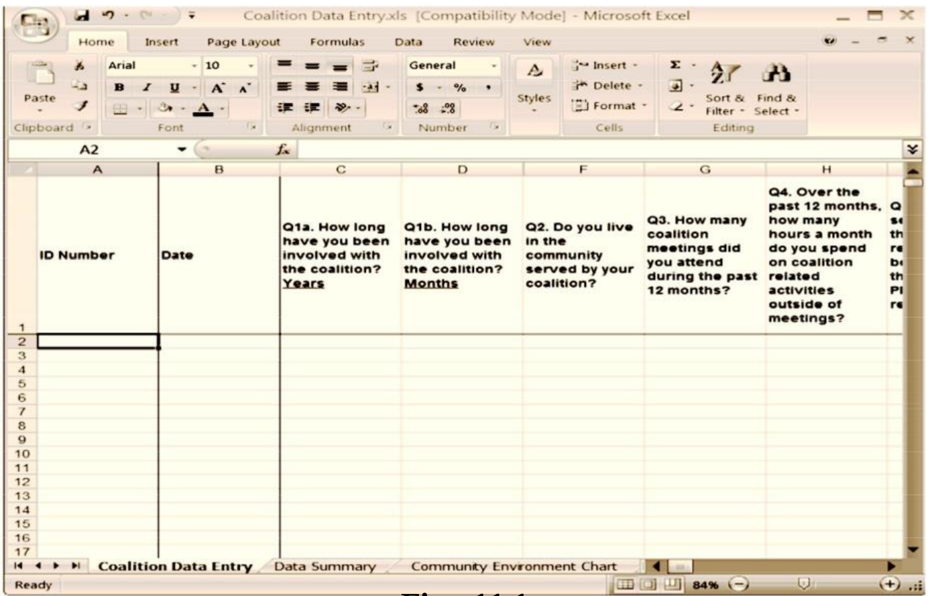


Fig : 11.1

11.3.1.1 : Step by Step Data Feeding Instructions

There are generally three kinds of data you will be entering: 1) continuous numbers (e.g., months, years, ages, and numbers of participants), 2) multiple choice responses (e.g., strongly agree, agree, disagree, strongly disagree), and 3) text data. Each of the Excel files Wilder has created for data entry will reflect the response options of the corresponding survey. For example, for multiple choice questions, the possible categories will be available in the Excel file for you to select from a drop-down menu. For text and continuous items, the cells in the Excel file will likely be blank, and you will simply type into the appropriate cell what was written on the survey.

The following provides a step-by-step process for completing data entry into Excel:

The first time you open the database, be sure you save it in a secure location on your computer. This will preferably be on a password protected network or drive. From the database screen, you can simply click under the File menu (Office 2003 or earlier, see figure on the left below) or the Office button (Office 2007, see figure on the right below) located at the top left corner of the screen. Under this menu, select "Save as" and navigate to the location you would like to save the database. It may be easiest for you to save all of the data files in the same location. Add the current date to the name of the file and click "Save." Adding the date will help you keep track of updates and versions of the database in the future. After saving the file, you are ready to begin entering your data.

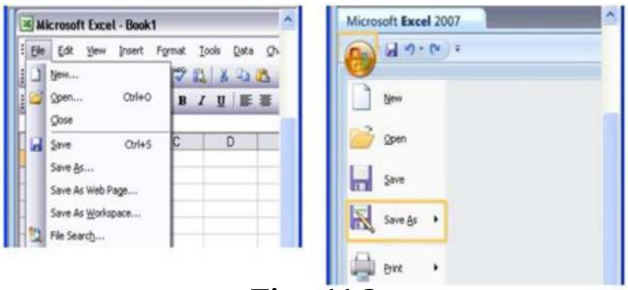


Fig : 11.2

At the top of the spreadsheet, you will see either the full question or the abbreviated name of the corresponding survey item listed along the top row (e.g. Q1. First time or Q1a. How long have you been involved with the coalition?). The key words are to help you keep track of what question you are on as you complete your data entry. They do not represent the full meaning of the question. Be aware that some questions have several parts to them and therefore, multiple data entry columns (e.g., Coalition Q1a. and Q1b., and Responsible Beverage Server Training Q2a. and Q2b., see figure below).

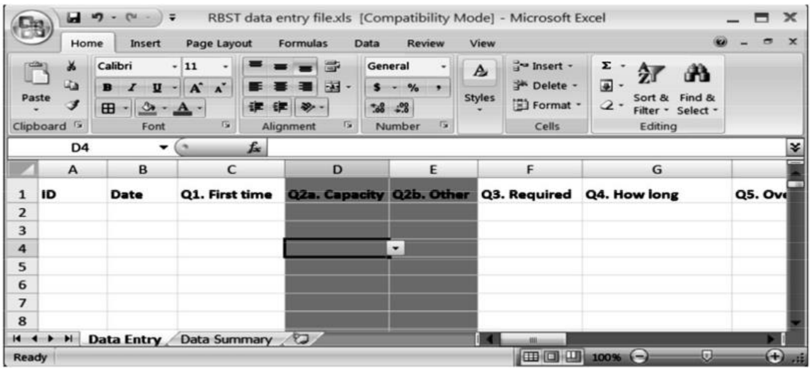


Fig : 11.3

The item names across the top are listed in the same order as the items appear in the survey, to facilitate data entry.

In the first blank cell on the left under the column labeled "ID," enter the ID number assigned to the survey that you are entering.

Type in the date the form or survey was completed, listed at the top of the survey when noted on the form. You can type dates with abbreviations or numbers – they will automatically format to "mo/day/yr" style (see figure below).

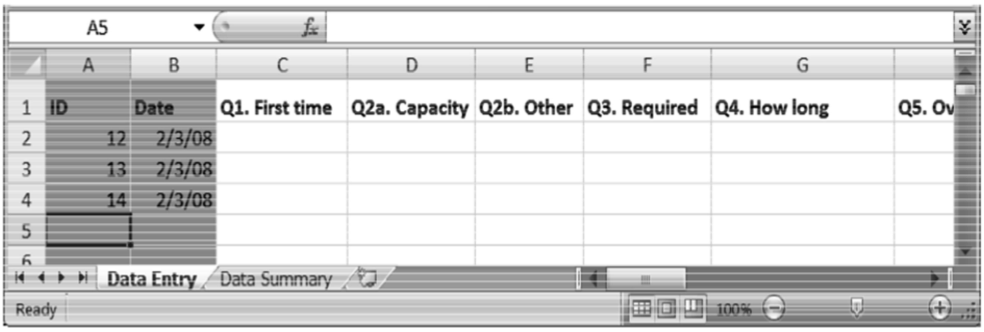


Fig : 11.4

For multiple choice questions, a gray arrow will appear to the right of the cell when your mouse hovers over the cell. Click on this arrow to access a drop-down menu of responses. Select the response that corresponds to the response noted on the survey (see figure below). Note: be sure to select the response carefully and check your answers. It is easy to make errors at this step in the process.

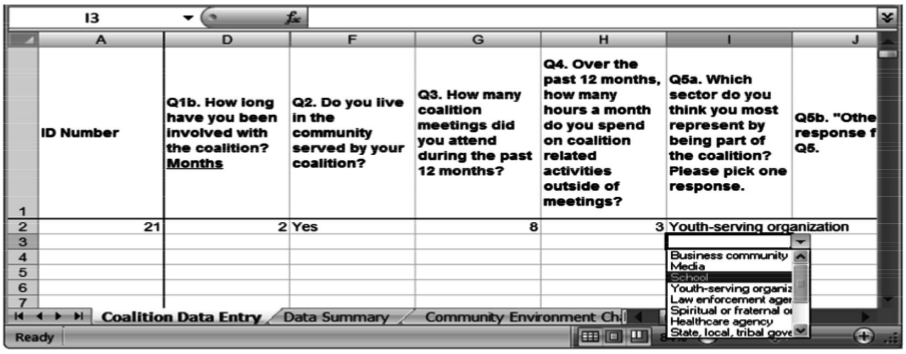


Fig : 11.5

For continuous numbers, simply enter the number as it appears on the written survey (e.g., if the respondent noted he or she has attended 6 meetings to date, type the number "6" into the cell). There are no drop-down menus for these items.

For text data (e.g., the information written on the survey that specifies what a response option of "other" means), there will usually be a column immediately following the column in which you selected the response option "other" that allows you to enter the written text response for "other." If nothing is written in on the survey, you can leave that cell blank.

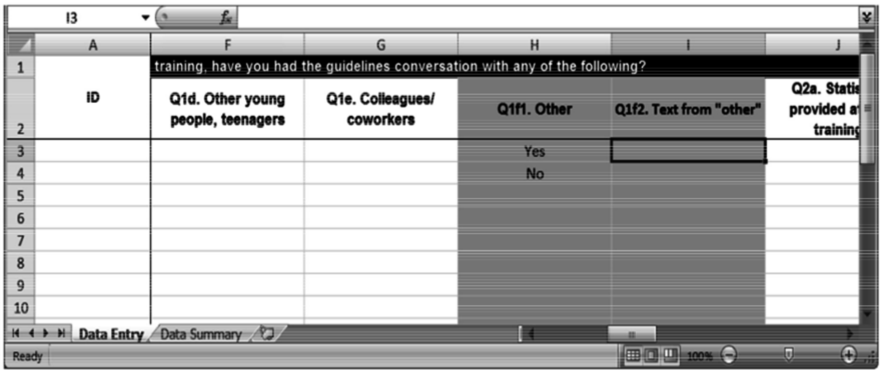


Fig : 11.6

For surveys that contain open-ended responses, those responses will not be included in the summary analysis that is automatically generated in Excel. You may choose to enter more lengthy text responses using a program that is more familiar to you, such as Microsoft Word. Just be sure to keep track of where you enter it, as well as the survey ID and any other relevant information.

Enter all responses from a single survey before entering data from another survey.

Repeat steps for all surveys. 8. Be sure to save your data often during the data entry process

11.3.1.2 Data Feeding Decisions

No matter how clearly a survey is written, there will be some survey respondents who do unexpected things. For example, respondents may choose multiple answers even when asked to choose only one, they may skip questions, or it may just be difficult to understand their intended response. The following are some common issues that you may discover and some ideas for navigating those difficult surveys. Once you have made a decision about how to treat a particular issue, make note of it in a separate document, or even in the margin of these instructions, to reference later and help ensure consistency in your decision-making process.

If data are missing or unintelligible, just leave the space blank in the database. Youn should not try to guess what the respondent might have been thinking.

A participant may respond to a numerical question with a range of numbers (e.g., "1n or 2" or "5-7") or a vague reference (e.g., "a couple" or "several") instead of a single number. In these cases, the response is too vague to translate into a single representative number, so you will simply leave this cell blank.

Sometimes respondents will be unable to choose between two options such as 'agree'n and 'strongly agree', and will select both! Again, you'll have to decide what to do in this situation. Unless it is clear that one of the responses was the intended response (e.g., the other is crossed out or one is especially obviously indicated), you are safer to just leave that space blank in the database. We do not want to try to guess participants' intentions.

For those surveys in which an 'other' category is possible, you will have to decide hown to treat these answers. You may enter them into the spreadsheet as described above in Step 8. However, sometimes respondents choose 'other', but then provide an answer that closely aligns with one of your response categories.

11.3.2 Coding of Data

After collecting data, it must be reduced to some form suitable for analysis so that conclusions or findings can be reported to target population.

For analysing data researchers must decide –

Whether the tabulation of data will be performed by hand or by computer.

How information can be converted into a form that will allow it to be processed efficiently.

What statistical tools or methods will be employed?

Now a day's computers have become an essential tool for the tabulation and analysis of data. Even in simple statistical procedures computer tabulation is encouraged for easy and flexible handling of data. Micro and laptop computers can produce tables of any dimension and perform statistical operations much more easily and usually with far less error than is possible manually. If the data is large and the processing undertaken by computer the following issues are considered.

Data preparation which includes editing, coding, and data entry.

Exploring, displaying and examining data which involves breaking down, examining and rearranging data so as to search for meaningful description, patterns and relationships.

Coding refers to the process of assigning numerals or other symbols to answers so that responses can be put into a limited number of categories or classes. Such classes should be appropriate to the research problem under consideration. They must also possess the characteristic of exhaustiveness (i.e., there must be a class for every data item) and also that of mutual exclusively which means that a specific answer can be placed in one and only one cell in a given category set. Another rule to be observed is that of unidimensionality by which is meant that every class is defined in terms of only one concept. Coding is necessary for efficient analysis and through it the several replies may be reduced to a small number of classes which contain the critical information required for analysis. Coding decisions should usually be taken at the designing stage of the questionnaire. This makes it possible to pre-code the questionnaire choices and which in turn is helpful for computer tabulation as one can straight forward key punch from the original questionnaires. But in case of hand coding some standard method may be used. One such standard method is to code in the margin with a coloured pencil. The other method can be to transcribe the data from the questionnaire to a coding sheet. Whatever method is adopted, one should see that coding errors are altogether eliminated or reduced to the minimum level.

Therefore, for coding, the first level of distinction is whether a set of data is qualitative or quantitative in nature. For qualitative data a further distinction is whether the information is descriptive in nature (e.g. a description of a service to a community, a case history) or is generated through discrete qualitative categories. For example, the following information about a respondent is in discrete qualitative categories: income – above average, average, below average; gender – male, female; religion – Christian, Hindu, Muslim, Buddhist, etc.; or attitude towards an issue – strongly favourable, favourable, uncertain, unfavourable, strongly unfavourable. Each of these variables is measured either on a nominal scale or an ordinal scale. Some of them could also have been measured on a ratio scale or an interval scale. For example, income can be measured in dollars (ratio scale), or an attitude towards an issue can be measured on an interval or a ratio scale. The way you proceed with the coding depends upon the measurement scale used in the measurement of a variable and whether a question is open-ended or closed. In addition, the types of statistical procedures that can be applied to a set of information to a large extent depend upon the measurement scale on which a variable was measured in the research instrument. For example, you can find out different statistical descriptors such as mean, mode and median if income is measured on a ratio scale, but not if it is measured on an ordinal or a nominal scale. It is extremely important to understand that the way you are able to analyse a set of information is dependent upon the measurement scale used in the research instrument for measuring a variable. It is therefore important to visualise-particularly at the planning stage when constructing the research instrument-the way you are going to communicate your findings. How you can analyse information obtained in response to a question depends upon how a question was asked, and how a respondent answered it. In other words, it depends upon the measurement scale on which a response can be measured/classified. If you study answers given by your respondents in reply to a question, you will realise that almost all responses can be classified into one of the following three categories:

Quantitative responses

Categorical responses (which may be quantitative or qualitative)

Descriptive responses (which are invariably qualitative-keep in mind that this is qualitative data collected as part of quantitative research and not the qualitative research)

For the purpose of analysis, quantitative and categorical responses need to be dealt with differently from descriptive ones. Both quantitative and categorical information go through a process that is primarily aimed at transforming the information into numerical values, called codes, so that the information can be easily analysed, either manually or by computers. On the other hand, descriptive information first goes through a process called content analysis, whereby you identify the main themes that emerge from the descriptions given by respondents in answer to questions.

Having identified the main themes, there are three ways that you can deal with them: (1) You can examine verbatim responses and integrate them with the text of your report to either support or contradict your argument.

You can assign a code to each theme and count how frequently each has occurred.

Y ou can combine both methods to communicate your findings. This is your choice, and it is based on your impression of the preference of your readers.

11.4 RELIABILITY OF DATA

The two most important and fundamental characteristics of any measurement procedure are reliability and validity. A sound measurement must meet the tests of validity, reliability and practicality. In fact, these are the three major considerations one should use in evaluating a measurement tool. Validity refers to the extent to which a test measures what we actually wish to measure. Reliability has to do with the accuracy and precision of a measurement procedure. Practicality is concerned with a wide range of factors of economy, convenience, and interpretability. In this lesson we briefly take up the relevant details concerning these tests of reliability and validity. These two principles are discussed below.

The test of reliability is another important test of sound measurement. A measuring instrument is reliable if it provides consistent results. Reliable measuring instrument does contribute to validity, but a reliable instrument need not be a valid instrument. For instance, a scale that consistently overweighs objects by five kgs, is a reliable scale, but it does not give a valid measure of weight. But the other way is not true i.e., a valid instrument is always reliable. Accordingly reliability is not as valuable as validity, but it is easier to assess reliability in comparison to validity. If the quality of reliability is satisfied by an instrument, then while using it we can be confident that the transient and situational factors are not interfering. Two aspects of reliability viz., stability and equivalence deserve special mention. The stability aspect is concerned with securing consistent results with repeated measurements of the same person and with the same instrument. We usually determine the degree of stability by comparing the results of repeated measurements. The equivalence aspect considers how much error may get introduced by different investigators or different samples of the items being studied. A good way to test for the equivalence of measurements by two investigators is to compare their observations of the same events. Reliability can be improved in the following two ways:

By standardising the conditions under which the measurement takes place i.e., we must ensure that external sources of variation such as boredom, fatigue, etc., are minimised to the extent possible. That will improve stability aspect.

By carefully designing directions for measurement with no variation from group to group, by using trained and motivated persons to conduct the research and also by broadening the sample of items used. This will improve equivalence aspect.

11.4.1 Concept of Reliability

Joppe (2000) defines reliability as the extent to which results are consistent over time and an accurate representation of the total population under study is referred to as reliability and if the results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable. It is the extent to which a questionnaire, test, observation or any measurement procedure produces the same results on repeated trials. Reliability relates to numbers or scores and not humans. Therefore, it would be wrong in the field of research to say that someone is reliable. Take for instance, in a mathematics quiz competition for schools. The degree to which the scores of the panel of judges for each contestant agree is an indication of reliability. Similarly, the degree to which an individual’s responses (i.e., their scores) on a survey would stay the same over time is also a measure of reliability.

A measure can be reliable without being valid. A measure cannot be valid without being reliable. Consider a tape rule that always measures my height as 175cm, taller than my true height. This tape-rule (though invalid as it incorrectly assesses height) is perfectly reliable as it consistently measures my height as 175cm, taller than I truly are. A research example of this phenomenon would be a questionnaire designed to assess the impact of corporate social responsibility on employee commitment that asked questions such as, "Do you like to swim?", "What do you like to eat more, pizza or hamburgers?" and "What is your favourite colour?". As you can readily imagine, the responses to these questions would probably remain stable over time, thus, demonstrating highly reliable scores. However, are the questions valid when one is attempting to measure impact of corporate social responsibility on employee commitment? Of course not, as they have nothing to do with the research objective.

11.4.2 Types of Reliability

There are three important aspects of reliability, viz., equivalence, stability and internal consistency (homogeneity). It is important to understand the distinction between these three as it will guide researcher in the proper assessment of reliability given the research protocol.

Test-Retest reliability (also called Stability): This answers the question, "Will the scores or results be stable over the time a measure is administered". Stability is said to occur when the same or similar scores are obtained with repeated testing with the same group of respondents. Stability is assessed through a test-retest procedure that involves administering the same measurement instrument to the same individuals under the same conditions after some period of time. The reliability coefficient is expected to be highly correlated. For example, if a classroom achievement test is administered today and the test is given two weeks later, it is expected to have a reliability coefficient of $r = 0.75$ or more.

When researchers measure a construct that they assume to be consistent across time, then the scores they obtain should also be consistent across time. Test-retest reliability is the extent to which this is actually the case. For example, intelligence is generally thought to be consistent across time. A person who is highly intelligent today will be highly intelligent next week. This means that any good measure of intelligence should produce roughly the same scores for this individual next week as it does today. Clearly, a measure that produces highly inconsistent scores over time cannot be a very good measure of a construct that is supposed to be consistent. Assessing test-retest reliability requires using the measure on a group of people at one time, using it again on the same group of people at a later time, and then looking at test-retest correlation between the two sets of scores. This is typically done by graphing the data in a scatter plot and computing Pearson's r. In general, a test-retest correlation of +.80 or greater is considered to indicate good reliability.

Internal Consistency: A second kind of reliability is internal consistency, which is the consistency of people's responses across the items on a multiple-item measure. In general, all the items on such measures are supposed to reflect the same underlying construct, so people's scores on those items should be correlated with each other. On the Rosenberg Self-Esteem Scale, people who agree that they are a person of worth should tend to agree that that they have a number of good qualities. If people's responses to the different items are not correlated with each other, then it would no longer make sense to claim that they are all measuring the same underlying construct. This is as true for behavioural and physiological measures as for self-report measures. For example, people might make a series of bets in a simulated game of roulette as a measure of their level of risk seeking. This measure would be internally consistent to the extent that individual participants' bets were consistently high or low across trials. This type of reliability also known as parallel forms reliability or equivalence. This answers the question, "Are the two forms of the test or measure equivalent?" If different forms of the same test or measure are administered to the same group; one would expect that the reliability coefficient will be high. Equivalence measures the level of agreement between two or more instruments that are administered at nearly the same point in time. It is measured through a parallel forms procedure in which one administers alternative forms of the same measure to either the same group or different group of respondents. Internal consistency reliability or homogeneity answers the question, "How well does each item measure the content or construct under consideration?" It is an indicator of reliability for a test or measure which is administered once. One expects the correlation between responses to each test item to be highly correlated with the total test score. For example, an employee job satisfaction (attitude scale) or a classroom achievement test which is administered once. Internal consistency is estimated via the split-half reliability index, coefficient alpha (Cronbach, 1951) index.

Like test-retest reliability, internal consistency can only be assessed by collecting and analysing data. One approach is to look at a split-half correlation. This involves splitting the items into two sets, such as the first and second halves of the items or the even- and odd-numbered items. Then a score is computed for each set of items, and the relationship between the two sets of scores is examined. A split-half correlation of +.80 or greater is generally considered good internal consistency.

Inter-Rater Reliability: Different raters, using a common rating form, measure the object of interest consistently. Inter-rater agreement answers the question, "Are the raters consistent in their ratings?" One expects that the reliability coefficient will be high, if the observers rated similarly. For example, three senior sales trainers rating the closing skills of a novice sales representative or master teachers rating the teaching effectiveness of a first or second year teacher.

At this point, it is important to understand the two main questions reliability helps to answer;

What is considered a 'good' or 'adequate' value? and

How does one improve the reliability of a survey instrument?

The general convention in research has been prescribed by Nunnally and Bernstein (1994) who state that one should strive for reliability values of .70 or higher. Reliability values increase as test length increases (Gulliksen, 1950) That is, the more items you have in your scale to measure the construct of interest the more reliable your scale will become. Many behavioural measures involve significant judgment on the part of an observer or a rater. Inter-rater reliability is the extent to which different observers are consistent in their judgments. For example, if you were interested in measuring university students' social skills, you could make video recordings of them as they interacted with another student whom they are meeting for the first time. Then you could have two or more observers watch the videos and rate each student's level of social skills. To the extent that each participant does in fact have some level of social skills that can be detected by an attentive observer, different observers' ratings should be highly correlated with each other. Inter-rater reliability is often assessed using Cronbach's á when the judgments are quantitative or an analogous statistic called Cohen's ê (the Greek letter kappa) when they are categorical.

11.4.3 Indices for Assessing the Reliability

Various indices for assessing the reliability of measures have been proposed. We shall look at them in relations to the type of reliability they measure. Various reliability indices are:

Parallel Forms Procedure: This procedure measures Equivalence. Here, one administers alternative forms of the same measure to either the same group or different group of respondents. This administration of the various forms occurs at the same time or following some time delay. The higher the degree of correlation between the two forms, the more equivalent they are. In practice the parallel forms procedure is seldom implemented, as it is difficult, if not impossible, to verify that two tests are indeed parallel (i.e., have equal means, variances, and correlations with other measures). Indeed, it is difficult enough to have one well-developed instrument to measure the construct of interest let alone two. Another situation in which equivalence will be important is when the measurement process entails subjective judgments or ratings being made by more than one person.

Test-Retest Procedure: This is used to measure stability. It involves administering the same measurement instrument to the same individuals under the same conditions after some period of time. Test-rest reliability is estimated with correlations between the scores at Time 1 and those at Time 2 (to Time x). Two assumptions underlie the use of the test-retest procedure. The first required assumption is that the characteristic that is measured does not change over the time period. The second assumption is that the time period is long enough that the respondents' memory of taking the test at Time 1 does not influence their scores at the second and subsequent test administrations.

Split-half reliability index: This index is used in estimating internal consistency. The split-half estimate entails dividing up the test into two parts (e.g., odd/even items or first half of the items/second half of the items), administering the two forms to the same group of individuals and correlating the responses.

index, the Kuder-Richardson formula 20 (KR-20) and the (Kuder & Richardson, 1937) index. These indices represent the average of all possible split-half estimates. The difference between the two is when they would be used to assess reliability. Specifically, coefficient alpha is typically used during scale development with items that have several response options (i.e., 1 = strongly disagree to 5 = strongly agree) whereas KR-20 is used to estimate reliability for dichotomous (i.e., Yes/No; True/False) response scales.

11.5 VALIDITYOFDATA

Validity is the extent to which the scores from a measure represent the variable they are intended to. But how do researchers make this judgment? We have already considered one factor that they take into account-reliability. When a measure has good test-retest reliability and internal consistency, researchers should be more confident that the scores represent what they are supposed to. There has to be more to it, however, because a measure can be extremely reliable but have no validity whatsoever. As an absurd example, imagine someone who believes that people's index finger length reflects their self-esteem and therefore tries to measure self-esteem by holding a ruler up to people's index fingers. Although this measure would have extremely good test-retest reliability, it would have absolutely no validity. The fact that one person's index finger is a centimetre longer than another's would indicate nothing about which one had higher self-esteem.

11.5.1 Concept of Validity

Validity examines how truthful the research results are. It is the extent to which the instrument measures what it purports to measure. In other words, does the research instrument allow you to hit "the bull's eye" of your research object? For example, a test that is used to screen applicants for M. Sc admissions into a UK university is valid if its scores are directly related to future academic performance of students either in research thesis or course work. Researchers generally determine validity by asking a series of questions, and will often look for the answers in the research of others.

There are various types of validity. They include, content validity, face validity, criterion-related validity (or predictive validity), construct validity, and discriminant validity. These will be explained here as under.

11.5.2 Types of Validity

Content Validity

This pertains to the degree to which the instrument fully assesses or measures the construct of interest. For example, say we are interested in evaluating employees' attitudes toward a training program within an organization. We would want to ensure that our questions fully represent the domain of attitudes toward the training program. The development of a content valid instrument is typically achieved by a rational analysis of the instrument by raters (ideally 3 to 5) familiar with the construct of interest. Content validity is the extent to which a measure "covers" the construct of interest. For example, if a researcher conceptually defines test anxiety as involving both sympathetic nervous system activation (leading to nervous feelings) and negative thoughts, then his measure of test anxiety should include items about both nervous feelings and negative thoughts. Or consider that attitudes are usually defined as involving thoughts, feelings, and actions toward something. By this conceptual definition, a person has a positive attitude toward exercise to the extent that he or she thinks positive thoughts about exercising, feels good about exercising, and actually exercises. So to have good content validity, a measure of people's attitudes toward exercise would have to reflect all three of these aspects. Like face validity, content validity is not usually assessed quantitatively. Instead, it is assessed by carefully checking the measurement method against the conceptual definition of the construct.

Face Validity

This component of content validity and is established when an individual reviewing the instrument concludes that it measures the characteristic or trait of interest. For instance, if a quiz in this class comprised items that asked questions pertaining to research methods you would most likely conclude that it was face valid. In short, it looks as if it is indeed measuring what it is designed to measure. Face validity is the extent to which a measurement method appears "on its face" to measure the construct of interest. Most people would expect a self-esteem questionnaire to include items about whether they see themselves as a person of worth and whether they think they have good qualities. So a questionnaire that included these kinds of items would have good face validity. The finger-length method of measuring self-esteem, on the other hand, seems to have nothing to do with self-esteem and therefore has poor face validity. Although face validity can be assessed quantitatively-for example, by having a large sample of people rate a measure in terms of whether it appears to measure what it is intended to-it is usually assessed informally.

Face validity is at best a very weak kind of evidence that a measurement method is measuring what it is supposed to. One reason is that it is based on people's intuitions about human behaviour, which are frequently wrong. It is also the case that many established measures in psychology work quite well despite lacking face validity. The Minnesota Multi-phasic Personality Inventory-2 (MMPI-2) measures many personality characteristics and disorders by having people decide whether each of over 567 different statements applies to them-where many of the statements do not have any obvious relationship to the construct that they measure. For example, the items "I enjoy detective or mystery stories" and "The sight of blood doesn't frighten me or make me sick" both measure the suppression of aggression. In this case, it is not the participants' literal answers to these questions that are of interest, but rather whether the pattern of the participants' responses to a series of questions matches those of individuals who tend to suppress their aggression.

Criterion-related

Criterion related validity is assessed when one is interested in determining the relationship of scores on a test to a specific criterion. An example is that scores on a job test for fresh graduate should be related to relevant criteria such as youth service corps completion (for students in Nigeria), class of degree, etc. Conversely, an instrument that measures colour inclinations would most assuredly demonstrate very poor criterion-related validity with respect to graduate job placement. Criterion validity is the extent to which people's scores on a measure are correlated with other variables (known as criteria) that one would expect them to be correlated with. For example, people's scores on a new measure of test anxiety should be negatively correlated with their performance on an important school exam. If it were found that people's scores were in fact negatively correlated with their exam performance, then this would be a piece of evidence that these scores really represent people's test anxiety. But if it were found that people scored equally well on the exam regardless of their test anxiety scores, then this would cast doubt on the validity of the measure.

A criterion can be any variable that one has reason to think should be correlated with the construct being measured, and there will usually be many of them. For example, one would expect test anxiety scores to be negatively correlated with exam performance and course grades and positively correlated with general anxiety and with blood pressure during an exam. Or imagine that a researcher develops a new measure of physical risk taking. People's scores on this measure should be correlated with their participation in "extreme" activities such as snowboarding and rock climbing, the number of speeding tickets they have received, and even the number of broken bones they have had over the years. When the criterion is measured at the same time as the construct, criterion validity is referred to as concurrent validity; however, when the criterion is measured at some point in the future (after the construct has been measured), it is referred to as predictive validity (because scores on the measure have "predicted" a future outcome).

Criteria can also include other measures of the same construct. For example, one would expect new measures of test anxiety or physical risk taking to be positively correlated with existing measures of the same constructs. This is known as convergent validity.

Assessing convergent validity requires collecting data using the measure. Researchers John Cacioppo and Richard Petty did this when they created their self-report Need for Cognition Scale to measure how much people value and engage in thinking. In a series of studies, they showed that people's scores were positively correlated with their scores on a standardised academic achievement test, and that their scores were negatively correlated with their scores on a measure of dogmatism (which represents a tendency toward obedience).

Construct Validity

Wainer and Braun (1998) describe the validity in quantitative research as "construct validity". The construct is the initial concept, notion, question or hypothesis that determines which data is to be gathered and how it is to be gathered. Construct validity is the degree to which an instrument measures the trait or theoretical construct that it is intended to measure. Also, indicates how well the scale measures the construct it was designed to measure. For example, if one were to develop an instrument to measure intelligence that does indeed measure IQ, then this test is construct valid.

Discriminant validity

Discriminant validity, on the other hand, is the extent to which scores on a measure are not correlated with measures of variables that are conceptually distinct. For example, self-esteem is a general attitude toward the self that is fairly stable over time. It is not the same as mood, which is how good or bad one happens to be feeling right now. So people's scores on a new measure of self-esteem should not be very highly correlated with their moods. If the new measure of self-esteem were highly correlated with a measure of mood, it could be argued that the new measure is not really measuring self-esteem; it is measuring mood instead.

For example, they found only a weak correlation between people's need for cognition and a measure of their cognitive style-the extent to which they tend to think analytically by breaking ideas into smaller parts or holistically in terms of "the big picture." Let's say you were researching depression in college students. In order to measure depression (the construct), you use two measurements: a survey and participant observation. If the scores from your two measurements are close enough (i.e. they converge), this demonstrates that they are measuring the same construct. If they don't converge, this could indicate they are measuring different constructs (for example, anger and depression or self-worth and depression.

11.6 SUMMARY

In this lesson we have discussed about processing of data including feeding and coding of data as well as about the reliability and validity of data. Irrespective of the method of data collection, qualitative or quantitative, the information is called 'raw data' or simply 'data'. The processing of data includes all operations undertaken from when a set of data is collected until it is ready to be analysed either manually or by a computer. Data processing in quantitative studies starts with data editing, which is basically 'cleaning' your data. This is followed by the coding of data, which entails developing a code book, pre-testing it, coding per se and verifying the coded data.

Validity is the most critical criterion and indicates the degree to which an instrument measures what it is supposed to measure. Validity can also be thought of as utility. In other words, validity is the extent to which differences found with a measuring instrument reflect true differences among those being tested. But the question arises: how can one determine validity without direct confirming knowledge? The answer may be that we seek other relevant evidence that confirms the answers wehave found with our measuring tool. What is relevant, evidence often depends upon the nature of the research problem and the judgement of the researcher. But one can certainly consider three types of validity in this connection: (i) Content validity; (ii) Criterion-related validity and (iii) Construct validity.(i) Content validity is the extent to which a measuring instrument provides adequate coverage of the topic under study. If the instrument contains a representative sample of the universe, the content validity is good. Its determination is primarily judgemental and intuitive. It can also be determined by using a panel of persons who shall judge how well the measuring instrument meets the standards, but there is no numerical way to express it.(ii) Criterion-related validity relates to our ability to predict some outcome or estimate the existence of some current condition. This form of validity reflects the success of measures used for some empirical estimating purpose. (iii) Construct validity is the most complex and abstract. A measure is said to possess construct validity to the degree that it confirms to predicted correlations with other theoretical propositions. Construct validity is the degree to which scores on a test can be accounted for by the explanatory constructs of a sound theory. For determining construct validity, we associate a set of other propositions. Reliability refers to the consistency of a measure. Psychologists consider three types of consistency: over time (test-retest reliability), across items (internal consistency), and across different researchers (inter-rater reliability). These two principles are discussed in this lesson to understand the concept clearly. When choosing participants, the characteristics of the candidates were a factor and is usually the same for each assessment administered, as with this particular one. Cut-off scores could be used if needed, however, are not inherently necessary for this assessment measure. Therefore, it is important to note that one's ability to answer a research question is only as good as the instruments developed or ones method of data collection. A well-developed survey instrument will better provide a researcher with quality data with which to answer a question or solve a problem. Finally, it must be remember that for something to be valid it must be reliable but it must also measure what it is intended to measure.

11.7 GLOSSARY

Data Displaying: Also known as data visualisation, a data display is a visual representation of raw or processed data that aims to communicate a small number of insights about the behaviour of an underlying table, which is otherwise difficult or impossible to understand to the naked eye.

Data Processing: Data processing is, generally, "the collection and manipulation of items of data to produce meaningful information."

Closed Ended Questions: Closed-ended questions are questions that can only be answered by selecting from a limited number of options, usually multiple-choice questions with a single-word answer , 'yes' or 'no', or a rating scale (e.g. from strongly agree to strongly disagree).

Coding of Data: Coding of data refers to the process of transforming collected information or observations to a set of meaningful, cohesive categories. It is a process of summarizing and re-presenting data in order to provide a systematic account of the recorded or observed phenomenon.

Editing of Data: Data editing is the process of "improving" collected survey data.

Open-ended Questions: Open-ended questions are phrased as a statement which requires a longer response. The response can be compared to information that is already known to the questioner.

Pre-test: Pretesting is the stage in survey research when survey questions and questionnaires are tested on members of target population/study population, to evaluate the reliability and validity of the survey instruments prior to their final distribution.

Reliability: The term reliability in psychological research refers to the consistency of a research study or measuring test.

Validity: Validity is the extent to which a concept, conclusion or measurement is well-founded and likely corresponds accurately to the real world.

Test Re-test: The test-retest reliability method is one of the simplest ways of testing the stability and reliability of an instrument over time.

Face Validity: Face validity refers to the extent to which a test appears to measure what it is intended to measure. Face validity is a test of internal validity.

Content Validity: content validity refers to the extent to which a measure represents all facets of a given construct.

11.8 SELFASSESSMENT QUESTIONS

Tick ( ) the correct option:-

Alternate form reliability is also known as:

a) Convergent reliability b) Test-retest reliability

c) Split half reliability d) Parallel forms

Why might test-retest reliability become less important?

If convergent validity was not high

If participants become familiar with a test and all perform better on the second occasion.

If the concept being measured is not expected to be stable over time

If alternate-form reliability cannot be calculated.

If care has been taken to ensure that a measure is measuring all aspects of something fully, it has high.

a) Face validity b) Criterion validity

c) Content validity d) Internal validity

If a measure correlates highly with other established measures of the same thing, it is said to have high.

a) Criterion validity b) Face validity

c) Content validity d) Construct validity

If a measure is based on a full and close examination of the underlying concept, along with the related theoretical approaches, then it is said to have high

a) Convergent validity b) Content validity

c) Face validity d) Construct validity

11.9 LESSON END EXERCISE

Discuss what is meant by measurement validity. Provide an example of measurement validity via a social work example.

____

____

____

Discuss how measurement validity is different from, and similar to, the quantitative and qualitative research approaches.

____

____

____

Discuss why measurement is fundamental to social work research.

____

____

____

Discuss the relationship between measurement validity and measurement reliability as presented in this lesson.

____

____

____

Discuss how a measuring instrument is assessed for its content validity. How can you tell if an instrument is content valid?

____

____

____

Discuss how content validity is different from, and similar to, the quantitative and qualitative research approaches.

What is face validity? What is the difference between content validity and face validity? Provide a social work example throughout your discussion.

What is the difference between concurrent validity and predictive validity?

Describe, in brief, the importance of editing, coding, classification, tabulation and presentation of data in the context of research study.

Discuss the fundamental rules of code construction.

Discuss with the help of suitable examples various steps involved in data processing.

11.10 SUGGESTED READING

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

DATAANALYSIS & INTERPRETATION

UNIT-III **Lesson No. 12**

OVERVIEW OFUNIVARIATE, BIVARIATEAND MULTIVARIATE DATA ANALYSIS THROUGH SPSS

STRUCTURE

12.1 Introduction

12.2 Objectives

12.3 Univariate Data Analysis

12.3.1 Concept

12.3.2 Analysing Data Using Univariate Analysis

12.3.2 Univariate Analysis Using SPSS

12.4 Bivariate Data Analysis

12.4.1 Bivariate Analysis Using SPSS

12.5 Multivariate Data Analysis

12.5.1 Multivariate Analysis Using SPSS

12.6 Summary

12.7 Glossary

12.8 Self Assessment Questions

12.9 Lesson End Exercise

12.10 Suggested Readings

12.1 INTRODUCTION

In data analysis, the data gathered are statistically analyzed to see if the hypotheses that were generated have been supported. For e.g. if we want to see that unresponsiveness of employees affects customer switching, we might want to do a correlational analysis to determine the relationship between these variables. Also, in the field of data, there is nothing more important than understanding the data that we are trying to analyse and for understanding the data it is important to understand the purpose of the analysis because this will help us save time and dictate how to go about analysing the data. Moreover, when it comes to the level of analysis in statistics, there are three different analysis techniques that exist. These are –

Univariate analysis

Bivariate analysis

Multivariate analysis

The selection of the data analysis technique is dependent on the number of variables, types of data and focus of the statistical inquiry. Also, there are different tools, techniques and methods that can be used to conduct data analysis. We could use software libraries, visualization tools and statistic testing methods. In this lesson we will be focusing on Univariate, Bivariate and Multivariate data analysis through SPSS.

12.2 OBJECTIVES

After studying this lesson, you will be able to:

Understand the concept of Univariate data analysis.

Know about different tools and techniques used for Univariate data analysis.

Understand the concept of bivariate data analysis.

Know about different tools and techniques used for bivariate data analysis.

Understand the concept of Multivariate data analysis.

Know about different tools and techniques used for Multivariate data analysis.

12.3 UNIVARIATE DATAANALYSIS

Before discussing about the univariate data analysis first we must understand the types of variables:

Categorical Variables: variables that have a finite number of categories or distinct groups. Examples: gender, method of payment, horoscope, etc.

Numerical Variables: variables that consist of numbers. There are two main numerical variables.

Discrete Variables: variables that can be counted within a finite time. Examples: the change in your pocket, number of students in a class, numerical grades, etc.

Continuous Variables: variables that are infinite in number often measured on a scale of sort. Examples: weight, height, temperature, date and time of a payment, etc.

However, depending on the type of variable, it can also be changed to another variable for ease of use. For example the date and time could be broken down to year, month and time could be categorised into AM and PM sales. A common technique for continuous variables is “binning” the variables into categories. For example, the weight of a person can be categorised into “below average”/”slim”, “average” and “above average”/”obese” by setting ranges.

12.3.1 Concept

Univariate data analysis is the simplest of the three analyses where the data you are analysing is only one variable. There are many different ways people use univariate analysis. The most common univariate analysis is checking the central tendency (mean, median and mode), the range, the maximum and minimum values, and standard deviation of a variable. Common visual technique used for univariate analysis is a histogram, which is a frequency distribution graph.

The analysis of univariate data is thus the simplest form of analysis since the information deals with only one quantity that changes. It does not deal with causes or relationships and the main purpose of this analysis is to describe the data and find patterns that exist within it. The example of a univariate data can be height.

Heights (in cm)	164	167.3	170	174.2	178	180	186

Table : 12.1

Suppose that the heights of seven students of a class are recorded (as above), there is only one variable that is height and it is not dealing with any cause or effect relationship. The description of patterns found in this type of data can be made by drawing conclusions using central tendency measures (such as mean, median and mode), dispersion or spread of data (range, minimum, maximum, quartiles, variance and standard deviation) and by using frequency distribution tables, histograms, pie-charts, frequency polygon and bar charts.

In another example, in a survey of a class room, the researcher may be looking to count the number of boys and girls. In this instance, the data would simply reflect the number, i.e. a single variable and its quantity. The key objective of Univariate analysis is to simply describe the data to find patterns within the data. This is be done by looking into the mean, median, mode, dispersion, variance, range, standard deviation etc.

Univariate analysis is conducted through several ways which are mostly descriptive in nature.

Frequency Distribution Tables
Histograms
Frequency Polygons
Pie Charts
Bar Charts

12.3.2 Analysing Data Using Univariate Analysis

Data analyst uses Univariate data analysis as it explores each variable in a data set separately. It looks at the range of values, as well as the central tendency of the values. It describes the pattern of response to the variable. It describes each variable on its own.

Descriptive statistics describe and summarise data. Univariate descriptive statistics describe individual.

Raw Data

Obtain a printout of the raw for all the variables. Raw data resembles a matrix, with the variable names heading the columns and the information for each case or record displayed across the rows.

Example: Raw data for a study of injuries among county workers (first 10 cases)

Injury . Report No	***County Name***	***Cause of Injury***	***Severity of*** Injury
1	County A	Fall	3
2	County B	Auto	4
3	County C	Fall	6
4	County C	Fall	4
5	County B	Fall	5
6	County A	Violence	9
7	County A	Auto	3
8	County A	Violence	2
9	County A	Violence	9
10	County B	Auto	3

Table : 12.2

It is difficult to tell what is going on with each variable in this data set. Raw data is difficult to grasp, especially with large number of cases or records. Univariate descriptive statistics can summarize large quantities of numerical data and reveal patterns in the raw data. In order to present the information in a more organised format, start with univariate descriptive statistics for each variable.

For example, the variable Severity of Injury:

Severity of Injury
3

4
6
4
5
9
3
2
9
3

Table : 12.3

Frequency Distribution

Obtain a frequency distribution of the data for the variable. This is done by identifying the lowest and highest values of the variable, and then putting all the values of the variable in order from lowest to highest. Next, count the number of appearance of each value of the variable. This is a count of the frequency with which each value occurs in the data set.

For example, for the variable "Severity of Injury," the values range from 2 to 9.

Severity of Injury	***Number of Injuries with this Severity***
2	1
3	3
4	2
5	1
6	1
9	2
Total	10

Table : 12.4

Grouped Data

Decide on whether the data should be grouped into classes. The severity of injury ratings can be collapsed into just a few categories or groups. Grouped data usually has from 3 to 7 groups. There should be no groups with a frequency of zero (for example, there are no injuries with a severity rating of 7 or 8). One way to construct groups is to have equal class intervals (e.g., 1-3, 4-6, 7-9).

Another way to construct groups is to have about equal numbers of observations in each group. Remember that class intervals must be both mutually exclusive and exhaustive.

Severity of Injury	***Number of Injuries with this Severity***
Mild (1-3)	4
Moderate (4-6)	4
Severe (6-9)	2
Total	10

Table : 12.5

Cumulative Distributions

Cumulative frequency distributions include a third column in the table (this can be done with either simple frequency distributions or with grouped data):

Severity of Injury	***Number of Injuries***	***Cumulative frequency***
2	1	1
3	3	4
4	2	6
5	1	7
6	1	8
9	2	10

Table : 12.6

A cumulative frequency distribution can answer questions such as, how many of the injuries were at level 5 or lower? Answer here is = 7

Percentage Distributions

Frequencies can also be presented in the form of percentage distributions and cumulative percentages.

Severity of Injury	*Percent of Injuries*	*Cumulative percentages*
2	10	10
3	30	40
4	20	50
5	10	70
6	10	80
9	20	100

Table : 12.7

Graphing is a way of visually presenting the data. Many people can grasp the information presented in a graph better than in a text format. The purpose of graphing is to:

present the data
enhance textual descriptions
make comparisons easy
provoke thought about the data
summarize the data
describe and explore the data
avoid distortion

Bar Graphs

Bar graphs are used to display the frequency distributions for variables measured at the nominal and ordinal levels. Bar graphs use the same width for all the bars on the graph, and there is space between the bars. Label the parts of the graph, including the title, the left (Y) or vertical axis, the right (X) or horizontal axis, and the bar labels.

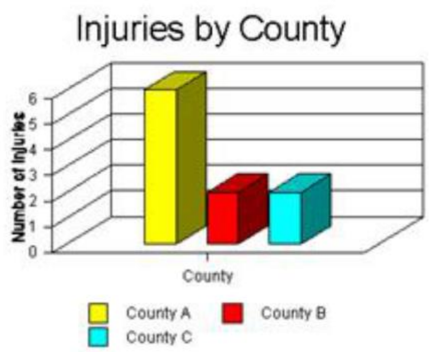


Table : 12.8

Bar graphs can also be rotated so that the bars are parallel to the horizontal orientation of the page.
For example:

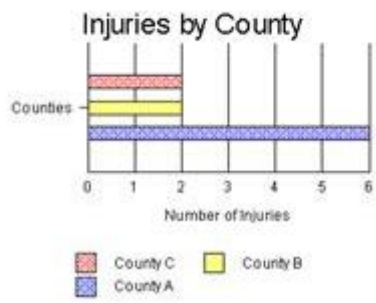


Table : 12.9

Histogram

A histogram is a chart that is similar to a bar chart, but it is used for interval and ratio level variables. With a histogram, the width of the bar is important, since it is the total area under the bar that represents the proportion of the phenomenon accounted for by each category.

The bars convey the relationship of one group or class of the variable to the other(s).

For example, in the case of the counties and employee injuries, we might have information on the rate of injury according to the number of workers in each county in State X.

County Name	Rate of Injury per 1,000 workers
County A	5.5
County B	4.2
County C	3.8
County D	3.6
County E	3.4
County F	3.1
County G	1.8
County H	1.7
County I	1.6
County J	1.0
County K	0.9
County L	0.4

Table : 12.10

If we group the injury rates into three groups, then a low rate of injury would be 0.0-1.9 injuries per 1,000 workers; moderate would be 2.0-3.9; and high would be 4.0 and above (in this case, up to 5.9). This could be graphed as follows:

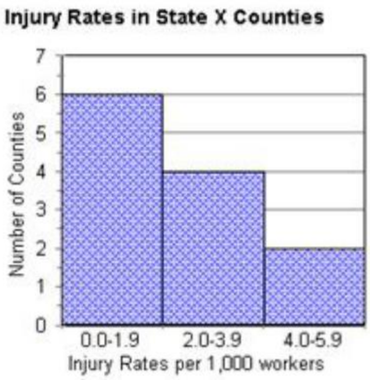


Table :12.11

Frequency Polygon

A frequency polygon is another way of displaying information for an interval or ratio level variable. A frequency polygon displays the area under the curve that is represented by the values of the variable. This type of chart is also used to show time series graphs or the changes in rates over time. For example, the following table shows the average injury rate per 1,000 employees for counties in State X for the years 1980 to 1990.

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Rate	3.6	4.2	3.4	5.5	3.8	3.1	1.7	1.8	1.0	1.6	0.9

Table : 12.12

A cumulative frequency polygon is used to display the cumulative distribution of values for a variable.

Table : 12.13

Pie Chart

Another way to show the relationships between classes or categories of a variable is in a pie or circle chart. In a pie chart, each "slice" represents the proportion of the total phenomenon that is due to each of the classes or groups.

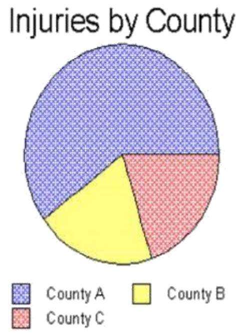


Table : 12.14

Rates and Ratios

Other ways to look at the sub-groups or classes within one variable is by the relation of each sub-group or class to the whole.

This can be calculated with a proportion. A proportion is obtained by dividing the frequency of observations counted for one group or class (written as f) by the total number of observations counted for the variable (written as N).

This can be expressed as f / N

A percentage is the same as a proportion, multiplied by 100.

This can be expressed as f / N x 100

A rate is the relationship between two different numbers, for example, the number of injuries among county workers and the population of the county. This can be calculated as the first number (N_1, or injuries) divided by the second number (N_2, or population).

This can be expressed as N1 / N2

Many health statistics are expressed as rates, for example, the birth rate is the number of births per some population, such as number of births per 1,000 women.

12.3.3 Univariate Data Analysis through SPSS

Univariate analysis, looking at single variables, is typically the first procedure one does when examining first time data. There are a number of reasons why it is the first procedure, and most of the reasons we will cover at the end of this chapter, but for now let us just say we are interested in the "basic" results. If we are examining a survey, we are interested in how many people said, "Yes" or "No", or how many people "Agreed" or "Disagreed" with a statement. We aren't really testing a traditional hypothesis with an independent and dependent variable; we are just looking at the distribution of responses.

The SPSS tools for looking at single variables include the following procedures: Frequencies, Descriptive and Explore all located under the Analyse menu.

SPSS is a simple analytics tool which is very easy to maintain and use.

To begin the process start SPSS, then open the data file. Under the Analyze menu, choose Descriptive Statistics and the procedure desired: Frequencies, Descriptives, Explore, Crosstabs.

Frequencies

Generally a frequency is used for looking at detailed information in a nominal (category) data set that describes the results. Categorical data is for variables such as gender i.e. males are coded as "1" and females are coded as "2." Frequencies options include a table showing counts and percentages, statistics including percentile values, central tendency, dispersion and distribution, and charts including bar charts and histograms. The steps for using the frequencies procedure is to click the Analyze menu choose Descriptive Statistics then from the submenu choose Frequencies and select your variables for analysis. You can then choose statistics options, choose chart options, choose format options, and have SPSS calculate your request.

Choosing Frequencies Procedure: From the Analyze menu, highlight Descriptive Statistics, see figure 12.1 below, then move to the sub menu and click on Frequencies. A dialog box, see figure 12.2 below, will appear providing a scrollable list of the variables on the left, a Variable(s) choice box, and buttons for Statistics, Charts and Format options.

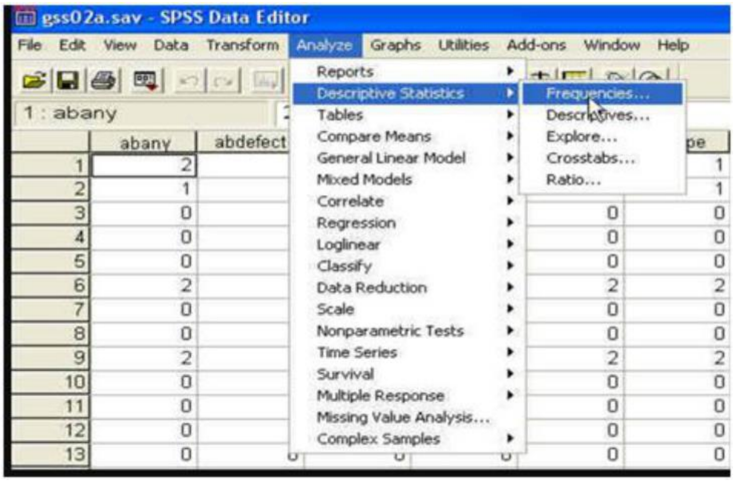


Figure : 12.1

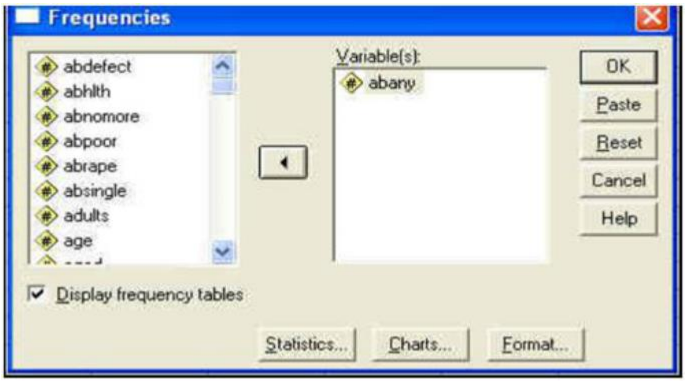


Figure 12.2

Descriptive

Descriptive (Analysis, Descriptive Statistics, Descriptive, Figure 12.3) is used to obtain summary information about the distribution, variability, and central tendency of continuous variables. Possibilities for Descriptive include mean, sum, standard deviation, variance, range, minimum, maximum, S.E. mean, kurtosis and skewness. For this example we are going to look at the distribution of age and education for the General Social Survey sample.

Since both these variables were measured at interval/ratio level, different statistics from our previous example will be used.

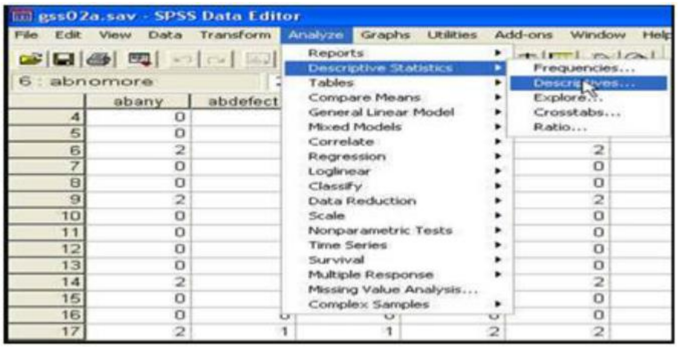


Figure 12.3

First click the Analyze menu and select Descriptive Statistics, then move across to the sub menu and select Descriptive (see Figure 12.3).

Explore

Explore is primarily used to visually examine the central tendency and distributional characteristics of continuous variables. Explore statistics include M estimators, outliers, and percentiles. Grouped frequency tables and displays, as well as Stem and leaf and box plots, are available. Explore will aid in checking assumptions with Normality plots and Spread vs. Level with the Levene test.

From the Analyze menu choose Descriptive Statistics, drag to the sub menu and select Explore.

As in the other procedures, find and click the variable you want to explore, and then click the select arrow to include your variable in the Dependent List box. Choose the variable and the dialog box should look like Figure 12.4.

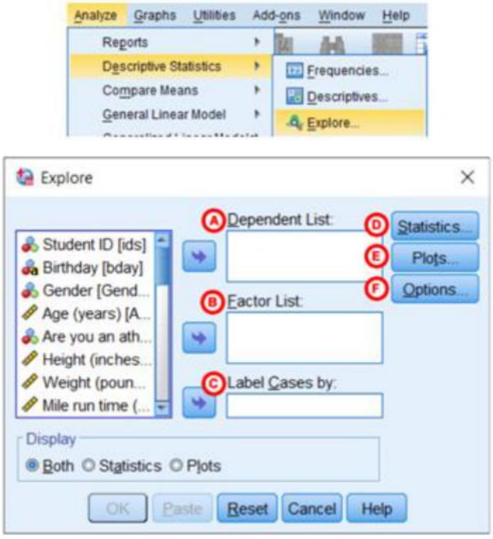


Figure 12.4

The Descriptive statistics output should look like Figure 12.5 given below:

Descriptive Statistics

	N	Range	Minimu	Maximu	Mean	Std.	Varianc	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Number of Older Siblings	46	4	0	4	1.26	1.255	1.575	.678	.350	-.543	.688
Valid N (listwise)	46										

Figure 12.5

12.4 BIVARIATE DATAANALYSIS

Bivariate analysis is where we are comparing two variables to study their relationships. These variables could be dependent or independent to each other. In Bivariate analysis there is always a Y-value for each X-value. The most common visual technique for bivariate analysis is a scatter plot, where one variable is on the x-axis and the other on the y-axis. In addition to the scatter plot, regression plot and correlation coefficient are also frequently used to study the relationship of the variables.

Further, the analysis of this type of data deals with causes and effect relationships and the analysis is done to find out the relationship among the two variables only. Example of bivariate data can be temperature and ice cream sales in summer season.

Temperature (in Celsius)	***Ice Cream Sales***
20	2000
25	2500
35	5000
43	7800

Suppose the temperature and ice cream sales are the two variables of a bivariate data (as above). Here, the relationship is visible from the table that temperature and sales are directly proportional to each other and thus related because as the temperature increases, the sales also increase. Thus, bivariate data analysis involves comparisons, relationships, causes and explanations. These variables are often plotted on X and Y axis on the graph for better understanding of data and one of these variables is independent while the other is dependent.

In another example, in a survey of a classroom, the researcher may be looking to analysis the ratio of students who scored above 85% corresponding to their genders. In this case, there are two variables: gender = X (independent variable) and result = Y (dependent variable). A Bivariate analysis is will measure the correlations between the two variables. Bivariate analysis is generally conducted by using correlation and regression.

12.4.1 Bivariate Data Analysis Using SPSS

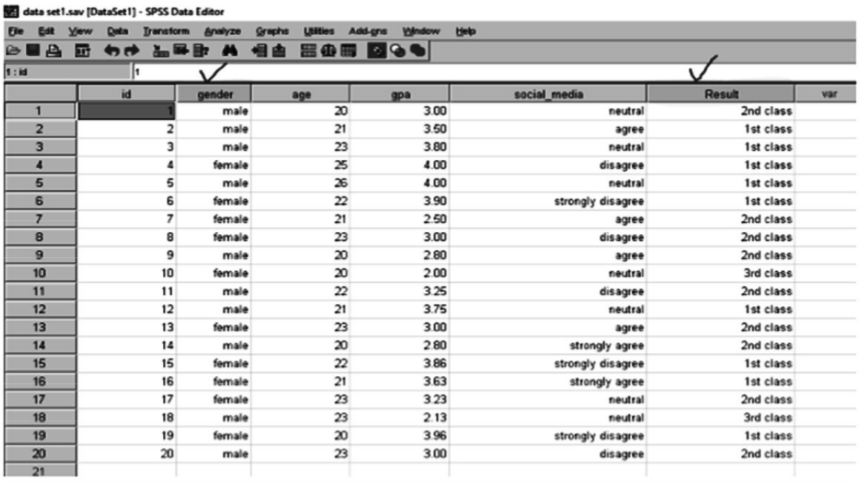
Bivariate analysis using SPSS is very simple procedure of finding the association between two variables. SPSS is a very simple data analytics tool by which we can perform our analysis very easily.

Bivariate analysis is the analysis of two random variable and find their association. For bivariate analysis we mainly use crosstabs and to show the association we use chi-square test. In chi-square table, we interpret the p-values. P-values interpretation is following-

A small p-value (typically dH 0.05) indicates strong evidence against the null hypothesis, so we reject the null hypothesis.

A large p-value (> 0.05) indicates weak evidence against the null hypothesis, so we fail to reject the null hypothesis.

p- values very close to the cut off (0.05) are considered to be marginal. For bivariate analysis in SPSS we use the following data set and conduct bivariate analysis for the two highlighted variable (gender and result) and we also found their association.



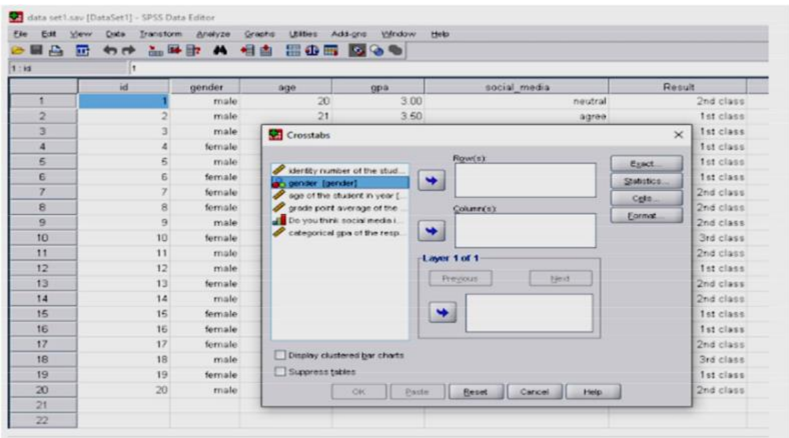
	id	gender	age	gpa	social_media	Result	var
1	1	male	20	3.00	neutral	2nd class	
2	2	male	21	3.50	agree	1st class	
3	3	male	23	3.80	neutral	1st class	
4	4	female	25	4.00	disagree	1st class	
5	5	male	26	4.00	neutral	1st class	
6	6	female	22	3.90	strongly disagree	1st class	
7	7	female	21	2.50	agree	2nd class	
8	8	female	23	3.00	disagree	2nd class	
9	9	male	20	2.80	agree	2nd class	
10	10	female	20	2.00	neutral	3rd class	
11	11	male	22	3.25	disagree	2nd class	
12	12	male	21	3.75	neutral	1st class	
13	13	female	23	3.00	agree	2nd class	
14	14	male	20	2.80	strongly agree	2nd class	
15	15	female	22	3.86	strongly disagree	1st class	
16	16	female	21	3.63	strongly agree	1st class	
17	17	female	23	3.23	neutral	2nd class	
18	18	male	23	2.13	neutral	3rd class	
19	19	female	20	3.96	strongly disagree	1st class	
20	20	male	23	3.00	disagree	2nd class	
21							

Data set

Figure : 12.6

First we have to go to–

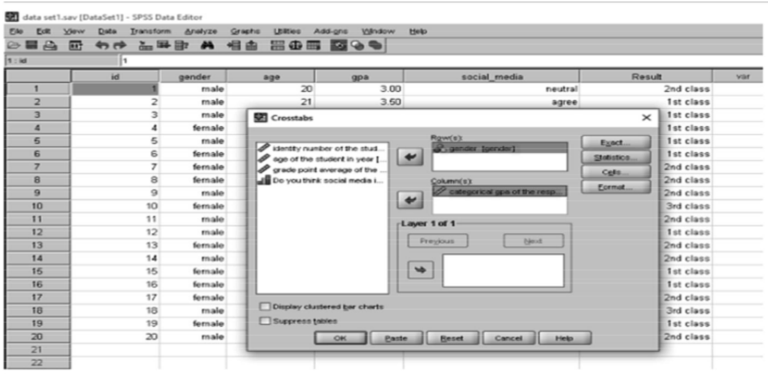
Analyze Descriptive statistics Crosstabs



Crosstabs

Figure : 12.7

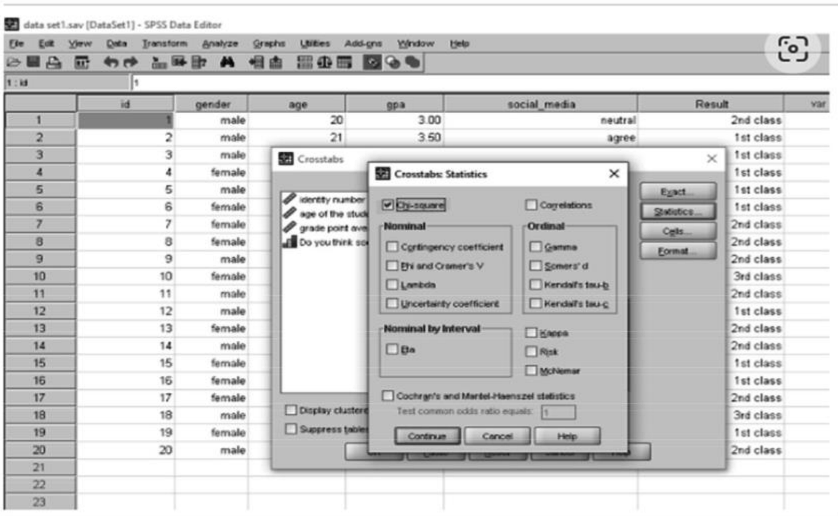
Now we can see two box named Row and Column and we have to enter one variable into Row box and another into column box and click on statistics.



Row and column box

Figure : 12.8

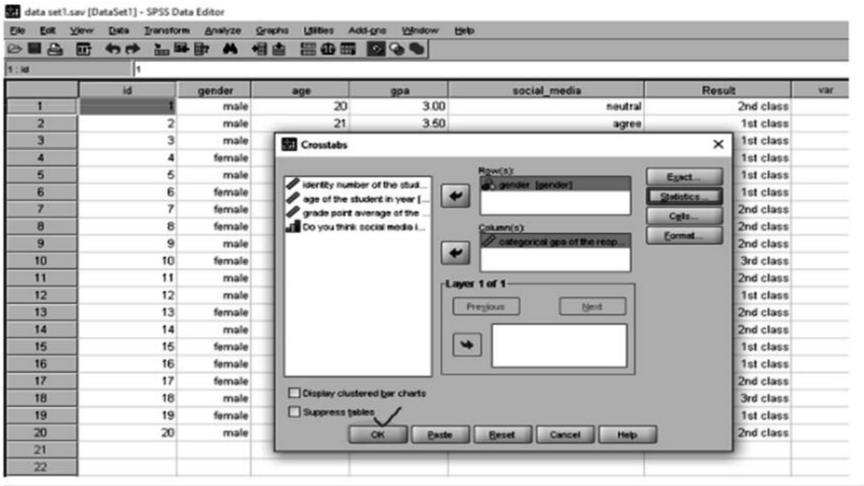
In statistics we have to select chi-square and click on continue.



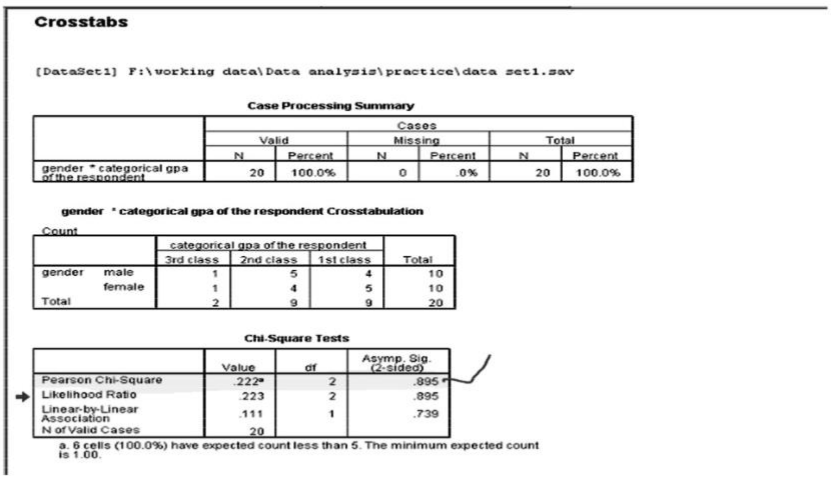
Chi-square

Figure : 12.9

4. Finally we have to click on OK and then we the see the output window.



click on OK



Crosstabs

[DataSet1] F:\working data\Data analysis\practice\data set1.sav

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
gender * categorical gpa of the respondent	20	100.0%	0	.0%	20	100.0%

gender * categorical gpa of the respondent Crosstabulation

Count

		categorical gpa of the respondent			Total
		3rd class	2nd class	1st class	
gender	male	1	5	4	10
	female	1	4	5	10
Total		2	9	9	20

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	.222[a]	2	.895
Likelihood Ratio	.223	2	.895
Linear-by-Linear Association	.111	1	.739
N of Valid Cases	20		

a. 6 cells (100.0%) have expected count less than 5. The minimum expected count is 1.00.

Figure : 12.10

OUTPUTOFBIVARIATE DATAANALYSIS Interpretation:

To interpret the chi square value we have build a hypothesis for association test. Let, H_0: There is no association between gender and result, vs

H_1: H0 is not true.

Here we have a p-value 0.895 which is greater than 0.05. we can not reject our null hypothesis.

That means our null hypothesis is accepted and there is no association between gender and result.

A small p-value (typically dH 0.05) indicates strong evidence against the null hypothesis, so we reject the null hypothesis.

A largep-value (> 0.05) indicates weak evidence against the null hypothesis, so we fail to reject the null hypothesis.

p-values very close to the cutoff (0.05) are considered to be marginal.

12.5 MULTIVARIATE DATAANALYSIS

Multivariate analysis is similar to bivariate analysis but here in multivariate data analysis we are comparing more than two variables. For three variables, you can create a 3-D model to study the relationship (sometimes also known as trivariate analysis). However, since we cannot visualize anything above the third dimension, we often rely on other software and techniques for us to be able to grasp the relationship in the data

Therefore, when the data involves three or more variables, it is categorised under multivariate data analysis. Example of this type of data is suppose an advertiser wants to compare the popularity of four advertisements on a website, then their click rates could be measured for both men and women and relationships between variables can then be examined. The ways to perform analysis on this data depends on the goals to be achieved. Some of the techniques are regression analysis, path analysis, factor analysis and multivariate analysis of variance (MANOVA), cluster analysis, etc.

Here is another example of multivariate data analysis – A doctor has collected data on cholesterol, blood pressure, and weight. She also collected data on the eating habits of the subjects (e.g., how many ounces of red meat, fish, dairy products, and chocolate consumed per week). She wants to investigate the relationship between the three measures of health and eating habits? In this instance, a multivariate data analysis would be required to understand the relationship of each variable with each other.

12.5.1 Multivariate Data Analysis Using SPSS

Multiple Regression

Multiple regression is an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other variables. The variable we want to predict is called the dependent variable (or sometimes, the outcome, target or criterion variable). The variables we are using to predict the value of the dependent variable are called the independent variables (or sometimes, the predictor, explanatory or regressor variables).

For example, you could use multiple regression to understand whether exam performance can be predicted based on revision time, test anxiety, lecture attendance and gender. Alternately, you could use multiple regression to understand whether daily cigarette consumption can be predicted based on smoking duration, age when started smoking, smoker type, income and gender.

Multiple regression also allows you to determine the overall fit (variance explained) of the model and the relative contribution of each of the predictors to the total variance explained. For example, you might want to know how much of the variation in exam performance can be explained by revision time, test anxiety, lecture attendance and gender "as a whole", but also the "relative contribution" of each independent variable in explaining the variance.

This "quick start" guide shows you how to carry out multiple regression using SPSS Statistics, as well as interpret and report the results from this test. However, before we introduce you to this procedure, you need to understand the different assumptions that your data must meet in order for multiple regression to give you a valid result. We discuss these assumptions next.

ASSUMPTIONS

When you choose to analyse your data using multiple regression, part of the process involves checking to make sure that the data you want to analyse can actually be analysed using multiple regression. You need to do this because it is only appropriate to use multiple regression if your data "passes" eight assumptions that are required for multiple regression to give you a valid result. Let's take a look at these eight assumptions:

Assumption 1: Our dependent variable should be measured on a continuous scale (i.e., it is either an interval or ratio variable). Examples of variables that meet this criterion include revision time (measured in hours), intelligence (measured using IQ score), exam performance (measured from 0 to 100), weight (measured in kg), and so forth.

Assumption 2: The two or more independent variables, can be either continuous (i.e., an interval or ratio variable) or categorical (i.e., an ordinal or nominal variable).

Assumption 3: We should have independence of observations (i.e., independence of residuals), which we can easily check using the Durbin-Watson statistic, which is a simple test to run using SPSS Statistics. We explain how to interpret the result of the Durbin-Watson statistic, as well as showing you the SPSS Statistics procedure required, in our enhanced multiple regression guide.

Assumption 4: There needs to be a linear relationship between:

the dependent variable and each of your independent variables, and

the dependent variable and the independent variables collectively.

we suggest creating scatterplots and partial regression plots using SPSS Statistics, and then visually inspecting these scatterplots and partial regression plots to check for linearity. If the relationship displayed in your scatterplots and partial regression plots are not linear, you will have to either run a non-linear regression analysis or "transform" your data, which you can do using SPSS Statistics.

Assumption 5: Our data needs to show homoscedasticity, which is where the variances along the line of best fit remain similar as you move along the line. When we analyse our own data, we will need to plot the standardised residuals against the unstandardised predicted values.

Assumption 6: Data must not show multicollinearity, which occurs when you have two or more independent variables that are highly correlated with each other. This leads to problems with understanding which independent variable contributes to the variance explained in the dependent variable, as well as technical issues in calculating a multiple regression model.

Assumption 7: There should be no significant outliers, high leverage points or highly influential points. Outliers, leverage and influential points are different terms used to represent observations in our data set that are in some way unusual when you wish to perform a multiple regression analysis.

Assumption 8: Finally, we need to check that the residuals (errors) are approximately normally distributed. Two common methods to check this assumption include using: (a) a histogram (with a superimposed normal curve) and a Normal P-P Plot; or (b) a Normal Q-Q Plot of the standardised residuals.

Example

A health researcher wants to be able to predict "VO_2max", an indicator of fitness and health. Normally, to perform this procedure requires expensive laboratory equipment and necessitates that an individual exercise to their maximum (i.e., until they can longer continue exercising due to physical exhaustion). This can put off those individuals who are not very active/fit and those individuals who might be at higher risk of ill health (e.g., older unfit subjects). For these reasons, it has been desirable to find a way of predicting an individual's VO_2max based on attributes that can be measured more easily and cheaply. To this end, a researcher recruited 100 participants to perform a maximum VO_2max test, but also recorded their "age", "weight", "heart rate" and "gender". Heart rate is the average of the last 5 minutes of a 20 minute, much easier, lower workload cycling test. The researcher's goal is to be able to predict VO_2max based on these four attributes: age, weight, heart rate and gender.

SETUPIN SPSS STATISTICS

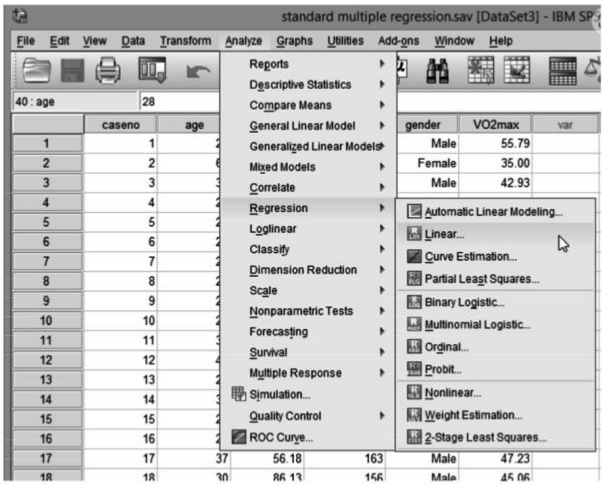
In SPSS Statistics, we created six variables: (1) VO_2max, which is the maximal aerobic capacity; (2) age, which is the participant's age; (3) weight, which is the participant's weight (technically, it is their 'mass'); (4) heart_rate, which is the participant's heart rate; (5) gender, which is the participant's gender; and (6) caseno, which is the case number. The caseno variable is used to make it easy for you to eliminate cases (e.g., "significant outliers", "high leverage points" and "highly influential points") that you have identified when checking for assumptions. In our enhanced multiple regression guide, we show you how to correctly enter data in SPSS Statistics to run a multiple regression when you are also checking for assumptions.

TEST PROCEDURE IN SPSS STATISTICS

The seven steps below show you how to analyse your data using multiple regression in SPSS Statistics when none of the eight assumptions in the previous section, Assumptions, have been violated.

Steps:

Click Analyze > Regression > Linear... on the main menu, as shown below:



You will be presented with the Linear Regression dialogue box below:

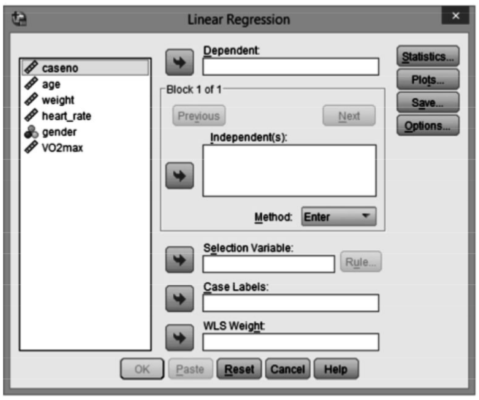


Figure : 12.11

Transfer the dependent variable, VO_2max, into the Dependent: box and the independent variables, age, weight, heart_rate and gender into the Independent(s): box, using the  buttons, as shown below (all other boxes can be ignored):

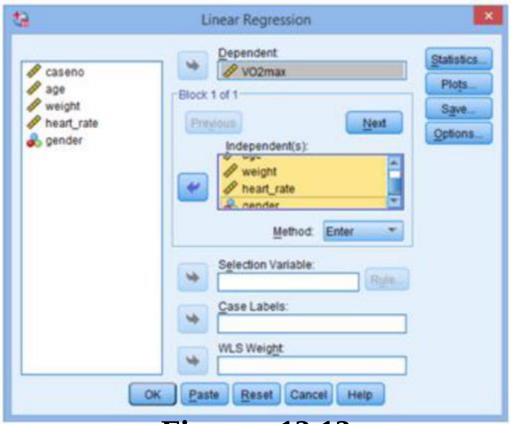
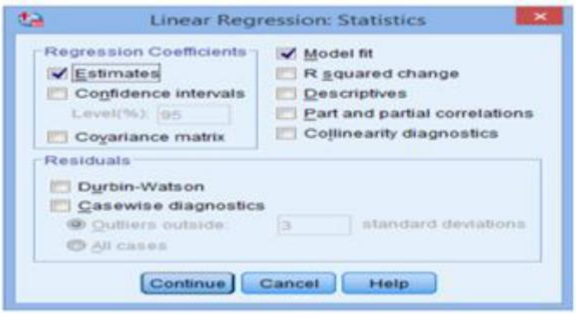


Figure : 12.12

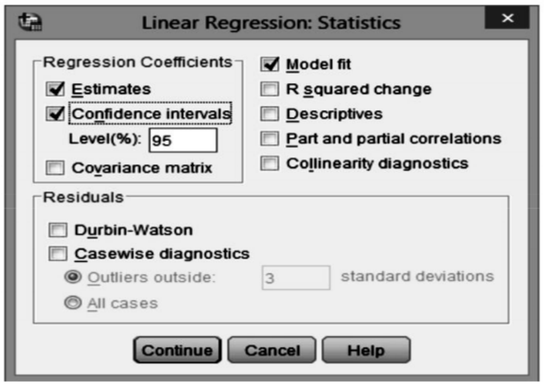
Click on the 

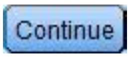
button.

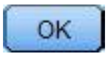
You will be presented with the Linear Regression: Statistics dialogue box, as shown below:



In addition to the options that are selected by default, select Confidence intervals in the –Regression Coefficients– area leaving the Level(%): option at “95”. You will end up with the following screen:



Click on the  button. You will be returned to the Linear Regression dialogue box.

Click on the  button. This will generate the output.

12.8 SUMMARY

In this lesson, we have discussed about the various types of analysis, viz., univariate data analysis, bivariate data analysis and multivariate data analysis. Univariate analysis looks at one variable, bivariate analysis looks at two variables and their relationship. Multivariate analysis looks at more than two variables and their relationship. Aside from the techniques mentioned above, there are numerous software's and packages that helps the data analysts and data scientists uncover the story of the data. For e.g. SPSS.

12.9 GLOSSARY

Univariate- Univariate is a term commonly used in statistics to describe a type of data which consists of observations on only a single characteristic or attribute.

Bivariate - Bivariate analysis is one of the simplest forms of quantitative analysis. It involves the analysis of two variables, for the purpose of determining the empirical relationship between them.

Multivariate - Multivariate analysis is concerned with the interrelationships among several variables. The data may be metrical, categorical, or a mixture of the two.

SPSS - SPSS means “Statistical Package for the Social Sciences” and was first launched in 1968.

Multiple Regression- Multiple regression is a statistical technique that can be used to analyze the relationship between a single dependent variable and several independent variables.

Chi-Square- Chi-square is a statistical test used to examine the differences be-tween categorical variables from a random sample in order to judge goodness of fit between expected and observed results.

Cross Tab: A crosstab is a table showing the relationship between two or more variables.

12.10 SELFASSESSMENT QUESTIONS

Tick the correct answer:-

Which of the following cannot be covered under univariate analysis of data?

Association between two variables
Computation of mean, median and mode
Preparation of frequency table
Computation of percentage frequency for a variable

For which type of measurement, the coefficient of variation can be computed.

a. Nominal scale b. Ordinal scale
c. Interval scale d. Ratio scale

There are 20% female students in a class-this is an example of:

a. Descriptive analysis b. Inferential analysis
c. Gender bias d. All of the above are correct

The uses of a frequency distribution are:

To get the extent of non response.
To detect the presence of extreme cases (outliers in the distribution).
To get the extent of illegitimate responses.

d. All of the above

Factor analysis is a technique used for:

a. Univariate data b. Bivariate data

c. Multivariate data d. Both (a) and (b)

e. Both (a) and (c) f. Both (b) and (c)

12.11 LESSON END EXERCISE

Explain the concept of Univariate data analysis with some examples.

____

____

____

Briefly discuss exploratory factor analysis.

____

____

____

How discriminant analysis is different from cluster analysis?

____

____

____

Explain different types of correlation under bivariate data analysis.

____

____

____

Differentiate between correlation and regression analysis.

____

____

____

Discuss the procedure of analysing data using univariate data analysis technique through SPSS.

12.12 SUGGESTED READINGS

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

William G. Z. Business Research Methods. Thomson, India.

Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

UNIT-III **Lesson No. 13**

CONCEPT AND APPLICATIONS OF EXPLORATORY FACTORY ANALYSIS, CONFIRMATORY FACTOR ANALYSIS AND SEM IN BUSINESS RESEARCH

STRUCTURE

13.1 INTRODUCTION

Data is everywhere. From data research to artificial intelligence technology, data has become an essential commodity that is being perceived as a link between our past and future. Also, data is the key solution to this problem. While the world has moved on to technology, it still is unaware of the fact that data is the building block of all these technological advancements that have together made the world so advanced. Number of tools and techniques are put to work to arrange, organise, and accumulate data the way one wants to. Factor Analysis is one of them. A data reduction technique, Factor Analysis is a statistical method used to reduce the number of observed factors for a much better insight into a given dataset. But first, we should know what a factor is. A factor is a set of observed variables that have similar responses to an action. Since variables in a given dataset can be too much to deal with, Factor Analysis condenses these factors or variables into fewer variables that are actionable and substantial to work upon. It is works on narrowing the availability of variables in a given data set, allowing deeper insights and better visibility of patterns for data research and is most commonly used to identify the relationship between various variables in statistics. Thus, "FA is considered an extension of principal component analysis since the ultimate objective for both techniques is a data reduction".

Developed in 1904 by Spearman, Factor Analysis is broadly divided into various types based upon the approach to detect underlying variables and establish a relationship between them. While there are a variety of techniques to conduct factor analysis like Principal Component Analysis or Independent Component Analysis, Factor Analysis can be divided into 2 types i.e. exploratory factor analysis and confirmatory factor analysis. In the case of Exploratory Factor Statistical Analysis, the purpose is to determine/ explore the underlying latent structure of a large set of variables. EFA, unlike CFA, tends to uncover the relationship, if any, between measured variables of an entity (for example -

height, weight, etc. in a human figure). Conducting Exploratory Factor Analysis involves figuring the total number of factors involved in a dataset. "EFA is generally considered to be more of a theory-generating procedure than a theory-testing procedure. Confirmatory Factor Analysis (CFA) determines whether a relationship between factors or a set of observed variables and their underlying components exists. It helps to confirm whether there is a connection between two components of variables in a given dataset. Usually, the purpose of CFA is to test whether certain data fit the requirements of a particular hypothesis. The process begins with a researcher formulating a hypothesis that is made to fit along the lines of a certain theory. If the constraints imposed on a model do not fit well with the data, then the model is rejected, and it is confirmed that no relationship exists between a factor and its underlying construct. Therefore, Confirmatory factor analysis (CFA) is generally based on a strong theoretical and/or empirical foundation that allows the researcher to specify an exact factor model in advance. This CFA works on finding a relationship between a set of observed variables and their underlying structure, this works to uncover a relationship between various variables within a given dataset. Therefore, in this lesson our focus is on two important types of factor analysis, viz., exploratory factor analysis and confirmatory factor analysis.

13.2 OBJECTIVES

After studying this lesson, you will be able to:

Define the concept of factor analysis.

Explain the type of factor analysis.

Discuss the concept and applications of exploratory factor analysis.

Discuss the concept and applications of confirmatory factor analysis.

Discuss the concept and applications of structural equation modelling.

13.3 EXPLORATORYFACTORANALYSIS

Exploratory factor analysis is a statistical technique that is used to reduce data to a smaller set of summary variables and to explore the underlying theoretical structure of the phenomena. It is used to identify the structure of the relationship between the variable and the respondent. Exploratory factor analysis can be performed by using the following two methods:

R-type factor analysis: When factors are calculated from the correlation matrix, then it is called R-type factor analysis.

Q-type factor analysis: When factors are calculated from the individual respondent, then it said to be Q-type factor analysis.

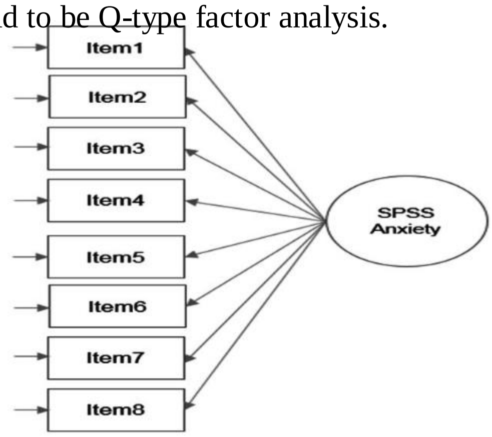


Fig: 13.1: Exploratory Factor Analysis

13.3.1 Concept

In multivariate statistics, exploratory factor analysis (EFA) is a statistical method used to uncover the underlying structure of a relatively large set of variables. EFA is a technique within factor analysis whose overarching goal is to identify the underlying relationships between measured variables. It is commonly used by researchers when developing a scale (a scale is a collection of questions used to measure a particular research topic) and serves to identify a set of latent constructs underlying a battery of measured variables. It should be used when the researcher has no a priori hypothesis about factors or patterns of measured variables. Measured variables are any one of several attributes of people that may be observed and measured. Examples of measured variables could be the physical height, weight, and pulse rate of a human being. Usually, researchers would have a large number of measured variables, which are assumed to be related to a smaller number of "unobserved" factors. Researchers must carefully consider the number of measured variables to include in the analysis. EFA procedures are more accurate when each factor is represented by multiple measured variables in the analysis.

There are two methods for driving factors, these two methods are as follows:

Principle component factor analysis method: This method is used when we need to drive the minimum number of factors and explain the maximum portion of variance in the original variable.

Common factor analysis: This method is used when the researchers do not know the nature of the factor to be extracted and the common error variance.

Selection of Factors to be Extracted

Theory is the first criteria to determine the number of factors to be extracted. From theory, we know that the number of factors extracted does make sense. Most researchers use the Eigen value criteria for the number of factors to be extracted. Value of the percentage and variance explained method is also used for exploratory factor analysis. We can use the scree test criteria for the selection of factors. In this method, Eigen value is plotted on a graph and factors are selected.

Orthogonal Rotation

In this method, axis are maintained at 90 degrees, thus the factors are uncorrelated to each other. In orthogonal rotation, the following three methods are available based on the rotation:

Quartimax: Rows are simplified so that the variable should be loaded on a single factor.

Varimax: Used to simplify the column of the factor matrix so that the factor extracts are clearly associated and there should be some separation among the variables.

Equimax: The combination of the above two methods. This method simplifies row and column at a single time.

Criteria for Practical and Statistical Significance of Factor Loadings

Factor loadings can be classified based on their magnitude:

Greater than + .30 - minimum consideration level

.40 - more important

.50 - practically significant

Power and Significance level

The researcher can determine the statistical power and significance level. For instance, in order to achieve a factor loading of .55 with a power of .80, a sample of 100 is needed. Factor analysis can be performed in SPSS by clicking on "analysis" from menu, and then selecting "factor" from the data reduction option.

13.3.2 Assumptions

Variables used should be metric. Dummy variables can also be considered, but only in special cases.

Sample size: Sample size should be more than 200. In some cases, sample size may be considered for 5 observations per variable.

Homogeneous sample: A sample should be homogenous. Violation of this assumption increases the sample size as the number of variables increases. Reliability analysis is conducted to check the homogeneity between variables.

In exploratory factor analysis, multivariate normality is not required.

Correlation: At least 0.30 correlations are required between the research variables.

There should be no outliers in the data.

13.3.3 Applications of Exploratory Factor Analysis in Business Research

Factor analysis is a statistical technique in which a multitude of variables is reduced to a lesser number of factors. In the marketing world, it is used to collectively analyse several successful marketing campaigns to derive common success factors. This, in turn, helps companies understand the customer better. Factor analysis is sometimes termed a "data reduction" technique because the method is frequently used to extract a few underlying components (or factors) from a large initial set of observed variables. It is extensively used in psychological research concerned with the construction of scales intended to measure attitudes, perceptions, motivations, and so forth.

Business-related applications are numerous and examples include the development of scales used to measure customer satisfaction with products and employee work attitudes. Factor analysis, however, has applicability outside of the realm of psychological research. It may be used, for example, by financial analysts to identify groups of stocks in which prices fluctuate in similar ways. And factor analysis often plays a crucial role in establishing the validity of employment tests and performance appraisal methods, thus helping a firm defend itself against employment discrimination charges.

An example of the use of factor analysis might involve research designed to construct a scale of employee job satisfaction. Initially, a researcher or consultant may assemble a large set of questionnaire items that seem to be related to job satisfaction. These items will generally be presented to subjects along with some type of numeric or verbal scale.

The job satisfaction questionnaire may include several dozen such questions. What is really of interest, however, are employee views regarding underlying dimensions of job satisfaction. Typically, there are only a few such dimensions, which are psychological states that cannot be directly measured. Such dimensions are called “factors” and factor analysis is used to assess them indirectly.

In this case, the basic factor analysis model assumes that employee responses to each of the job satisfaction items in the questionnaire can be condensed into one or more underlying factors. Sometimes the researcher will have some expectation as to the number of factors, although such an assumption is not necessary. The factors are assumed to be related to the score of each item on the questionnaire in a linear manner. Suppose that all of the answers to the job satisfaction items derived from two underlying factors, plus some random element (perhaps due to measurement error). If so, then a respondent’s answer to a particular item could be decomposed into those basic factors according to the following equation: where: scorei = subject’s score on questionnaire item i; a_{1i} = coefficient relating factor 1 to $score_i$; $factor_1$ = value of the first factor for the subject; a_{2i} = coefficient relating factor 2 to $score_i$; $factor_2$ = value of the second factor for the subject; random = random error.

The underlying factors can be thought of as the subject's true feelings with respect to his or her job. The researcher, however, has only the subject's answers to specific questions regarding the characteristics of the employee's job. Since these are measured with some error, there is also a random component to the observed value. Separate equations can be written for all of the items in the questionnaire. For each item, the coefficients (a $_{1i}$ and a $_{2i}$) will probably be different. These coefficients are usually referred to as "factor loadings." While the factors cannot be measured directly, it is possible to estimate factor loadings indirectly. Estimates of factor loadings are derived from the matrix containing the inter correlations of all of the observed scores for a large number of subjects. Each subject in the sample answers all of the questionnaire items and the matrix contains all possible correlations between pairs of these items. Most general statistical programs, such as the Statistical Package for the Social Services (SPSS), SYSDAT, and SAS, perform factor analysis.

With immense use in various fields in real life, this segment presents a list of applications of Factor Analysis and the way FA is used in day-to-day operations.

Marketing: Marketing is defined as the act of promoting a good or a service or even a brand. When it comes to Factor Analysis in marketing, one can benefit immensely from this statistical method. In order to boost marketing campaigns and accelerate success, in the long run, companies employ Factor Analysis techniques that help to find a correlation between various variables or factors of a marketing campaign. Moreover, FA also helps to establish connections with customer satisfaction and consequent feedback after a marketing campaign in order to check its efficacy and impact on the audiences. That said, the realm of marketing can largely benefit from Factor Analysis and trigger sales with respect to much-enhanced feedback and customer satisfaction reports.

Data Mining: In data mining, Factor Analysis can play a role as important as that of artificial intelligence. Owing to its ability to transform a complex and vast dataset into a group of filtered out variables that are related to each other in some way or the other, FA eases out the process of data mining. For data scientists, the tedious task of finding relationships and establishing correlation among various variables has always been full of obstacles and errors. However, with the help of this statistical method, data mining has become much more advanced.

Machine Learning: Machine Learning and data mining tools go hand in hand. Perhaps this is the reason why Factor Analysis finds a place among Machine Learning tools and techniques. As Factor Analysis in machine learning helps in reducing the number of variables in a given dataset to procure a more accurate and enhanced set of observed factors, various machine learning algorithms are put to use to work accordingly. They are trained well with humongous data to rightly work in order to give way to other applications. An unsupervised machine learning algorithm, FA is largely used for dimensionality reduction in machine learning. Thereby, machine learning can very well collaborate with Factor Analysis to give rise to data mining techniques and make the task of data research massively efficient.

Nutritional Science: Nutritional Science is a prominent field of work in the contemporary scenario. By focusing on the dietary practices of a given population, Factor Analysis helps to establish a relationship between the consumption of nutrients in an adult's diet and the nutritional health of that person. Furthermore, an individual's nutrient intake and consequent health status have helped nutritionists to compute the appropriate quantity of nutrients one should intake in a given period of time.

Business: The application of Factor Analysis in business is rather surprising and satisfactory. Remember the times when business firms had to employ professionals to dig out patterns from past records in order to lay a road ahead for strategic business plans? Well, gone are the days when so much work had to be done. Thanks to Factor Analysis, the world of business can use it for eliminating the guesswork and formulating more accurate and straightforward decisions in various aspects like budgeting, marketing, production, and transport.

Market Research Surveys: If a marketer looking to research a market through surveys to determine people's satisfaction with a product. Satisfaction is hard to quantitate and it's even harder to determine what factors lead to it. In the marketing world, factor analysis has been used to develop things such as perception maps, advanced SWOT analyses, etc. It gathers unobservable variables such as customer happiness and quantitate their values against observed variables such as increased sales. For this the marketer need to design a questionnaire that would focus on questions pertaining to customer satisfaction. This would include variables like product durability, packaging, chance of reuse etc. After the customers fill the surveys, marketer would then do a factor analysis that would extract the variance of each variable and tell which factor matters the most on customer satisfaction. If he/she find highly correlated answers for separate variables/questions, he would combine them in a single factor for future analyses.

13.4 CONFIRMATORYFACTORANALYSIS

"Confirmatory factor analysis" is a much different process. Here, the researcher generally has an idea both as to the number of underlying factors and the factor structure. This means he or she specifies which of the observed variables are associated with which of the hypothesised factors. The factors themselves may be correlated. If certain mathematical conditions are met, then the values of all of the factor loadings identified can be estimated; no subsequent rotation of the factor loadings is required. Confirmatory factor analysis methods also allow for various tests of statistical significance, both for the model as a whole and for individual parameters, such as factor loadings. Thus tests of statistical significance exist for individual factor loadings (as for coefficients in multiple regression analysis) and several different overall fit statistics are used to assess the general adequacy of the model.

13.4.1 Concept

The CFA is a multivariate statistic that serves to estimate the structure of an instrument, verifying how well the measured variables represent the number of constructs. That is, it verifies whether an instrument's structure can be, but is not necessarily, true. For this, we need to state which structure we want to test. Generally, the CFA is used when there is a previous study that tells us the dimensionality of that instrument. For instance, we would have a North American study that uses an EFA to verify the instrument's dimensionality and you use a CFA to verify how well this structure happens with Brazilian data. However, this is not the only way you can use the CFA! You can, for example, have the EFA in the same study (to explore the dimensionality), but still test different theoretical models using the CFA.

Thus, both EFA and CFA are applied when you want to estimate the dimensionality of an instrument (note that I said estimate, not explore/discover dimensionality). For example, we can apply the CFA in self-report instruments, where items represent behaviours, thoughts, or feelings. Another example, we can apply it in a set of other measures, such as psychophysical measures of anxiety. Thus, the CFA applies to instruments that measure some attributes such as well-being, anxiety, prejudice, etc.

In using confirmatory factor analysis, the researcher must posit one or more specific models, defining the number of factors and which variables are hypothesized to be determined by (i.e., to have loadings with) which factors. Unlike exploratory factor analysis, in which all observed variables are assumed to load on all factors, many or all of the observed variables in confirmatory factor analysis will generally be hypothesized to load on fewer than the specified number of factors, often on only one of the factors (i.e., the factor loads between an observed variable and all of the other factors is, in effect, fixed to a value of zero). Once a model is specified, its factor loadings and other parameters (e.g., error variances) can be estimated. This is accomplished via a mathematical process in which the estimated loadings and other parameters are used to compute a predicted correlation matrix for the observed variables. With a predicted correlation matrix, the parameters are systematically updated via an iterative process until the difference between the actual and predicted correlation matrices are minimized given the hypothesized structure of the model, at which point the results are said to "converge."

There are several goodness-of-fit statistics for the overall model, with the basic statistic following a "chi-square distribution." The chi-square test can be used to determine if the model as a whole reasonably reproduces the observed correlation matrix. If so, then the researcher would conclude that the analysis tends to confirm the validity of the hypothesized model (assuming the values of the factor loadings are also of the expected signs). Since the chi-square test is highly sensitive to the sample size, however, it often leads to rejection of the models. For this reason, other criteria, such as the "goodness-of-fit index," are often used to assess these models. Such alternative criteria are not size sensitive, but also do not have sampling distributions that can be used to perform tests of statistical significance. Thus, such tests are rather judgmental, though there is general agreement as to appropriate values for these indexes to achieve in order to confirm a model.

It is also possible to use confirmatory factor analysis to compare two different models in order to determine which one fits better. For example, a researcher may wish to determine if a one-factor model of job satisfaction, in which all observed indicators of satisfaction load on a single factor, is inferior to a two-factor model, in which satisfaction with economic aspects of the job load only on one factor and satisfaction with intrinsic aspects load only on a second factor. Both models would be estimated and the one with the better measures of fit would be the preferred.

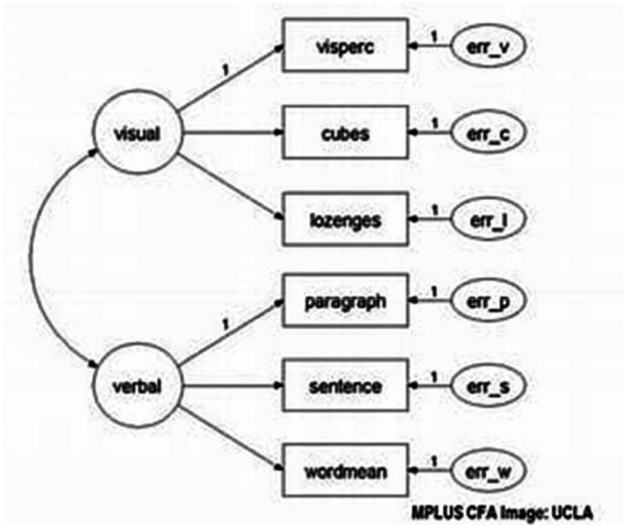


Fig 13.2: Confirmatory Factor Analysis

13.4.2 Applications Confirmatory Factor Analysis in Business Research

Confirmatory factor analysis (CFA) is a multivariate statistical procedure that is used to test how well the measured variables represent the number of constructs. Confirmatory factor analysis (CFA) and exploratory factor analysis (EFA) are similar techniques, but in exploratory factor analysis (EFA), data is simply explored and provides information about the numbers of factors required to represent the data. In exploratory factor analysis, all measured variables are related to every latent variable. But in confirmatory factor analysis (CFA), researchers can specify the number of factors required in the data and which measured variable is related to which latent variable. Therefore, Confirmatory factor analysis (CFA) is a tool that is used to confirm or reject the measurement theory.

Conditions for using confirmatory factor analysis (CFA

Developing the overall measurement model theory: In confirmatory factor analysis (CFA), we should consider the concept of unidimensionality between construct error variance and within construct error variance. At least four constructs and three items per constructs should be present in the research.

Designing a study to produce the empirical results: The measurement model must be specified. Most commonly, the value of one loading estimate should be one per construct. Two methods are available for identification; the first is rank condition, and the second is order condition.

Assessing the measurement model validity: Assessing the measurement model validity occurs when the theoretical measurement model is compared with the reality model to see how well the data fits. To check the measurement model validity, the number of the indicator helps us. For example, the factor loading latent variable should be greater than 0.7. Chi-square test and other goodness of fit statistics like RMR, GFI, NFI, RMSEA, SIC, BIC, etc., are some key indicators that help in measuring the model validity.

The assumptions of a CFA include multivariate normality, a sufficient sample size (n >200), the correct a priori model specification, and data must come from a random sample

Usually, statistical software like AMOS, LISREL, EQS and SAS are used for confirmatory factor analysis. In AMOS, visual paths are manually drawn on the graphic window and analysis is performed. In LISREL, confirmatory factor analysis can be performed graphically as well as from the menu. In SAS, confirmatory factor analysis can be performed by using the programming languages.

13.4.3 Evaluating Model Fit

In CFA, several statistical tests are used to determine how well the model fits to the data. A good fit between the model and the data does not mean that the model is “correct”, or even that it explains a large proportion of the covariance. A “good model fit” only indicates that the model is plausible. When reporting the results of a confirmatory factor analysis, one is urged to report:

the proposed models,

any modifications made,

which measures identify each latent variable,

correlations between latent variables,

any other pertinent information, such as whether constraints are used.

With regard to selecting model fit statistics to report, one should not simply report the statistics that estimate the best fit, though this may be tempting. Though several varying opinions exist, Kline (2010) recommends reporting the chi-squared test, the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the standardised root mean square residual (SRMR).

Absolute fit indices

Absolute fit indices determine how well the a priori model fits, or reproduces the data. Absolute fit indices include, but are not limited to, the Chi-Squared test, RMSEA, GFI, AGFI, RMR, and SRMR.

Chi-squared test : The chi-squared test indicates the difference between observed and expected covariance matrices. Values closer to zero indicate a better fit; smaller difference between expected and observed covariance matrices. Chi-squared statistics can also be used to directly compare the fit of nested models to the data. One difficulty with the chi-squared test of model fit, however, is that researchers may fail to reject an inappropriate model in small sample sizes and reject an appropriate model in large sample sizes. As a result, other measures of fit have been developed.

Root mean square error of approximation : The root mean square error of approximation (RMSEA) avoids issues of sample size by analysing the discrepancy between the hypothesized model, with optimally chosen parameter estimates, and the population covariance matrix. The RMSEA ranges from 0 to 1, with smaller values indicating better model fit. A value of .06 or less is indicative of acceptable model fit.

Root mean square residual and standardized root mean square residual :

The root mean square residual (RMR) and standardized root mean square residual (SRMR) are the square root of the discrepancy between the sample covariance matrix and the model covariance matrix. The RMR may be somewhat difficult to interpret, however, as its range is based on the scales of the indicators in the model (this becomes tricky when you have multiple indicators with varying scales; e.g., two questionnaires, one on a 0–10 scale, the other on a 1–3 scale). The standardized root mean square residual removes this difficulty in interpretation, and ranges from 0 to 1, with a value of .08 or less being indicative of an acceptable model.

Goodness of fit index and adjusted goodness of fit index : The goodness of fit index (GFI) is a measure of fit between the hypothesized model and the observed covariance matrix. The adjusted goodness of fit index (AGFI) corrects the GFI, which is affected by the number of indicators of each latent variable. The GFI and AGFI range between 0 and 1, with a value of over .9 generally indicating acceptable model fit.

Relative Fit Indices

Relative fit indices (also called "incremental fit indices" and "comparative fit indices") compare the chi-square for the hypothesised model to one from a "null", or "baseline" model. This null model almost always contains a model in which all of the variables are uncorrelated, and as a result, has a very large chi-square (indicating poor fit). Relative fit indices include the normed fit index and comparative fit index.

Normed fit index and non-normed fit index : The normed fit index (NFI) analyses the discrepancy between the chi-squared value of the hypothesised model and the chi-squared value of the null model. However, NFI tends to be negatively biased. The non-normed fit index (NNFI; also known as the Tucker–Lewis index, as it was built on an index formed by Tucker and Lewis, in 1973) resolves some of the issues of negative bias, though NNFI values may sometimes fall beyond the 0 to 1 range. Values for both the NFI and NNFI should range between 0 and 1, with a cut off of .95 or greater indicating a good model fit.

Comparative Fit Index : The comparative fit index (CFI) analyses the model fit by examining the discrepancy between the data and the hypothesised model, while adjusting for the issues of sample size inherent in the chi-squared test of model fit, and the normed fit index. CFI values range from 0 to 1, with larger values indicating better fit. Previously, a

CFI value of .90 or larger was considered to be an acceptable model fit. However, recent studies have indicated that a value greater than .90 is needed to ensure that misspecified models are not deemed acceptable. Thus, a CFI value of .95 or higher is presently accepted as an indicator of good fit.

13.4.4 Confirmatory Analysis Process

To complete a successful confirmatory factor analysis, it's important to follow several key steps. Although confirmatory factor analysis is mostly an automated software procedure, researchers are in charge of setting up the analysis, selecting variables and interpreting results. Here are the primary steps in the CFA;

Specify the Latent Variable

Start by determining what concept you want to analyse and establishing its theoretical definition. Establishing a baseline to describe the latent variable makes it possible for researcher to evaluate the accuracy of observed variables. We can define the latent variable by making a list of characteristics or collecting additional data. For example, if we want to use CFA to determine if an intake survey is a good assessment of self-esteem, start by defining self-esteem. We may use our professional background to determine that having self-esteem involves presenting traits like confidence, sociability, adaptability and having goals.

Determine Measurement Methods

Next step is to identify the measurement method we want to test and which observed variables to include. These variables are usually survey questions. We may include multiple questions from the same survey or choose questions from different surveys depending on the type of analysis we want to conduct. Here are some examples of observed variables from a survey assessing self-esteem:

Rank your confidence from one to five.

Agree or disagree: I am uncomfortable accepting compliments from others.

Rate your adaptability from one to five.

Collect the data

Gather the information we want to use in our confirmatory factor analysis. Decide if we want to collect original responses from our own research or if we want to use outside data from other sources. Try to secure a large sample size of information to ensure an accurate analysis. Once we have enough quality information, input it into statistical modelling software.

Establish Consistent Parameters

Using statistical modelling software, establish standardised parameters to evaluate the latent and observed variables. Decide what measurement system we want to use as the standard and allow the software to convert all other values to that measurement. For example, if we want to use the rating system of one to five as the standard measurement, we first need to convert all other types of questions to that same format. For questions that ask respondents to agree or disagree to a statement, assign a value of one to “disagree” and the value five to “agree.” This allows the software to compute all types of data consistently.

Compute the Data

Use statistical modelling software to compute the factor loadings for your data. Follow the prompts for specific software interface to produce results. Most factor analysis software produces this information in a table, although some generate graphs and tables to express the same information.

Interpretation

Review the factor loading column of the factor analysis table to determine how well each observed variable relates to the latent variable. Decide what factor loading value indicates a significant relationship, and use that to guide your interpretation. For example, you may decide that any variables with a factor loading of 0.75 are valid for assessing self-esteem. If all the survey questions have a factor loading over 0.75, you can conclude that your survey is a good overall measurement of self-esteem.

13.5 STRUCTURALEQUATION MODELLING (SEM)

Structural equation modelling software is typically used for performing confirmatory factor analysis. CFA is also frequently used as a first step to assess the proposed measurement model in a structural equation model. Many of the rules of interpretation regarding assessment of model fit and model modification in structural equation modelling apply equally to CFA. CFA is distinguished from structural equation modelling by the fact that in CFA, there are no directed arrows between latent factors. In other words, while in CFA factors are not presumed to directly cause one another, SEM often does specify particular factors and variables to be causal in nature. In the context of SEM, the CFA is often called 'the measurement model', while the relations between the latent variables (with directed arrows) are called 'the structural model'.

13.5.1 Concept

Structural equation modeling is a statistical technique used for testing and estimating causal relations using a combination of statistical data and qualitative causal assumptions. This concept of SEM was given by the Sewall Wright, Trygve Haavelmo and Herbert A. Simon, and formally defined by Judea Pearl using a calculus of counterfactuals.

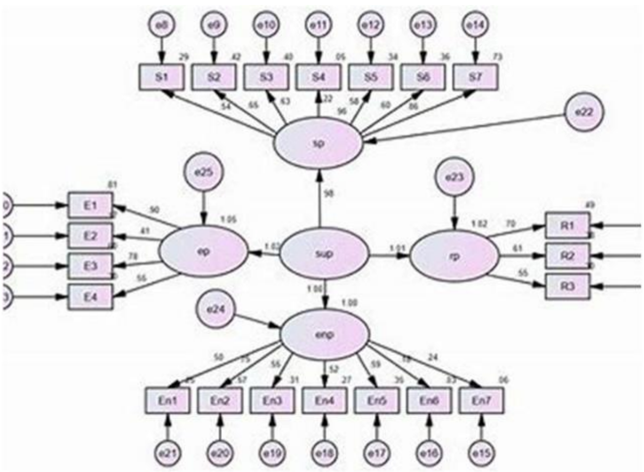


Figure 13.3: Structure Equation Modelling

According to Hoyle, Structural equation modeling may also be explained as a comprehensive statistical approach used for testing hypotheses exploring relations between observed and latent variables. It is a methodology for representing, estimating, and testing a theoretical network of (mostly) linear relations between variables. SEM is used to achieve two important goals:

To understand the patterns of correlation/covariance among a set of variables.

Explaining as much of their variance as possible with the model specified.

Structural equation modelling may also be defined as a multivariate statistical analysis technique that is used for analysing structural relationships. This technique may better be explained as a combination of factor analysis and multiple regression analysis.

SEM is used to analyse the structural relationship between measured variables and latent constructs. Largely preferred by the researchers SEM estimates the multiple and interrelated dependence in a single analysis. Structural equation models are inclusive of both confirmatory and exploratory modelling. Confirmatory modelling usually starts out with a hypothesis that gets represented in a causal model. The concepts used in the model must then be operationalised to allow testing of the relationships between the concepts in the model.

Structural Equation Modelling Examples can better be explained with Structural Equation Models (SEM). The models of Structural equation are a subset of graphical models.

Each Structural equation model is associated with a graph that represents the causal structure of the model and the form of the linear equations. There is a directed edge from X to Y (X Y) if the coefficient of X in the structural equation for Y is nonzero (i.e., X is a direct cause of Y). In addition, there is a bi-directed edge between the error terms [X and [Y if and only if the covariance between the error terms is nonzero.

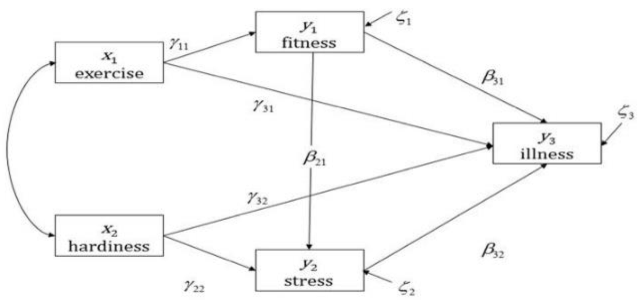
Here, in path diagrams, latent substantive variables are enclosed in ovals, and measured variables are enclosed in rectangles.

In the structural modelling, theoretical latent constructs are represented with ovals or circles and their measured empirical indicators are represented with rectangles. The lines connecting indicators to constructs and constructs to each other carry numerical values that quantify the degree of co-variation accounted for by the model components. Whereas some investigators claim that Structural Equation Modelling explains variation among latent variables; we may explain in a better way by saying that SEMs predict or account for variation among model components. Structural Equation Models are incapable of providing causal mechanism information because they were not designed to do so. The same is true for the path and regression models.

Structural Modelling falls into four broad categories. These structural equation models are Path Analysis, Latent Variable Structural Model, Growth Curve Model, and Latent Growth Model.

Path Analysis

Path Analysis, one of the major structural equation models in use is the application of structural equation modelling without latent variables. The best part about Path Analysis is that it includes the relationships among variables that serve as predictors in one single model. A classic example of this may be a mediation model.



Path Analysis – Only Observed Variables

Figure : 13.4

In a path analysis model from the correlation matrix, two or more casual models are compared. The path of the model is shown by a square and an arrow, which shows the causation. Regression weight is predicted by the model.

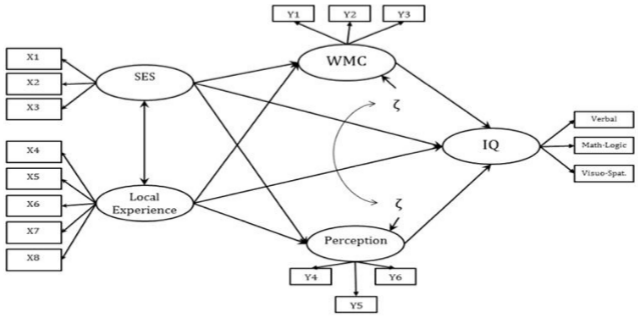
Confirmatory Factor Analysis

Confirmatory Factor Analysis, also known as CFA, is a way ahead of data reduction. CFA is also known within SEM as the measurement model because is the step taken to determine how the factors (å1 and å1) are measured by the indicators (x1 to x8). The Confirmatory Factor Model in SEM treats intelligence as a latent variable which can be measured on the basis of test scores. These are spread out in four areas: reading, writing, math, and analysis.

Latent Variable Structural Model

The next structural equation model for analysis is the Latent Variable Structural Model.

Very next step is to fit the structural model, which is what you probably think of when you hear about SEM. It is mainly using the measured latent variables within the path analysis framework. Once you have declared the latent variables you can hypothesize and test their relationships.



Latent Variable Structural Model

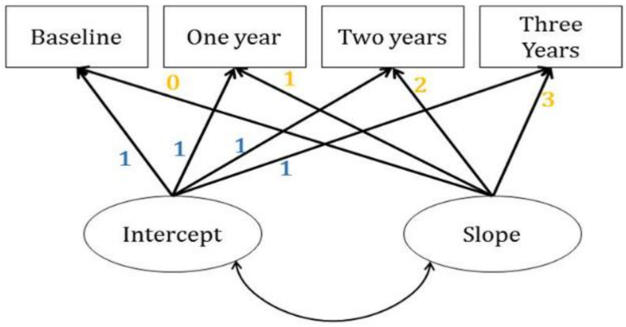
Figure : 13.5

Growth Curve Models

Another popular use of Structural Equation Modelling is longitudinal models, commonly referred to as Growth Curve Models. Let's say for instance you have multiple observations of the same variable over time, you may declare an intercept. A slope for the subjects 'papers over time as latent variables by constraining the path coefficients in a specific way.

Since the paths are constrained, we have to estimate on growth curve modelling the means of the latent variables. These means give us the overall intercept and the overall slope across all subjects.

Latent Growth Curve Models are related to and offer an alternative means of running Mixed Models on longitudinal data. These mixed models are often known as Individual Growth Curve Models.



Latent Growth Model

Figure : 13.6

13.5.2 Applications of SEM in Business Research

Structural Equation Modelling (SEM) is the combination of related methods and not a single technique. The methods are flexible and the framework is for data analysis. Researchers prefer these methods because it enables them to estimate multiple and interrelated dependencies in a single analysis. Structural equation modelling uses two types of variables, endogenous and exogenous. Further, in any organisation, marketing is very important. And to be successful in marketing, one must know about the consumers. They must know their attitude, opinions and personality traits. But these characteristics are latent and cannot be measured easily because they are often abstract. As of now, to measure them, we can conduct surveys, create an observation model, and so forth. But these processes are not that much fruitful because measuring and observing has the drawback of errors.

Structural equation modelling excels at both tasks and uses. Structural equation modelling uses factor analysis and multiple regression analysis simultaneously. If we use both these analytic methods individually, we miss out on the flexibility. So, SEM provides us with flexibility. It is suited for causal analysis, multi-collinearity, which is correlating independent variables.

SEM is the knot that ties the components and elements of the measurement model and relates the component and elements together or to other independent variables. In some cases, variables are combined on empirical grounds. The combining act happens prior to factor analysis, and the measurement model has no role. In other cases, when we are only concerned with raw variables, the observed variables are used. And lastly when there is no measurement model, then the structure model follows the path analysis. Structural equation modelling is also used to analyse survey data. It is not bound to one data source and can be used with customer transaction, economic, social media, and customers' transaction data. Recently it is used in neuroscience for MRI data. In its modern forms, it can be used with any data type-the model uses data types such as ratio, interval, ordinal, nominal, and count. They help to model curvilinear relationships among variables.

Furthermore, SEM is widely used for longitudinal, mixed, and hierarchical modelling and in segmentation also. The model accommodates multiple dependent variables such as the Conjoint Analysis. Structural equation modelling is also used to fix response style issues in consumer surveys.

There may be a business case that needs you to focus on consumer perceptions such as purchase interest, liking, in your product. Though this is a complex modelling task, structural equation modelling is apt for these objectives. Structural equation modelling is used for simpler jobs, such as for a consumer survey.

Structural Equation Mixture Modelling (SEMM) is another type of method to target the hidden segments of consumers with very numerous amounts of data. One must not assume that one type of model is suitable for any kind of analysis. Mixture modelling sometimes works only when the effort is made competently.

When you are working in an environment in which non-experimental designs were common such as industrial or organisational psychology, structural equation modelling is required. Structural equation modelling is widely used and is being used by reviewers for data analysis. The reviewers are often clueless about how to proceed further.

The major advantage of Structural equation modelling is that it allows to tests for theoretical propositions and also enables you to evaluate quantitative predictions.

Some Advanced Applications of Structural Equation Modelling

Measurement invariance

It is the technique that allows the joint estimation of multiple models, each with different sub-groups. Applications that include analysis of differences between groups such as cultures, gender, and so forth and behaviour genetics.

Latent growth modelling

Hierarchical/multilevel models

Mixture model (latent class) Structural Equation

Modelling Alternative estimation and testing techniques.

Robust inference

Survey sampling analyses

Multi-method

Multi-trait models

Structural Equation Model Trees

Therefore, SEM is a technique used for examining complex interactions between variables and reducing them to visual representations. Its key benefit is that it allows us to assess the model's efficiency. There are two types of variables employed in this analysis, endogenous variables and exogenous variables. Endogenous variables are dependent variables and exogenous variables are same as the independent variables. Thus, the techniques of SEM help in building the relationship among these variables. It helps present multiple and interconnected dependencies in a single analysis. SEM is most commonly applied in studies that are intended to validate a research study design rather than to analyse or explain a phenomenon. If a researcher is interested in the strength of the relationship between variables in a hypothesis, structural equation modelling offers a means to investigate those correlations without committing to a costly research effort.

13.5.3 Structural Equation Modelling Process

The Structural equation modelling analysis proceeds through the following steps:

Research the relevant theory

Review literature to support model specification

Specifies model such as diagram and equations

determines the number of degrees of freedom and the model identification to estimate the parameters to find unique values

Selecting the measurement methods for the variables represented in the model

Collect data

Perform preliminary descriptive statistical analysis such as missing data, scaling, and collinearity issues

Estimate the model parameters

Estimate model fit

Specify the meaningful mode

Interpret results

Present results

Structural Equation Modelling Specific Software

LISREL was the fitting structural equation models software in the 1970s.

The Open Mx R package is an R open-source that provides an open-source and an updated version of the Mx application.

The goals of structural equation modeling are to understand the correlated patterns among a set of variables and explain the variances as much as possible.

13.5.4 Advantages of Structural Equation Modelling (SEM)

It is a collection of statistical procedures for testing hypotheses based on numerous constructs that may be related indirectly or directly in both linear and nonlinear models. It differs from other types of analyses in that it may evaluate a large number of associations while also identifying measurement errors.

SEM offers a notable advantage over traditional multiple regression studies in that it has more statistical power. ‘Statistical power’ here refers to the likelihood of rejecting a false null hypothesis).

SEM allows for the examination of latent variables and their relationships, allowing for the examination of psychological construct dependencies without measurement errors.

It can also examine correlated measurement error to see how unknown causes influence shared error between variables. This could change the model’s predicted parameters and can also handle missing data effectively by fitting raw data rather than summary statistics.

13.5 SUMMARY

In this lesson, we have discussed about the concept and applications of exploratory factor analysis, confirmatory factor analysis and structural equation modelling. In “exploratory factor analysis,” the researcher is first confronted with the problem of determining the optimal number of factors to extract from the matrix of correlations among the observed variables (such as responses to questionnaire items). Various criteria are available, but the most common is based on the increment in common variance explained by extracting an additional factor. Exploratory methods typically define factors as uncorrelated to one another; the first factor extracted by the method will explain the greatest amount of common variance in the set of observed items, the second factor will be the uncorrelated (or “orthogonal”) factor that explains the next largest component of common variance, and so forth. “Confirmatory factor analysis” is a much different process. Here, the researcher generally has an idea both as to the number of underlying factors and the factor structure. This means he or she specifies which of the observed variables are associated with which of the hypothesised factors. The factors themselves may be correlated. If certain mathematical conditions are met, then the values of all of the factor loadings identified can be estimated; no subsequent rotation of the factor loadings is required. Confirmatory factor analysis methods also allow for various tests of statistical significance, both for the model as a whole and for individual parameters, such as factor loadings. Thus tests of statistical significance exist for individual factor loadings (as for coefficients in multiple regression analysis) and several different overall fit statistics are used to assess the general adequacy of the model.

Structural equation modelling (SEM) is a method that is very frequently applied by marketing and business researchers to assess empirically new theoretical proposals articulated by means of complex models. Suppose there are various factors that affect an organisation's skills with respect to time management and job stress, such as:

Time management factors: goal setting, operational planning, communication management and meeting management.

Job stress factors: submission, targeting and meetings.

In order to determine how these factors affect organisational skills, SEM can be used. Structural equation modelling is a more comprehensive and flexible method as compared to traditional analysis methods. It is suitable for investigating economic trends, achievements, family and peer dynamics, health issues, self-efficacy, self-concept, depression, exercise, and psychotherapy. Furthermore, traditional approaches are more of a default model-based practice while SEM offers a formulation of model based on requirements and no such default methodology (Suhr, 2006).

Traditional Methods	***Structural Equation Modelling***
More grid.	Flexible and comprehensive.
Follows a default model practice.	Freedom to formulate model, no default model.
Includes only measured variables.	Considers measured and unobserved variables.
Measurement error not considered.	Focuses on measurement error.
Straightforward model fitness.	Methodology of resolving multicollinearity.

13.6 GLOSSARY

Common Variance: Variance in a variable that is shared with other variables.

Communality: The proportion of a variable's variance explained by the extracted factor structure. Final communality estimates are the sum of squared loadings for a variable in an orthogonal factor matrix.

Correlation: The Pearson or product-moment correlation coefficient.

Correlation matrix: A table showing the linear correlations between all pairs of variables.

Data reduction: Reducing the number of variables (e.g., by using factor analysis to determine a smaller number of factors to represent a larger set of factors).

Eigen Value: Column sum of squared loadings for a factor. Represents the variance in the variables which is accounted for by a specific factor.

Exploratory factor analysis: A factor analysis technique used to explore the underlying structure of a collection of observed variables.

Extraction: The process for determining the number of factors to retain.

Factor: Linear combination of the original variables. Factors represent the underlying dimensions (constructs) that summarise or account for the original set of observed variables.

Factor analysis: A statistical technique used to estimate factors and/or reduce the dimensionality of a large number of variables to a fewer number of factors.

Factor loading: Correlation between a variable and a factor, and the key to understanding the nature of a particular factor. Squared factor loadings indicate what percentage of the variance in an original variable is explained by a factor.

Factor rotation: A process of adjusting the factor axes to achieve a simpler and pragmatically more meaningful factor solution - the goal is a usually a simple factor structure.

Measure of sampling adequacy (MSA): Measures which indicate the appropriateness of applying factor analysis.

Principal component analysis (PC or PCA): The factors are based on the total variance of all items.

Theory: A systematic set of causal relationships that provide the comprehensive explanation of a phenomenon.

Model: A specified set of dependent relationships that can be used to test the theory.

Path analysis: Used to test structural equations.

Path diagram: Shows the graphical representation of cause and effect relationships of the theory.

Endogenous variable: The resulting variables that are a causal relationship (also known as dependent variable).

Exogenous variable: The predictor/independent variables.

Confirmatory analysis: Used to test the pre-specified relationship.

Cronbach's alpha: Used to measure the reliability of two or more construct indicators.

Identification: Used to test whether or not there are a sufficient number of equations to solve the unknown coefficient. Identifications are of three types: (1) under-identified, (2) exact identified, and (3) over-identified.

Goodness of fit: The degree to which the observed input matrix is predicted by the estimated model.

Latent variables: Variables that are inferred, not directly observed, from other variables that are observed.

13.7 SELFASSESSMENT QUESTIONS

Multiple Choice Questions:

1. Which of the following is not the part of the exploratory factor analysis process?

    Extracting factors
    Determining the number of factors before the analysis
    Rotating the factors
    Refining and interpreting the factors

Which of the following criteria cannot be used to determine the number of factors in an EFA?

Asking a group of researchers before the analysis
Eigenvalue rule
Scree test
Parallel analysis

Which of the following criteria cannot be used to determine the number of factors in an EFA?

Asking a group of researchers before the analysis
Eigenvalue rule
Scree test
Parallel analysis

What will a factor loading in an orthogonal solution represent?

a. Correlation
b. Partial correlation
c. Multiple correlation
d. Eigenvalue

Which of the following is not a typical structural equation model?

a. Confirmatory factor analysis
b. Latent path analysis
c. Latent mean analysis
d. Exploratory factor analysis

Which of the following criteria are used to assess the quality of a measurement model in SEM?

a. Fit indices
b. Size and significance of factor loadings
c. Discriminant validity
d. Latent path coefficients

Which of the following is not a typical model fit index used in SEM?

Root mean squared error of approximation (RMSEA)
b. Adjusted R-square
c. Comparative fit index (CFI)
Tucker-Lewis index (TLI)

13.8 LESSON END EXERCISE

How EFA does differ from CFA?

____

____

____

Differentiate between EFA from other multivariate data analysis techniques.

____

____

____

Describe how to determine the number of factors to be extract?

____

____

____

How can factor analysis help the researcher to improve the results of other multivariate techniques?

____

____

____

Distinguish between Q type and R type factor analysis.

____

____

____

Distinguish EFA, CFA and SEM

____

____

____

Distinguish between variables and constructs.

____

____

____

13.9 SUGGESTED READINGS

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

William G. Z. Business Research Methods. Thomson, India.

Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

UNIT-III **Lesson No. 14**

TESTINGOFHYPOTHESIS

STRUCTURE

14.1 INTRODUCTION

Hypothesis is usually considered as the principal instrument in research. Its main function is to suggest new experiments and observations. In fact, many experiments are carried out with the deliberate object of testing hypotheses. Decision-makers often face situations wherein they are interested in testing hypotheses on the basis of available information and then take decisions on the basis of such testing. In social science, where direct knowledge of population parameter(s) is rare, hypothesis testing is the often used strategy for deciding whether a sample data offer such support for a hypothesis that generalisation can be made. Thus, hypothesis testing enables us to make probability statements about population parameter(s). The hypothesis may not be proved absolutely, but in practice it is accepted if it has withstood a critical testing. Before we explain how hypotheses are tested through different tests meant for the purpose, it will be appropriate to explain clearly the meaning of a hypothesis and the related concepts for better understanding of the hypothesis testing techniques.

14.2 OBJECTIVES

After studying this lesson, you will be able to:

- Explain the nature and logic of hypothesis testing.
- Describe what a statistically significant difference is.
- Explain the six-step hypothesis testing procedure.
- Discuss differences between parametric and nonparametric tests and when to use each.
- Explain the factors that influence the selection of an appropriate test of statistical significance.
- Explain how to interpret the various test statistics.

14.3 CONCEPTOFHYPOTHESIS

Ordinarily, when one talks about hypothesis, one simply means a mere assumption or some supposition to be proved or disproved. But for a researcher hypothesis is a formal question that he intends to resolve. Thus a hypothesis may be defined as a proposition or a set of proposition set forth as an explanation for the occurrence of some specified group of phenomena either asserted merely as a provisional conjecture to guide some investigation or accepted as highly probable in the light of established facts. Quite often a research hypothesis is a predictive statement, capable of being tested by scientific methods, that relates an independent variable to some dependent variable. For example, consider statements like the following ones:

"Students who receive counselling will show a greater increase in creativity than students not receiving counselling" Or

"the automobile A is performing as well as automobile B."

These are hypotheses capable of being objectively verified and tested. Thus, we may conclude that a hypothesis states what we are looking for and it is a proposition which can be put to a test to determine its validity.

14.3.1 Characteristics of Hypothesis

Hypothesis must possess the following characteristics:

Hypothesis should be clear and precise. If the hypothesis is not clear and precise, the inferences drawn on its basis cannot be taken as reliable.

Hypothesis should be capable of being tested. In a swamp of untestable hypotheses, many a time the research programmes have bogged down. Some prior study may be done by researcher in order to make hypothesis a testable one. A hypothesis "is testable if other deductions can be made from it which, in turn, can be confirmed or disproved by observation."

Hypothesis should state relationship between variables, if it happens to be a relational hypothesis.

Hypothesis should be limited in scope and must be specific. A researcher must remember that narrower hypotheses are generally more testable and he should develop such hypotheses.

Hypothesis should be stated as far as possible in most simple terms so that the same is easily understandable by all concerned. But one must remember that simplicity of hypothesis has nothing to do with its significance.

Hypothesis should be consistent with most known facts i.e., it must be consistent with a substantial body of established facts. In other words, it should be one which judges accept as being the most likely.

Hypothesis should be amenable to testing within a reasonable time. One should not use even an excellent hypothesis, if the same cannot be tested in reasonable time for one cannot spend a life-time collecting data to test it.

Hypothesis must explain the facts that gave rise to the need for explanation. This means that by using the hypothesis plus other known and accepted generalizations, one should be able to deduce the original problem condition. Thus hypothesis must actually explain what it claims to explain; it should have empirical reference.

14.3.2 Basic Concepts Concerning Testing of Hypotheses

Basic concepts in the context of testing of hypotheses need to be explained. These are:

Null hypothesis and alternative hypothesis: In the context of statistical analysis, we often talk about null hypothesis and alternative hypothesis. Null Hypothesis - It gives the statement contrary to working hypothesis. It is a negative statement and you find no relationship between dependent and independent variable here. It is denoted by 'H_0'. Null hypothesis can be simple or complex and directional or non directional. Null hypothesis testing is a formal approach to deciding between two interpretations of a statistical relationship in a sample. One interpretation is called the null hypothesis (often symbolized H_0 and read as "H-naught"). This is the idea that there is no relationship in the population and that the relationship in the sample reflects only sampling error. Informally, the null hypothesis is that the sample relationship "occurred by chance.'' For more clarity, a null hypothesis is a type of hypothesis used in statistics that proposes no statistical significance exists in a set of given observations. The null hypothesis attempts to show that no variation exists between variables or that a single variable is not different than its mean. It is presumed to be true until statistical evidence nullifies it for an alternative hypothesis. Further, it is the proposition that implies no effect or no relationship between phenomena or populations. Any observed difference would be due to sampling error (random chance) or experimental error. The null hypothesis is popular because it can be tested and found to be false, which then implies there is a relationship between the observed data. Also Known As: H_0, no-difference hypothesis. There are two ways to state a null hypothesis. One is to state it as a declarative sentence, and the other is to present it as a mathematical statement.

For example, say a researcher suspects that exercise is correlated to weight loss, assuming a diet remains unchanged. The average length of time to achieve a certain weight loss is an average of 6 weeks when a person works out five times a week. The researcher wants to test whether weight loss takes longer if the number of workouts is reduced to three times a week.

The first step to writing the null hypothesis is to find the (alternate) hypothesis. In a word problem like this, you're looking for what you expect as the outcome of the experiment. In this case, the hypothesis is "I expect weight loss to take longer than 6 weeks."

This can be written mathematically as: H_1: > 6

In this example, ì is the average.

Now, the null hypothesis is what you expect if this hypothesis does not happen. In this case, if weight loss isn't achieved in greater than 6 weeks, then it must occur at a time equal to or less than 6 weeks.

H_0: ≤ 6

The other way to state the null hypothesis is to make no assumption about the outcome of the experiment. In this case, the null hypothesis is simply that the treatment or change will have no effect on the outcome of the experiment. For this example, it would be that reducing the number of workouts would not affect time to achieve weight loss:

H_0: $=6$

"Hyperactivity is unrelated to eating sugar" is an example of a null hypothesis. If the hypothesis is tested and found to be false, using statistics, then a connection between hyperactivity and sugar ingestion may be indicated. A significance test is the most common statistical test used to establish confidence in a null hypothesis.

Another example of a null hypothesis would be, "Plant growth rate is unaffected by the presence of cadmium in the soil." A researcher could test the hypothesis by measuring the rate of plant growth of plants grown in a medium lacking cadmium compared with the rate of growth of plants grown in a medium containing different amounts of cadmium. Disproving the null hypothesis would set the groundwork for further research into the effects of different concentrations of the element in soil.

Sometimes, you may be wondering why you would want to test a hypothesis just to find it false. Why not just test an alternate hypothesis and find it true? The short answer is that it's part of the scientific method. In science, "proving" something doesn't occur. Science uses math to determine the probability a statement is true or false. It turns out it's much easier to disprove a hypothesis than to ever prove one. Also, while the null hypothesis may be simply stated, there's a good chance the alternate hypothesis is incorrect.

For example, if your null hypothesis is that plant growth is unaffected by duration of sunlight, you could state the alternate hypothesis several different ways. Some of these statements might be incorrect. You could say plants are harmed by more than 12 hours of sunlight or that plants need at least 3 hours of sunlight, etc. There are clear exceptions to those alternate hypotheses, so if you test the wrong plants, you could reach the wrong conclusion. The null hypothesis is a general statement that can be used to develop an alternate hypothesis, which may or may not be correct. Testing the null hypothesis can tell you whether your results are due to the effect of manipulating the dependent variable or due to chance. Rejecting a hypothesis does not mean an experiment was "bad" or that it didn't produce results. In fact, it is often one of the first steps toward further inquiry. A significance test is used to determine the likelihood that the results supporting the null hypothesis are not due to chance. A confidence level of 95 percent or 99 percent is common. Keep in mind, even if the confidence level is high, there is still a small chance the null hypothesis is not true, perhaps because the experimenter did not account for a critical factor or because of chance.

Alternative Hypothesis: The other interpretation is called the alternative hypothesis (often symbolised as H_1/H_a). This is the idea that there is a relationship in the population. The alternate hypothesis, H_A or H_1, proposes that observations are influenced by a non random factor. In an experiment, the alternate hypothesis suggests that the experimental or independent variable has an effect on the dependent variable.

The level of significance: This is a very important concept in the context of hypothesis testing. It is always some percentage (usually 5%) which should be chosen with great care, thought and reason. In case we take the significance level at 5 per cent, then this implies that H_0 will be rejected when the sampling result (i.e., observed evidence) has a less than 0.05 probability of occurring if H_0 is true. In other words, the 5 per cent level of significance means that researcher is willing to take as much as a 5 per cent risk of rejecting the null hypothesis when it (H_0) happens to be true. Thus the significance level is the maximum value of the probability of rejecting H_0 when it is true and is usually determined in advance before testing the hypothesis.

Decision rule or test of hypothesis: Given a hypothesis H0 and an alternative hypothesis Ha, we make a rule which is known as decision rule according to which we accept H_0 (i.e., reject Ha) or reject H_0 (i.e., accept Ha). For instance, if (H_0 is that a certain lot is good (there are very few defective items in it) against Ha) that the lot is not good (there are too many defective items in it), then we must decide the number of items to be tested and the criterion for accepting or rejecting the hypothesis. We might test 10 items in the lot and plan our decision saying that if there are none or only 1 defective item among the 10, we will accept H0 otherwise we will reject H_0 (or accept Ha). This sort of basis is known as decision rule.

Type I and Type II errors: In the context of testing of hypotheses, there are basically two types of errors we can make. We may reject H_0 when H_0 is true and we may accept H_0 when in fact H_0 is not true. The former is known as Type I error and the latter as Type II error. In other words, Type I error means rejection of hypothesis which should have been accepted and Type II error means accepting the hypothesis which should have been rejected. Type I error is denoted by á (alpha) known as á error, also called the level of significance of test; and Type II error is denoted by â (beta) known as â error.

The probability of Type I error is usually determined in advance and is understood as the level of significance of testing the hypothesis. If type I error is fixed at 5 per cent, it means that there are about 5 chances in 100 that we will reject H_0 when H_0 is true. We can control Type I error just by fixing it at a lower level. For instance, if we fix it at 1 per cent, we will say that the maximum probability of committing Type I error would only be 0.01.

But with a fixed sample size, n, when we try to reduce Type I error, the probability of committing Type II error increases. Both types of errors cannot be reduced simultaneously. There is a trade-off between two types of errors which means that the probability of making one type of error can only be reduced if we are willing to increase the probability of making the other type of error. To deal with this trade-off in business situations, decision-makers decide the appropriate level of Type I error by examining the costs or penalties attached to both types of errors. If Type I error involves the time and trouble of reworking a batch of chemicals that should have been accepted, whereas Type II error means taking a chance that an entire group of users of this chemical compound will be poisoned, then in such a situation one should prefer a Type I error to a Type II error. As a result one must set very high level for Type I error in one's testing technique of a given hypothesis. Hence, in the testing of hypothesis, one must make all possible effort to strike an adequate balance between Type I and Type II errors.

Two-tailed and One-tailed tests: In the context of hypothesis testing, these two terms are quite important and must be clearly understood. A two-tailed test rejects the null hypothesis if, say, the sample mean is significantly higher or lower than the hypothesised value of the mean of the population. Such a test is appropriate when the null hypothesis is some specified value and the alternative hypothesis is a value not equal to the specified value of the null hypothesis. In a two-tailed test, there are two rejection regions, one on each tail of the curve.

If the significance level is 5 per cent and the two-tailed test is to be applied, the probability of the rejection area will be 0.05 (equally splitted on both tails of the curve as 0.025) and that of the acceptance region will be 0.95 as shown in the above curve. If we take $\mu = 100$ and if our sample mean deviates significantly from 100 in either direction, then we shall reject the null hypothesis; but if the sample mean does not deviate significantly from μ , in that case we shall accept the null hypothesis.

But there are situations when only one-tailed test is considered appropriate. A one-tailed test would be used when we are to test, say, whether the population mean is either lower than or higher than some hypothesised value. For instance, if our H_0: $\mu = \mu_{H0}$ = and Ha: μ

μ_{H0}, then we are interested in what is known as left-tailed test (wherein there is one rejection region only on the left tail).

14.3.3 Procedure for Hypothesis Testing

To test a hypothesis means to tell (on the basis of the data the researcher has collected) whether or not the hypothesis seems to be valid. In hypothesis testing the main question is: whether to accept the null hypothesis or not to accept the null hypothesis? Procedure for hypothesis testing refers to all those steps that we undertake for making a choice between the two actions i.e., rejection and acceptance of a null hypothesis. The various steps involved in hypothesis testing are stated below:

Making a formal statement: The step consists in making a formal statement of the null hypothesis (H_0) and also of the alternative hypothesis (Ha). This means that hypotheses should be clearly stated, considering the nature of the research problem. For instance, Mr. Mohan of the Civil Engineering Department wants to test the load bearing capacity of an old bridge which must be more than 10 tons, in that case he can state his hypotheses as under:

Null hypothesis H_0: μ = 10 tons

Alternative Hypothesis Ha: μ > tons

Take another example. The average score in an aptitude test administered at the national level is 80. To evaluate a state's education system, the average score of 100 of the state's students selected on random basis was 75. The state wants to know if there is a significant difference between the local scores and the national scores. In such a situation the hypotheses may be stated as under

Null hypothesis H_0: μ = 80

Alternative Hypothesis Ha: μ 80

The formulation of hypotheses is an important step which must be accomplished with due care in accordance with the object and nature of the problem under consideration. It also indicates whether we should use a one-tailed test or a two-tailed test. If Ha is of the type greater than (or of the type lesser than), we use a one-tailed test, but when Ha is of the type "whether greater or smaller" then we use a two-tailed test.

Selecting a significance level: The hypotheses are tested on a pre-determined level of significance and as such the same should be specified. Generally, in practice, either 5% level or 1% level is adopted for the purpose. The factors that affect the level of significance are: (a) the magnitude of the difference between sample means; (b) the size of the samples; (c) the variability of measurements within samples; and (d) whether the hypothesis is directional or non-directional (A directional hypothesis is one which predicts the direction of the difference between, say, means). In brief, the level of significance must be adequate in the context of the purpose and nature of enquiry.

Deciding the distribution to use: After deciding the level of significance, the next step in hypothesis testing is to determine the appropriate sampling distribution. The choice generally remains between normal distribution and the t-distribution. The rules for selecting the correct distribution are similar to those which we have stated earlier in the context of estimation.

Selecting a random sample and computing an appropriate value: Another step is to select a random sample(s) and compute an appropriate value from the sample data concerning the test statistic utilizing the relevant distribution. In other words, draw a sample to furnish empirical data.

Calculation of the probability: One has then to calculate the probability that the sample result would diverge as widely as it has from expectations, if the null hypothesis were in fact true.

Comparing the probability: Yet another step consists in comparing the probability thus calculated with the specified value for á, the significance level. If the calculated probability is equal to or smaller than the á value in case of one-tailed test (and á /2 in case of two-tailed test), then reject the null hypothesis (i.e., accept the alternative hypothesis), but if the calculated probability is greater, then accept the null hypothesis. In case we reject

H0, we run a risk of (at most the level of significance) committing an error of Type I, but if we accept H0 , then we run some risk (the size of which cannot be specified as long as the H0 happens to be vague rather than specific) of committing an error of Type II.

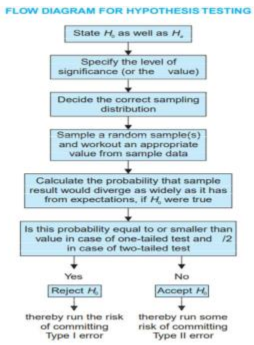


Figure 14.1: Flow Diagram of Hypothesis Testing

14.3.4 Measuring the Power of a Hypothesis Test

As stated above we may commit Type I and Type II errors while testing a hypothesis. The probability of Type I error is denoted as (the significance level of the test) and the probability of Type II error is referred to as . Usually the significance level of a test is assigned in advance and once we decide it, there is nothing else we can do about . But what can we say about ? We all know that hypothesis test cannot be foolproof; sometimes the test does not reject H_0 when it happens to be a false one and this way a Type II error is made. But we would certainly like that (the probability of accepting H_0 when H_0 is not true) to be as small as possible. Alternatively, we would like that 1 – (the probability of rejecting H_0 when H_0 is not true) to be as large as possible. If 1 – is very much nearer to unity (i.e., nearer to 1.0), we can infer that the test is working quite well, meaning thereby that the test is rejecting H0 when it is not true and if 1 – is very much nearer to 0.0, then we infer that the test is poorly working, meaning thereby that it is not rejecting H_0 when H_0 is not true. Accordingly 1 – value is the measure of how well the test is working or what is technically described as the power of the test. In case we plot the values of 1 – for each possible value of the population parameter (say µ, the true population mean) for which the H_0 is not true (alternatively the Ha is true), the resulting curve is known as the power curve associated with the given test. Thus power curve of a hypothesis test is the curve that shows the conditional probability of rejecting H_0 as a function of the population parameter and size of the sample.

The function defining this curve is known as the power function. In other words, the power function of a test is that function defined for all values of the parameter(s) which yields the probability that H_0 is rejected and the value of the power function at a specific parameter point is called the power of the test at that point. As the population parameter gets closer and closer to hypothesised value of the population parameter, the power of the test (i.e., 1

–) must get closer and closer to the probability of rejecting H_0 when the population parameter is exactly equal to hypothesised value of the parameter. We know that this probability is simply the significance level of the test, and as such the power curve of a test terminates at a point that lies at a height of (the significance level) directly over the population parameter.

Closely related to the power function, there is another function which is known as the operating characteristic function which shows the conditional probability of accepting H_0 for all values of population parameter(s) for a given sample size, whether or not the decision happens to be a correct one. If power function is represented as H and operating characteristic function as L, then we have L = 1 – H. However, one needs only one of these two functions for any decision rule in the context of testing hypotheses.

14.4 TESTSOFHYPOTHESES

Hypothesis testing determines the validity of the assumption (technically described as null hypothesis) with a view to choose between two conflicting hypotheses about the value of a population parameter. Hypothesis testing helps to decide on the basis of a sample data, whether a hypothesis about the population is likely to be true or false. Statisticians have developed several tests of hypotheses (also known as the tests of significance) for the purpose of testing of hypotheses which can be classified as: (a) Parametric tests or standard tests of hypotheses; and (b) Non-parametric tests or distribution-free test of hypotheses.

Parametric tests usually assume certain properties of the parent population from which we draw samples. Assumptions like observations come from a normal population, sample size is large, assumptions about the population parameters like mean, variance, etc., must hold good before parametric tests can be used. But there are situations when the researcher cannot or does not want to make such assumptions. In such situations we use statistical methods for testing hypotheses which are called non-parametric tests because such tests do not depend on any assumption about the parameters of the parent population. Besides, most non-parametric tests assume only nominal or ordinal data, whereas parametric tests require measurement equivalent to at least an interval scale. As a result, non-parametric tests need more observations than parametric tests to achieve the same size of Type I and Type II errors.

14.4.1 Parametric Tests

The important parametric tests are: (1) z-test; (2) t-test; and (3) F-test. All these tests are based on the assumption of normality i.e., the source of data is considered to be normally distributed. In some cases the population may not be normally distributed, yet the tests will be applicable on account of the fact that we mostly deal with samples and the sampling distributions closely approach normal distributions.

Z-test is based on the normal probability distribution and is used for judging the significance of several statistical measures, particularly the mean. The relevant test statistic, z, is worked out and compared with its probable value (to be read from table showing area under normal curve) at a specified level of significance for judging the significance of the measure concerned. This is a most frequently used test in research studies. This test is used even when binomial distribution or t-distribution is applicable on the presumption that such a distribution tends to approximate normal distribution as 'n' becomes larger. Z-test is generally used for comparing the mean of a sample to some hypothesised mean for the population in case of large sample, or when population variance is known. Z-test is also used for judging the significance of difference between means of two independent samples in case of large samples, or when population variance is known. Z-test is also used for comparing the sample proportion to a theoretical value of population proportion or for judging the difference in proportions of two independent samples when n happens to be large. Besides, this test may be used for judging the significance of median, mode, coefficient of correlation and several other measures.

T-test is based on t-distribution and is considered an appropriate test for judging the significance of a sample mean or for judging the significance of difference between the means of two samples in case of small sample(s) when population variance is not known (in which case we use variance of the sample as an estimate of the population variance). In case two samples are related, we use paired t-test (or what is known as difference test) for judging the significance of the mean of difference between the two related samples. It can also be used for judging the significance of the coefficients of simple and partial correlations. The relevant test statistic, t, is calculated from the sample data and then compared with its probable value based on t-distribution (to be read from the table that gives probable values of t for different levels of significance for different degrees of freedom) at a specified level of significance for concerning degrees of freedom for accepting or rejecting the null hypothesis. It may be noted that t-test applies only in case of small sample(s) when population variance is unknown.

F-test is based on F-distribution and is used to compare the variance of the two-independent samples. This test is also used in the context of analysis of variance (ANOVA) for judging the significance of more than two sample means at one and the same time. It is also used for judging the significance of multiple correlation coefficients. Test statistic, F, is calculated and compared with its probable value (to be seen in the F-ratio tables for different degrees of freedom for greater and smaller variances at specified level of significance) for accepting or rejecting the null hypothesis.

14.4.2 Non-Parametric Tests

In a statistical test, two kinds of assertions are involved viz., an assertion directly related to the purpose of investigation and other assertions to make a probability statement. The former is an assertion to be tested and is technically called a hypothesis, whereas the set of all other assertions is called the model. When we apply a test (to test the hypothesis) without a model, it is known as distribution-free test, or the nonparametric test. Non-parametric tests do not make an assumption about the parameters of the population and thus do not make use of the parameters of the distribution. In other words, under non-parametric or distribution-free tests we do not assume that a particular distribution is applicable, or that a certain value is attached to a parameter of the population. For instance, while testing the two training methods, say A and B, for determining the superiority of one over the other, if we do not assume that the scores of the trainees are normally distributed or that the mean score of all trainees taking method A would be a certain value, then the testing method is known as a distribution-free or nonparametric method. In fact, there is a growing use of such tests in situations when the normality assumption is open to doubt. As a result many distribution-free tests have been developed that do not depend on the shape of the distribution or deal with the parameters of the underlying population.

Tests of hypotheses with 'order statistics' or 'non-parametric statistics' or 'distribution-free' statistics are known as nonparametric or distribution-free tests. The following distribution-free tests are important and generally used:

Test of a hypothesis concerning some single value for the given data (such as one-sample sign test).

Test of a hypothesis concerning no difference among two or more sets of data (such as two-sample sign test, Fisher-Irwin test, Rank sum test, etc.).

Test of a hypothesis of a relationship between variables (such as Rank correlation, Kendall's coefficient of concordance and other tests for dependence.

Test of a hypothesis concerning variation in the given data i.e., test analogous to ANOVA viz., Kruskal-Wallis test.

Tests of randomness of a sample based on the theory of runs viz., one sample runs test.

Test of hypothesis to determine if categorical data shows dependency or if two classifications are independent viz., the chi-square test. The chi-square test can as well be used to make comparison between theoretical populations and actual data when categories are used.

The sign test is one of the easiest non-parametric tests. Its name comes from the fact that it is based on the direction of the plus or minus signs of observations in a sample and not on their numerical magnitudes. The sign test may be one of the following two types: The one sample sign test is a very simple non-parametric test applicable when we sample a continuous symmetrical population in which case the probability of getting a sample value less than mean is 1/2 and the probability of getting a sample value greater than mean is also 1/2. To test the null hypothesis $\mu = \mu_{H0}$ against an appropriate alternative on the basis of a random sample of size 'n', we replace the value of each and every item of the sample with a plus (+) sign if it is greater than $\mu H0$, and with a minus (–) sign if it is less than μ_{H0} . But if the value happens to be equal to μ_{H0} , then we simply discard it. After doing this, we test the null hypothesis that these + and – signs are values of a random variable, having a binomial distribution with p = 1/2*. For performing one sample sign test when the sample is small, we can use tables of binomial probabilities, but when sample happens to be large, we use normal approximation to binomial distribution.

The sign test has important applications in problems where we deal with paired data. In such problems, each pair of values can be replaced with a plus (+) sign if the first value of the first sample (say X) is greater than the first value of the second sample (say Y) and we take minus (–) sign if the first value of X is less than the first value of Y. In case the two values are equal, the concerning pair is discarded. (In case the two samples are not of equal size, then some of the values of the larger sample left over after the random pairing will have to be discarded.) The testing technique remains the same as started in case of one sample sign test.

Fisher-Irwin Test: Fisher-Irwin test is a distribution-free test used in testing a hypothesis concerning no difference among two sets of data. It is employed to determine whether one can reasonably assume, for example, that two supposedly different treatments are in fact different in terms of the results they produce. Suppose the management of a business unit has designed a new training programme which is now ready and as such it wishes to test its performance against that of the old training programme. This test (Fisher-Irwin test), is applicable for those situations where the observed result for each item in the sample can be classified into one of the two mutually exclusive categories. For instance, in the given example the worker's performance was classified as fail or pass and accordingly numbers failed and passed in each group were obtained. But supposing the score of each worker is also given and we only apply the Fisher-Irwin test as above, then certainly we are discarding the useful information concerning how well a worker scored. This in fact is the limitation of the Fisher-Irwin test which can be removed if we apply some other test, say, Wilcoxon test.

Wilcoxon Matched-pairs Test (or Signed Rank Test): In various research situations in the context of two-related samples (i.e., case of matched pairs such as a study where husband and wife are matched or when we compare the output of two similar machines or where some subjects are studied in context of before-after experiment) when we can determine both direction and magnitude of difference between matched values, we can use an important non-parametric test viz., Wilcoxon matched-pairs test. While applying this test, we first find the differences (di) between each pair of values and assign rank to the differences from the smallest to the largest without regard to sign. The actual signs of each difference are then put to corresponding ranks and the test statistic T is calculated which happens to be the smaller of the two sums viz., the sum of the negative ranks and the sum of the positive ranks.

Rank Sum Tests Rank sum tests are a whole family of test, but we shall describe only two such tests commonly used viz., the U test and the H test. U test is popularly known as Wilcoxon-Mann-Whitney test, whereas H test is also known as Kruskal-Wallis test. A brief description of the said two tests is given below:

4(a) Wilcoxon-Mann-Whitney test (or U-test): This is a very popular test amongst the rank sum tests. This test is used to determine whether two independent samples have been drawn from the same population. It uses more information than the sign test or the Fisher-Irwin test. This test applies under very general conditions and requires only that the populations sampled are continuous. However, in practice even the violation of this assumption does not affect the results very much. To perform this test, we first of all rank the data jointly, taking them as belonging to a single sample in either an increasing or decreasing order of magnitude. We usually adopt low to high ranking process which means we assign rank 1 to an item with lowest value, rank 2 to the next higher item and so on. In case there are ties, then we would assign each of the tied observation the mean of the ranks which they jointly occupy. In applying U-test we take the null hypothesis that the two samples come from identical populations. If this hypothesis is true, it seems reasonable to suppose that the means of the ranks assigned to the values of the two samples should be more or less the same. Under the alternative hypothesis, the means of the two populations are not equal and if this is so, then most of the smaller ranks will go to the values of one sample while most of the higher ranks will go to those of the other sample.

4(b) The Kruskal-Wallis test (or H test): This test is conducted in a way similar to the U test described above. This test is used to test the null hypothesis that 'k' independent random samples come from identical universes against the alternative hypothesis that the means of these universes are not equal. This test is analogous to the one-way analysis of variance, but unlike the latter it does not require the assumption that the samples come from approximately normal populations or the universes having the same standard deviation. In this test, like the U test, the data are ranked jointly from low to high or high to low as if they constituted a single sample. If the null hypothesis is true that there is no difference between the sample means and each sample has at least five items, then the sampling distribution of H can be approximated with a chisquare distribution with (k – 1) degrees of freedom. As such we can reject the null hypothesis at a given level of significance if H value calculated, as stated above, exceeds the concerned table value of chi-square.

14.5 LIMITATIONS OFTHE TESTS OFHYPOTHESES

We have described above some important test often used for testing hypotheses on the basis of which important decisions may be based. But there are several limitations of the said tests which should always be borne in mind by a researcher. Important limitations are as follows:

The tests should not be used in a mechanical fashion. It should be kept in view that testing is not decision-making itself; the tests are only useful aids for decision-making. Hence "proper interpretation of statistical evidence is important to intelligent decisions."

Test does not explain the reasons as to why does the difference exist, say between the means of the two samples. They simply indicate whether the difference is due to fluctuations of sampling or because of other reasons but the tests do not tell us as to which is/are the other reason(s) causing the difference.

Results of significance tests are based on probabilities and as such cannot be expressed with full certainty. When a test shows that a difference is statistically significant, then it simply suggests that the difference is probably not due to chance.

Statistical inferences based on the significance tests cannot be said to be entirely correct evidences concerning the truth of the hypotheses. This is specially so in case of small samples where the probability of drawing erring inferences happens to be generally higher. For greater reliability, the size of samples be sufficiently enlarged. All these limitations suggest that in problems of statistical significance, the inference techniques (or the tests) must be combined with adequate knowledge of the subject-matter along with the ability of good judgement.

14.6 SUMMARY

In the competitive and dynamic business world, those enterprises which are most likely to succeed, and indeed survive are those which are capable of maximising the use of various tools of testing hypothesis. This lesson has attempted to describe the meaning of hypothesis which helps the researcher in analysing the data in the field of business and management. A hypothesis test evaluates two mutually exclusive statements about a population to determine which statement is best supported by the sample data. A test result is statistically significant when the sample statistic is unusual enough relative to the null hypothesis that we can reject the null hypothesis for the entire population. Keep in mind that there is no magic significance level that distinguishes between the studies that have a true effect and those that don't with 100% accuracy. The common alpha values of 0.05 and 0.01 are simply based on tradition. For a significance level of 0.05, expect to obtain sample means in the critical region 5% of the time when the null hypothesis is true. In these cases, you won't know that the null hypothesis is true but you'll reject it because the sample mean falls in the critical region. That's why the significance level is also referred to as an error rate. When samples are small, i.e., less than 30 the large sample results do not hold good for small samples, i.e., the assumption of approximate normality of the distribution is not true; in fact another distribution (exact distribution) of the test statistics is to be used and the result modified accordingly. Generally, the student's t-test, Z-test, Chi-square test and F-test are exact tests or small test of interest. In the present lesson we students different parametric tests used for testing hypothesis, i.e., Z-test, t-test, and F-test. Student's t-test is used to test the significance of mean, significance of difference of two mean and significance of correlation coefficient are also discussed. Z-test is also used to test the significance of correlation coefficient.

14.7 GLOSSARY

Type I error: Probability of rejecting the null hypothesis when it should be accepted, that is, concluding that two means are significantly different when in fact they are the same. Denoted with á.

Type II error: Probability of failing to reject the null hypothesis when it should be rejected, that is, concluding that two means are not significantly different when in fact they are different. Denoted with â.

Null Hypothesis (H_0): Hypothesis which is tested for its possible rejection under the assumption is called as null hypothesis.

Alternative Hypothesis (H_1): Any hypothesis which is complementary to the null hypothesis.

Critical Region: Critical region is the region of values that corresponds to the rejection of the null hypothesis at some chosen probability level.

Level of Significance: The level of significance is defined as the probability of rejecting a null hypothesis by the test when it is really true, which is denoted as á.

Power of Hypothesis Test: A test's power is the probability of correctly rejecting the null hypothesis when it is false; a test's power is influenced by the choice of significance level for the test, the size of the effect being measured, and the amount of data available.

Standard Error: A measure of the statistical accuracy of an estimate, equal to the standard deviation of the theoretical distribution of a large population of such estimates.

Confidence interval: In statistics, a confidence interval (CI) is a type of interval estimate, computed from the statistics of the observed data that might contain the true value of an unknown population parameter. Most commonly, the 95% confidence level is used

T-test: The t test is one type of inferential statistics. It is used to determine whether there is a significant difference between the means of two groups.

Z-test: A z-test is a statistical test used to determine whether two population means are different when the variances are known and the sample size is large.

F-test: An F-test is any statistical test in which the test statistic has an F-distribution under the null hypothesis. It is most often used when comparing statistical models that have been fitted to a data set, in order to identify the model that best fits the population from which the data were sampled.

14.8 SELFASSESSMENT QUESTIONS

Fill in the blanks:

The null hypothesis asserts that there is no true difference in the.....and the .... in the particular matter under consideration.

Type I error is committed when the hypothesis is true but our test.....it.

Type II errors are made when we accept a null hypothesis which is.............

In ....tail test the rejection region is located in one tail.

______________is the value we receive when we run an ANOVA test on different groups.

The table value of Z-test at 5% level of significance is______________.

F-test is given by__________ and is also known as_____________________.

Z-test is use when our sample size is______________and population standard deviation is_____________.

Critical is also known as____________.

14.9 LESSON END EXERCISE

The procedure of testing hypothesis requires a researcher to adopt several steps. Describe in brief all such steps.

____

____

____

What is a hypothesis? What characteristics it must possess in order to be a good research hypothesis?

What do you mean by the power of a hypothesis test? How can it be measured? Describe and illustrate by an example.

Briefly describe the important parametric tests used in context of testing hypotheses. How such tests differ from non-parametric tests? Explain.

Distinguish between Simple hypothesis and composite hypothesis.

Distinguish between Simple hypothesis and composite hypothesis.

Distinguish between Null hypothesis and alternative hypothesis.

14.10 SUGGESTED READINGS

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

William G. Z. Business Research Methods. Thomson, India.

Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

UNIT-III **Lesson No. 15**

CONCEPT, IMPORTANCEAND PRE-REQUISITES OFDATA INTERPRETATION; ERRORS IN DATAINTERPRETATION

STRUCTURE

15.1 INTRODUCTION

Data analysis and interpretation have now taken centre stage with the advent of the digital age. In fact, a Digital Universe study found that the total data supply in 2012 was 2.8 trillion gigabytes. Based on that amount of data alone, it is clear the calling card of any successful enterprise in today's global world will be the ability to analyse complex data, produce actionable insights and adapt to new market needs. Business dashboards are the digital age tools for big data. Capable of displaying key performance indicators (KPIs) for both quantitative and qualitative data analyses, they are ideal for making the fast-paced and data-driven market decisions that push today's industry leaders to sustainable success. Through the art of streamlined visual communication, data dashboards permit businesses to engage in real-time and informed decision-making and are key instruments in data interpretation. Turning data into useful information a process that involves several steps:

Data Entry: This involves getting your raw data into a computer so that you can store it and retrieve it for analysis. It includes two steps: a. Entry: Data should be entered into a computer data management application. b. Validation: The entered data must be checked against the field and lab sheets to assure that it has been entered correctly.

Summary: The data is put into a form that allows you to view it as a whole, such as simple statistics, tables, and graphs.

Data Interpretation: This involves asking a series of questions about your data that relate to your study question(s). Your answers to these questions are organised as findings and conclusions. Based on these, you may develop recommendations for action or further study.

Presenting Your Results: Present your findings, conclusions, and recommendations in a form that best tells the story of your data. This story can be told in text and selected tables and graphs that are organised into an oral presentation and/or a written report. Your presentation or report should be geared to the audience you are trying to reach. This lesson focuses on the data interpretation step.

Therefore, based on the importance of data interpretation in research this lesson deals with various aspects of data interpretation, viz., concept, importance, factors considered etc.

15.2 OBJECTIVES

After reading this lesson, you will be able to:

Understand the concept and importance of data interpretation.

Explain different factors considered for data interpretation.

Know the errors/problems of data interpretation.

Understand the process of data interpretation.

15.3 DATA INTERPRETATION

Interpreting your data is a process that involves answering a series of questions about the data. The following steps should be followed when we start interpreting our data. 1) Review and interpret the data "in-house" to develop preliminary findings, conclusions and recommendations.

Review the data and your interpretation of it with an advisory group or technical committee. This group should involve local, regional, and state resource people who are familiar with monitoring and with our data/research. They can verify, add to, or correct your interpretation of the results.

Review the data and your interpretation of it with the people who will use your data - for example, the public, customers and government officials.

15.3.1 Concept

Data interpretation refers to the process of using diverse analytical methods to review data and arrive at relevant and meaningful conclusions. The interpretation of data helps researchers to categorise, manipulate, and summarise the information in order to answer critical questions.

Data is very likely to arrive from multiple sources and has a tendency to enter the analysis process with haphazard ordering. Data analysis tends to be extremely subjective. That is to say, the nature and goal of interpretation will vary from business to business, likely correlating to the type of data being analysed. While there are several different types of processes that are implemented based on individual data nature, the two broadest and most common categories are “quantitative analysis” and “qualitative analysis”. Before any serious data analysis can begin, the scale of measurement must be decided for the data as this will have a long-term impact on data interpretation. The varying scales include:

Nominal Scale: non-numeric categories that cannot be ranked or compared quantitatively. Variables are exclusive and exhaustive.

Ordinal Scale: exclusive categories that are exclusive and exhaustive but with a logical order. Quality ratings and agreement ratings are examples of ordinal scales (i.e., good, very good, fair, etc., OR agree, strongly agree, disagree, etc.).

Interval: a measurement scale where data is grouped into categories with orderly and equal distances between the categories. There is always an arbitrary zero point.

Ratio: contains features of all three.

Therefore, data interpretation is a data review process that utilises analysis, evaluation, and visualisation to provide in-depth findings to enhance data-driven decision-making. Further, there are many steps involved in data interpretation, as well as different types of data and data analysis processes that influence the larger data interpretation process.

15.3.2 Importance of Data Interpretation

The importance of data interpretation is not far from the importance of other data processes. Much like implementing data normalisation and understanding data quality, proper data interpretation offers real-time solutions and provides more in-depth insights than without it. More precisely, data interpretation can improve data identification, discover hidden correlations between datasets, find data outliers, and even help forecast trends.

Additionally, proper implementation of data interpretation offers immense benefits such as cost efficiency, enhanced decision making, and improved AI predictions. Namely, in a Business Intelligence survey, it was reported that companies that implemented data analysis and interpretation from big data datasets saw a ten percent reduction in costs. Thus, data analysis and interpretation helps to improve the processes and identify problems.

Informed decision-making: A decision is only as good as the knowledge that formed it. Informed data decision-making has the potential to set industry leaders apart from the rest of the market pack. Studies have shown that companies in the top third of their industries are, on average, 5% more productive and 6% more profitable when implementing informed data decision-making processes. Most decisive actions will arise only after a problem has been identified or a goal defined. Data analysis should include identification, thesis development, and data collection followed by data communication. The monitoring of data results will inevitably return the process to the start with new data and sights.

Anticipating needs with trends identification: Data insights provide knowledge, and knowledge is power. The insights obtained from market and consumer data analyses have the ability to set trends for peers within similar market segments. When industry trends are identified, they can then serve a greater industry purpose. Also, data gathering and interpretation processes can allow for industry-wide climate prediction and result in greater revenue streams across the market. Therefore, to achieve this benefit all institutions should follow the basic data cycle of collection, interpretation, decision making and monitoring.

Cost efficiency: Proper implementation of data analysis processes can provide businesses with profound cost advantages within their industries.

Clear Foresight: Companies that collect and analyse their data gain better knowledge about themselves, their processes, and performance. They can identify performance challenges when they arise and take action to overcome them. Data interpretation through visual representations processes findings faster and make better-informed decisions on the future of the company.

15.3.3 Pre-requisites of Data Interpretation

To interpret means to explain the meaning of data. Interpretation of data is that important part of the research which is associated with the drawing of inferences from the collected data. It is very useful and important part of research because it makes possible the use of research data. Figures themselves have no utility. It is the interpretation that makes it possible to utilise the collected data. If data properly interrelated, it gives right judgement and conclusions. Correct interpretation leads to correct forecasting too. Right interpretation helps in correct decisions and for that data must be handled carefully and following precautions must be taken to have good interpretation.

The Data Should be Homogenous

Homogenous means of the same or similar nature. It means the collection of data should be made on the same parameters or units, otherwise comparison will not be possible and drawing conclusions would be difficult. For example, if the data regarding buying capacity is collected then it must be in the terms of rupees only (and not in dollars). Homogeneity of data should also take into account the similarity of features in the two different areas. For example, the data regarding Mumbai and Nagpur. If the data is not homogenous, then the conclusions drawn can easily lead to wrong findings.

The Data Should be Adequate

Adequate means sufficient. Conclusions based on inadequate data cannot be right and proper, therefore the data must be full and complete. Incomplete data also makes it difficult to analyse the data properly. For example if data is collected only about consumers expenses without knowing their incomes it may lead to inappropriate conclusions.

The Data Should be Suitable

Suitability of data is very important right from drafting questions, collection of answers and interpretation. For example, in order to study the buying behaviour of car users, rich consumers should be consulted. Data regarding poor consumers is not suitable for this product and therefore may lease to false or wrong conclusions.

The Data Should be Scientifically Analysed

For drawing proper conclusions, the most appropriate and advanced statistical technique should be used to analyse the data. It will lead to correct and fast conclusions.

Data Analysis Involves Critical Thinking

Data analysis involves critical thinking. This is done only after collecting all the data and always focused on the research problem. The data are described and interpreted in detail leading to the ultimate conclusion. Tables, graphs and illustrations are used to present the data more clearly and economically.

15.3.4 Data Interpretation Methods

The interpretation of data is designed to help people make sense of numerical data that has been collected, analysed, and presented. Having a baseline method (or methods) for interpreting data will provide the analyst teams with a structure and consistent foundation. Indeed, if several departments have different approaches to interpret the same data while sharing the same goals, some mismatched objectives can result. Different methods will lead to duplicated efforts, inconsistent solutions, wasted energy, and inevitably-time and money.

Qualitative Data Interpretation

Qualitative data analysis can be summed up in one word – categorical. With qualitative analysis, data is not described through numerical values or patterns, but through the use of descriptive context (i.e., text). Typically, narrative data is gathered by employing a wide variety of person-to-person techniques. These techniques include:

Observations: detailing behavioural patterns that occur within an observation group. These patterns could be the amount of time spent in an activity, the type of activity, and the method of communication employed.

Focus groups: Group people and ask them relevant questions to generate a collaborative discussion about a research topic.

Secondary Research: much like how patterns of behaviour can be observed, different types of documentation resources can be coded and divided based on the type of material they contain.

Interviews: one of the best collection methods for narrative data. Inquiry responses can be grouped by theme, topic, or category. The interview approach allows for highly-focused data segmentation.

A key difference between qualitative and quantitative analysis is clearly noticeable in the interpretation stage. Qualitative data, as it is widely open to interpretation, must be "coded" so as to facilitate the grouping and labelling of data into identifiable themes. As person-to-person data collection techniques can often result in disputes pertaining to proper analysis, qualitative data analysis is often summarised through three basic principles: notice things, collect things, think about things.

Quantitative Data Interpretation

If quantitative data interpretation could be summed up in one word that word would be "numerical." There are few certainties when it comes to data analysis, but you can be sure that if the research you are engaging in has no numbers involved, it is not quantitative research. Quantitative analysis refers to a set of processes by which numerical data is analysed. It involves the use of statistical modelling such as standard deviation, mean and median.

Mean: a mean represents a numerical average for a set of responses. When dealing with a data set (or multiple data sets), a mean will represent a central value of a specific set of numbers. It is the sum of the values divided by the number of values within the data set. Other terms that can be used to describe the concept are arithmetic mean, average and mathematical expectation.

Standard deviation: this is another statistical term commonly appearing in quantitative analysis. Standard deviation reveals the distribution of the responses around the mean. It describes the degree of consistency within the responses; together with the mean, it provides insight into data sets.

Frequency distribution: this is a measurement gauging the rate of a response appearance within a data set. When using a survey, for example, frequency distribution has the capability of determining the number of times a specific ordinal scale response appears (i.e., agree, strongly agree, disagree, etc.). Frequency distribution is extremely keen in determining the degree of consensus among data points.

Typically, quantitative data is measured by visually presenting correlation tests between two or more variables of significance. Different processes can be used together or separately, and comparisons can be made to ultimately arrive at a conclusion. Other signature interpretation processes of quantitative data include; regression analysis, Cohort analysis, predictive analysis, prescriptive analysis, conjoint analysis, clusters analysis etc.

15.4 ERRORS IN DATA INTERPRETATION

The often repeated mantra of those who fear data advancements in the digital age is "big data equals big trouble." While that statement is not accurate, it is safe to say that certain data interpretation problems or "pitfalls" exist and can occur when analysing data, especially at the speed of thought. Some of the most common data misinterpretation risks are:

Correlation mistaken for causation: Our first misinterpretation of data refers to the tendency of data analysts to mix the cause of a phenomenon with correlation. It is the assumption that because two actions occurred together, one caused the other. This is not accurate as actions can occur together absent a cause and effect relationship. Remedy: attempt to eliminate the variable you believe to be causing the phenomenon.

Confirmation bias: Our second data interpretation problem/error occurs when you have a theory or hypothesis in mind but are intent on only discovering data patterns that provide support to it while rejecting those that do not.

Digital age example: your boss asks you to analyze the success of a recent multi-platform social media marketing campaign. While analysing the potential data variables from the campaign (one that you ran and believe performed well), you see that the share rate for Facebook posts was great, while the share rate for Twitter Tweets was not. Using only the Facebook posts to prove your hypothesis that the campaign was successful would be a perfect manifestation of confirmation bias.

Remedy: as this pitfall is often based on subjective desires, one remedy would be to analyze data with a team of objective individuals. If this is not possible, another solution is to resist the urge to make a conclusion before data exploration has been completed. Remember to always try to disprove a hypothesis, not prove it.

Irrelevant data: The third data misinterpretation pitfall is especially important in the digital age. As large data is no longer centrally stored, and as it continues to be analysed at the speed of thought, it is inevitable that analysts will focus on data that is irrelevant to the problem they are trying to correct.

Digital age example: In attempting to gauge the success of an email lead generation campaign, you notice that the number of homepage views directly resulting from the campaign increased, but the number of monthly newsletter subscribers did not. Based on the number of homepage views, you decide the campaign was a success when really it generated zero leads.

Truncating an Axes: When creating a graph to start interpreting the results of your analysis it is important to keep the axes truthful and avoid generating misleading visualisations. Starting the axes in a value that doesn't portray the actual truth about the data can lead to false conclusions.

Digital age example: In the image below we can see a graph from Fox News in which the Y-axes start at 34%, making it seem that the difference between 35% and 39.6% is way higher than it actually is. This could lead to a misinterpretation of the tax rate changes.

15.4.1 Process of Data Interpretation

Here in this section we can take a closer look at the interpretation process from which involve the five key steps starting from the baseline establishment, data collection, interpretation (qualitative or quantitative analysis), visualization, and upto reflection.

Baseline Establishment

Similar to the first step when conducting a competitive analysis, it is important to establish your baseline when conducting data interpretation. This can include setting objectives and outlining long-term and short-term goals that will be directly affected by any actions that result from your data interpretation. For example, investors utilising data interpretation may want to set goals regarding the ROI of companies they are evaluating. It is important to note that this step also includes the determination of which data type you wish to analyse and interpret.

Data Collection

Now that a baseline is established and the goals of your data interpretation process are known, you can start collecting data. As previously mentioned, the data collection process includes two major collecting types: web scraping and information exchange. Both of these methods are successful at collecting both qualitative and quantitative data. However, depending on the scope of your data interpretation process, you most likely will only require one method.

For example, if you are looking for specific information within a very particular demographic, you will want to target particular attributes within the larger demographic you are interested in. Particularly, let's say you want to collect sentiment surrounding an application used by a particular job type; you will want to target individuals with a specific job type attribute and utilise information exchange. Both of these collection methods can be quite extensive, and for that reason, you may want to enrich your data collection or even fully utilise high-quality data from a data provider. Notably, once your data is collected, you must clean and organise your data before you can proceed to analysis. This can be achieved through data cleaning and data normalisation processes.

Interpretation (qualitative or quantitative)

This step is arguably the most crucial one in the data interpretation process, and it involves the analysis of the data you've collected. This is where your decision to conduct a qualitative or quantitative analysis comes into play. Qualitative analysis will require you to use a more subjective lens. If you are using AI-based data analysis tools, extensive "coding" will be necessary so that the data can be understood subjectively as sentiment experienced by individuals that cannot be defined numerically. On the other hand, qualitative analysis requires that the data can be analysed through a numerical and mathematical approach. As mentioned earlier, raw numerical data will be analysed, resulting in mean, standard deviation, and ratios, which can then be analysed further via statistical modelling to better understand and predict behaviours.

Visualisation

When your analysis is complete, you can now start to visualise your data and draw insights from various perspectives. Today, many companies have implemented "dashboards" as a part of the visualisation stage. Dashboards essentially provide you with quick insights via programmable algorithms. Even without dashboards formatting your data for visualisation is relatively straightforward. To do this, you must input and format your data into a format that supports visualisation. Some of the more common visualisation formats include:

Bar charts
Tables
Scatter plots
Line graphs
Pie charts
Histogram

Reflection

At last, once you have created adequate visualisation types that meet your previously decided objectives, you can reflect. While a rather simple process, relative to the earlier steps, the reflection process can make or break your data interpretation process. During this step, you should reflect on the data analysis process as a whole, look for hidden correlations, and identify outliers or errors that may have affected your visualisation charts (but could have been missed during the data cleaning stage). It is crucial that during this step you differentiate between correlation and causation, identify bias, and take note of any missed insights.

15.5 SUMMARY

This lesson is designed to discuss about the concept of data interpretation, importance as well as steps and its pre-requisites. Data interpretation is an extremely important part of data-driven decision-making and should be done regularly as a part of a larger iterative interpretation process. The interpretation of data helps researchers to categorise, manipulate, and summarize the information in order to answer critical questions. There are four steps involves in the process of data interpretation: 1) assemble the information you'll need, 2) develop findings, 3) develop conclusions, and 4) develop recommendations.

15.6 GLOSSARY

Data Analysis: Data Analysis is the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data.

Data Interpretation: Data interpretation refers to the process of using diverse analytical methods to review data and arrive at relevant conclusions.

Qualitative analysis: Qualitative analysis uses subjective judgment to analyze a company's value or prospects based on non-quantifiable information, such as management expertise, industry cycles, strength of research and development, and labour relations.

Quantitative analysis: Quantitative analysis (QA) is a technique that uses mathematical and statistical modelling, measurement, and research to understand behaviour. Quantitative analysts represent a given reality in terms of a numerical value.

15.7 SELFASSESSMENT QUESTIONS

Multiple Choice Questions. Tick ( ) the correct answer:

Data Analytics uses ___ to get insights from data.

a. Statistical figures b. Numerical aspects
c. Statistical methods d. None of the mentioned above

Linear Regression is the supervised machine learning model in which the model finds the best fit ___ between the independent and dependent variable.

a. Linear line b. Nonlinear line
c. Curved line d. All of the mentioned above

Amongst which of the following is / are the types of Linear Regression,

a. Simple Linear Regression b. Multiple Linear Regressions
c. Both A and B d. None of the mentioned above

___ are used when we want to visually examine the relationship between two quantitative variables.

a. Bar graph b. Scatterplot
c. Line graph d. Pie chart

Data Analysis is a process of,

a. Inspecting data b. Data Cleaning
c. Transforming of data d. All of the mentioned above

Which of the following is not a major data analysis approaches?

a. Data Mining b. Predictive Intelligence
c. Business Intelligence d. Text Analytics

Descriptive Statistics means__________

a. Describes what data looks like b. Descriptive analysis
c. Both of above d. None of the above

15.8 LESSON END EXERCISE

Explain the concept and process of data interpretation in detail.

____
____
____

Differentiate between data analysis and data interpretation.

____
____
____

Discuss the pre-requisites of data interpretation.

____
____
____

Why data analysis and interpretation is important?

____
____
____

15.9 SUGGESTED READING

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

William G. Z. Business Research Methods. Thomson, India.

Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

UNIT-IV **Lesson No. 16**

CONCEPTAND TYPES OFRESEARCH REPORTS; ESSENTIALS OFAGOOD RESEARCH REPORT

STRUCTURE

16.1 INTRODUCTION

Research report is considered a major component of the research study for the research task remains incomplete till the report has been presented and/or written. As a matter of fact even the most brilliant hypothesis, highly well designed and conducted research study, and the most striking generalisations and findings are of little value unless they are effectively communicated to others. The purpose of research is not well served unless the findings are made known to others. There are people who do not consider writing of report as an integral part of the research process. But the general opinion is in favour of treating the presentation of research results or the writing of report as part and parcel of the research project. Writing of report is the last step in a research study and requires a set of skills somewhat different from those called for in respect of the earlier stages of research. This task should be accomplished by the researcher with utmost care; he may seek the assistance and guidance of experts for the purpose. One of the reasons for carrying out research is to add to the existing body of knowledge. Therefore, when conducting research, you need to document your processes and findings in a research report. With a research report, it is easy to outline the findings of your systematic investigation and any gaps needing further inquiry. Knowing how to create a detailed research report will prove useful when you need to conduct research. Here our aim is to write clearly and concisely about the research topic so that the reader can easily understand the purpose and results of the research. In this lesson we shall discuss the concept of research report, types of research reports, and essentials of a good research report.

16.2 OBJECTIVES

After reading this lesson, you will be able to:

Understand the concept and types of research reports.

Define the parts of a research report following a standard

format Know the characteristics of a good research report.

16.3 CONCEPT OFRESEARCH REPORT

A research report is a well-crafted document that outlines the processes, data, and findings of a systematic investigation. It is an important document that serves as a first-hand account of the research process, and it is typically considered an objective and accurate source of information. In other ways, a research report can be considered as a summary of the research process that clearly highlights the findings, implications, and other important details. Reading a well-written research report should provide you with all the information you need to know about the core areas of the research process. Some of the basic features that define a research report are:

It is a detailed presentation of research processes and findings, and it usually includes tables and graphs.

It is written in a formal language.

A research report is usually written in the third person.

It is informative and based on first-hand verifiable information.

It is formally structured with headings, sections, and bullet points.

It always includes recommendations for future actions.

16.4 TYPES OFRESEARCH REPORTS

The research report is classified based on two things, i.e. on the basis of nature of research and on the basis of target audience.

On the Basis of Nature of Research

Qualitative Research Report

This is the type of report written for qualitative research. It outlines the methods, processes, and findings of a qualitative method of systematic investigation. In educational research, a qualitative research report provides an opportunity for one to apply his or her knowledge and develop skills in planning and executing qualitative research projects. A qualitative research report is usually descriptive in nature. Hence, in addition to presenting details of the research process, you must also create a descriptive narrative of the information.

Quantitative Research Report

A quantitative research report is a type of research report that is written for quantitative research. Quantitative research is a type of systematic investigation that pays attention to numerical or statistical values in a bid to find answers to research questions. In this type of research report, the researcher presents quantitative data to support the research process and findings. Unlike a qualitative research report that is mainly descriptive, a quantitative research report works with numbers; that is, it is numerical in nature.

On the Basis of Target Audience

Also, a research report can be said to be technical or popular based on the target audience. If you're dealing with a general audience, you would need to present a popular research report, and if you're dealing with a specialised audience, you would submit a technical report.

Technical Research Report

A technical research report is a detailed document that you present after carrying out industry-based research. This report is highly specialized because it provides information for a technical audience; that is, individuals with above-average knowledge in the field of study. In a technical research report, the researcher is expected to provide specific information about the research process, including statistical analyses and sampling methods. Also, the use of language is highly specialized and filled with jargon. Examples of technical research reports include legal and medical research reports.

Popular Research Report

A popular research report is one for a general audience; that is, for individuals who do not necessarily have any knowledge in the field of study. A popular research report aims to make information accessible to everyone. It is written in very simple language, which makes it easy to understand the findings and recommendations. Examples of popular research reports are the information contained in newspapers and magazines.

16.4.1 Steps in Writing Reports

Research reports are the product of slow, painstaking, accurate inductive work. The usual steps involved in writing report are: (a) logical analysis of the subject-matter; (b) preparation of the final outline; (c) preparation of the rough draft; (d) rewriting and polishing; (c) preparation of the final bibliography; and (f) writing the final draft. Though all these steps are self explanatory, a brief mention of each one of these will be appropriate for better understanding.

Logical Analysis of the Subject Matter: It is the first step which is primarily concerned with the development of a subject. There are two ways in which to develop a subject (a) logically and (b) chronologically. The logical development is made on the basis of mental connections and associations between the one thing and another by means of analysis. Logical treatment often consists in developing the material from the simple possible to the most complex structures. Chronological development is based on a connection or sequence in time or occurrence. The directions for doing or making something usually follow the chronological order.

Preparation of the Final Outline: It is the next step in writing the research report "Outlines are the framework upon which long written works are constructed. They are an aid to the logical organisation of the material and a reminder of the points to be stressed in the report."

Preparation of the Rough Draft: This follows the logical analysis of the subject and the preparation of the final outline. Such a step is of utmost importance for the researcher now sits to write down what he has done in the context of his research study. He will write down the procedure adopted by him in collecting the material for his study along with various limitations faced by him, the technique of analysis adopted by him, the broad findings and generalisations and the various suggestions he wants to offer regarding the problem concerned.

Rewriting and Polishing of the Rough Draft: This step happens to be most difficult part of all formal writing. Usually this step requires more time than the writing of the rough draft. The careful revision makes the difference between a mediocre and a good piece of writing. While rewriting and polishing, one should check the report for weaknesses in logical development or presentation. The researcher should also "see whether or not the material, as it is presented, has unity and cohesion; does the report stand upright and firm and exhibit a definite pattern, like a marble arch? Or does it resemble an old wall of mouldering cement and loose brick." In addition the researcher should give due attention to the fact that in his rough draft he has been consistent or not. He should check the mechanics of writing-grammar, spelling and usage.

Preparation of the Final Bibliography: Next in order comes the task of the preparation of the final bibliography. The bibliography, which is generally appended to the research report, is a list of books in some way pertinent to the research which has been done. It should contain all those works which the researcher has consulted. The bibliography should be arranged alphabetically and may be divided into two parts; the first part may contain the names of books and pamphlets, and the second part may contain the names of magazine and newspaper articles. Generally, this pattern of bibliography is considered convenient and satisfactory from the point of view of reader, though it is not the only way of presenting bibliography. The entries in bibliography should be made adopting the following order:

For books and pamphlets the order may be as under:

Name of author, last name first.

Title, underlined to indicate italics.

Place, publisher, and date of publication.

Number of volumes

Example Kothari, C.R., Quantitative Techniques, New Delhi, Vikas Publishing House Pvt. Ltd., 1978.

For magazines and newspapers the order may be as under:

Name of the author, last name first.

Title of article, in quotation marks.

Name of periodical, underlined to indicate italics.

The volume or volume and number.

The date of the issue.

The pagination.

Example Robert V. Roosa, “Coping with Short-term International Money Flows”, The Banker, London, September, 1971, p. 995.

The above examples are just the samples for bibliography entries and may be used, but one should also remember that they are not the only acceptable forms. The only thing important is that, whatever method one selects, it must remain consistent

Writing the Final Draft: This constitutes the last step. The final draft should be written in a concise and objective style and in simple language, avoiding vague expressions such as "it seems", "there may be", and the like ones. While writing the final draft, the researcher must avoid abstract terminology and technical jargon. Illustrations and examples based on common experiences must be incorporated in the final draft as they happen to be most effective in communicating the research findings to others. A research report should not be dull, but must enthuse people and maintain interest and must show originality. It must be remembered that every report should be an attempt to solve some intellectual problem and must contribute to the solution of a problem and must add to the knowledge of both the researcher and the reader.

16.4.2 Importance of a Research Report

One of the reasons for carrying out research is to contribute to the existing body of knowledge, and this is made possible with a research report. A research report serves as a means to effectively communicate the findings of a systematic investigation to all.

With a research report, we would be able to identify knowledge gaps for further inquiry. A research report shows what has been done while hinting at other areas needing systematic investigation.

In market research, a research report would help us to understand the market needs and peculiarities at a glance.

A research report allows us to present information in a precise and concise manner.

It is time-efficient and practical because, in a research report, we do not have to spend time detailing the findings of our research work in person. We can easily send out the report via email and have stakeholders look at it.

16.5 ESSENTIALS OF A GOOD RESEARCH REPORT

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

Selectiveness: It is important to exclude the matter, which is known to all. Only necessary contents should be included to save time, costs, and energy. However, care should be taken that the vital points should not be missed.

Comprehensiveness: Report must be complete. It must include all the necessary contents. In short, it must contain enough detail to covey meaning.

Cost Consideration: It must be prepared within the budgeted amount. It should not result into excessive costs.

Accuracy: As far as possible, research report must be prepared carefully. It must be free from spelling mistakes and grammatical errors.

Objectivity: Report must be free from personal bias, i.e., it must be free from one's personal liking and disliking. The report must be prepared for impersonal needs. The facts must be stated boldly. It must reveal the bitter truth. It must suit the objectives and must meet expectations of the relevant audience/readers.

Clarity: Report must reveal the facts clearly. Contents and conclusions drawn must be free from ambiguities. In short, outcomes must convey clear-cut implications.

Preciseness: Research report must not be unnecessarily lengthy. It must contain only necessary parts with adequate description.

Simplicity: Report must be simple to understand. Unnecessary technical words or terminologies (jargons) should be avoided.

Proper Language: Researcher must use a suitable language. Language should be selected as per its target users.

Reliability: Research report must be reliable. Manager can trust on it. He can be convinced to decide on the basis of research reposts.

Proper Format: An ideal repost is one, which must be prepared as per commonly used format. One must comply with the contemporary practices; completely a new format should not be used.

Attractive: Report must be attractive in all the important regards like size, colour, paper quality, etc. Similarly, it should use liberally the charts, diagrams, figures, illustrations, pictures, and multiple colours.

16.6 SUMMARY

A research report is an oral presentation and/or written statement whose purpose is to communicate research results, strategic recommendations, and/or other conclusions to management or other specific audiences. Although this chapter deals primarily with the final written report required by an extensive research project, remember that the final report is not likely to be the only kind prepared. For a small project, a short oral or written report on the results may be all that is needed. Extensive projects may involve many written documents, interim reports, a long final written report, and several oral presentations. In addition, technical materials may be posted on an organisation's intranet.

16.7 GLOSSARY

Research Proposal: A research proposal is a document proposing a research project, generally in the sciences or academia, and generally constitutes a request for sponsorship of that research.

Research Report: A research report is a well-crafted document that outlines the processes, data, and findings of a systematic investigation.

Abstract: A summary of the contents of a book, article, or speech.

Oral Research report: Research can be report in two ways; written and oral. This is a technique to report research in oral (seminar, workshop, conference etc.)

16.8 SELFASSESSMENT QUESTIONS

Multiple Choice Questions:

What is a legal report?

A report which deals specifically with a legal problem
It's a brief statement of the most significant facts necessary for the investigation
Both a and b
Only a

Guidelines for report writing includes:

Knowledge of the research material
Organizing of research material
Rough drafts
Bibliography, footnotes and head notes
All of the above

Bibliography is :

At the end of the study arranged in alphabetical order
Anywhere in the study
Beginning of the study
Not necessary to be included

Which of the following become the initial part of a report:

Preface
Foreword
Table of cases
List of abbreviations
All of the above

The conceptual framework and background of research which will become the source for formulation of the hypothesis is known as:

Methodology
Review of literature
Data analysis
None of the above

Appendix to the report includes:

Questionnaires
Sample information
Mathematical derivations
All of the above

16.9 LESSON END EXERCISE

Explain the concept and importance of research report.

Distinguish between research report and research proposal.

Explain the types of research resports.

Discuss the principles that need to be followed while writing a research report.

Discuss the steps or the process of writing a research report.

16.10 SUGGESTED READING

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

William G. Z. Business Research Methods. Thomson, India.

Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

RESEARCH REPORT WRITING

UNIT-IV **Lesson No. 17**

STRUCTURE OFTHE RESEARCH REPORT: PRELIMINARYSECTION, MAIN REPORT, MANAGERIALIMPLICATIONS, REFERENCES, BIBLIOGRAPHYANDANNEXURE

STRUCTURE

17.1 INTRODUCTION

Anybody, who is reading the research report, must necessarily be conveyed enough about the study so that he can place it in its general scientific context, judge the adequacy of its methods and thus form an opinion of how seriously the findings are to be taken. For this purpose there is the need of proper layout of the report. The layout of the report means as to what the research report should contain. A comprehensive layout of the research report should comprise (A) preliminary pages; (B) the main text; and (C) the end matter.

17.2 OBJECTIVES

After reading this lesson, you will be able to:

Explain the structure of the research report.

Discuss about the preliminary section of the research report.

Explain the contents of the main body of the report.

Explain the relevance or significance of managerial implications in research.

17.3 STRUCTURE OF THE RESEARCH REPORT

Although every research report is custom-made for the project it represents, some conventions of report format are universal. They represent a consensus about the parts necessary for a good research report and how they should be ordered. This consensus is not a law, however. Every book on report writing suggests the use of its own unique format, and every report writer has to pick and choose the sections and order that will work best for the project at hand. Many companies and universities also have in-house report formats or writing guides for writers to follow. The structure of a research report may need to be adjusted for two reasons: (1) to obtain the proper level of formality and

to decrease the complexity of the report.

17.3.1 Preliminary Section

In its preliminary sections the report should carry a title and date, followed by acknowledgements in the form of 'Preface' or 'Foreword'. Then there should be a table of contents followed by list of tables and illustrations so that the decision-maker or anybody interested in reading the report can easily locate the required information in the report.

Title Page: The title page should state the title of the report, for whom the report was prepared, by whom it was prepared, and the date of release or presentation. The title should give a brief but complete indication of the purpose of the research project. Addresses and titles of the preparer and recipient may also be included. On confidential reports, the title page may list the people to whom the report should be circulated. For the most formal reports, the title page is preceded by a title fly page, which contains only the report's title.

Letter of Transmittal: Relatively formal and very formal reports include a letter of transmittal. Its purpose is to release or deliver the report to the recipient. It also serves to establish some rapport between the reader and the writer. This is the one part of the formal report in which a personal or even slightly informal tone should be used. The transmittal should not dive into the report findings except in the broadest terms.

Letter of Authorisation: The letter of authorisation is a letter to the researcher that approves the project, details who has responsibility for it, and describes the resources available to support it. Because the researcher would not write this letter personally, writing guidelines will not be discussed here. In many situations, simply referring to the authorization in the letter of transmittal is sufficient. If so, the letter of authorisation need not be included in the report. In some cases, though, the reader may be unfamiliar with the authorization or may need detailed information about it. In such cases, the report should include this letter, preferably an exact copy of the original.

Table of Contents: A table of contents is essential to any report more than a few pages long. It should list the divisions and subdivisions of the report with page references. The table of contents is based on the final outline of the report, but it should include only the first-level subdivisions. For short reports it is sufficient to include only the main divisions. If the report includes many figures or tables, a list of these should immediately follow the table of contents.

Summary: The summary, also known as the executive summary, briefly explains why the research project was conducted, what aspects of the problem were considered, what the outcome was, and what should be done. It is a vital part of the report. Studies have indicated that nearly all managers read a report's summary, while only a minority read the rest of the report. Thus, the writer's only chance to produce an impact may be in the summary. The summary should be written only after the rest of the report has been completed. It is like the "greatest hits" of the report. It should be just one or two pages long, so the writer must carefully sort out what is important enough to be included in it. Several pages of the full report may have to be condensed into one summarizing sentence. Some parts of the report may be condensed more than others; the number of words in the summary need not be in proportion to the length of the section being discussed. The summary should be written to be self-sufficient. In fact, the summary is often detached from the report and circulated by itself. The summary contains four elements. First, it states the objectives of the report, including the most important background information and the specific purposes of the project. Second, it presents the methodology and the major results. Finally, the conclusions of the report are presented. These are opinions based on the results and constitute an interpretation of the results. Finally come recommendations, or suggestions for action, based on the conclusions. In many cases, managers prefer not to have recommendations included in the report or summary. Whether or not recommendations are to be included should be clear from the particular context of the report. An additional element that can be included in the summary is a short justification for the research study and report itself.

17.3.2 Main Report

The main text provides the complete outline of the research report along with all details. Title of the research study is repeated at the top of the first page of the main text and then follows the other details on pages numbered consecutively, beginning with the second page. Each main section of the report should begin on a new page. The main text of the report should have the following sections: (i) Introduction; (ii) Statement of findings and recommendations; (iii) The results; (iv) The implications drawn from the results; and (v) The summary

Introduction: The purpose of introduction is to introduce the research project to the readers. It should contain a clear statement of the objectives of research i.e., enough background should be given to make clear to the reader why the problem was considered worth investigating. A brief summary of other relevant research may also be stated so that the present study can be seen in that context. The hypotheses of study, if any, and the definitions of the major concepts employed in the study should be explicitly stated in the introduction of the report. The methodology adopted in conducting the study must be fully explained. The scientific reader would like to know in detail about such thing: How was the study carried out? What was its basic design? If the study was an experimental one, then what were the experimental manipulations? If the data were collected by means of questionnaires or interviews, then exactly what questions were asked (The questionnaire or interview schedule is usually given in an appendix)? If measurements were based on observation, then what instructions were given to the observers? Regarding the sample used in the study the reader should be told: Who were the subjects? How many were there? How were they selected? All these questions are crucial for estimating the probable limits of generalisability of the findings. The statistical analysis adopted must also be clearly stated. In addition to all this, the scope of the study should be stated and the boundary lines be demarcated. The various limitations, under which the research project was completed, must also be narrated.

Statement of findings and recommendations: After introduction, the research report must contain a statement of findings and recommendations in non-technical language so that it can be easily understood by all concerned. If the findings happen to be extensive, at this point they should be put in the summarised form.

Results: A detailed presentation of the findings of the study, with supporting data in the form of tables and charts together with a validation of results, is the next step in writing the main text of the report. This generally comprises the main body of the report, extending over several chapters. The result section of the report should contain statistical summaries and reductions of the data rather than the raw data. All the results should be presented in logical sequence and splitted into readily identifiable sections. All relevant results must find a place in the report. But how one is to decide about what is relevant is the basic question. Quite often guidance comes primarily from the research problem and from the hypotheses, if any, with which the study was concerned. But ultimately the researcher must rely on his own judgement in deciding the outline of his report. "Nevertheless, it is still necessary that he states clearly the problem with which he was concerned, the procedure by which he worked on the problem, the conclusions at which he arrived, and the bases for his conclusions."

Implications of the results: Toward the end of the main text, the researcher should again put down the results of his research clearly and precisely. He should, state the implications that flow from the results of the study, for the general reader is interested in the implications for understanding the human behaviour. Such implications may have three aspects as stated below: (a) A statement of the inferences drawn from the present study which may be expected to apply in similar circumstances. (b) The conditions of the present study which may limit the extent of legitimate generalisations of the inferences drawn from the study. (c) The relevant questions that still remain unanswered or new questions raised by the study along with suggestions for the kind of research that would provide answers for them. It is considered a good practice to finish the report with a short conclusion which summarises and recapitulates the main points of the study. The conclusion drawn from the study should be clearly related to the hypotheses that were stated in the introductory section. At the same time, a forecast of the probable future of the subject and an indication of the kind of research which needs to be done in that particular field is useful and desirable.

Summary: It has become customary to conclude the research report with a very brief summary, resting in brief the research problem, the methodology, the major findings and the major conclusions drawn from the research results

17.3.3 Managerial Implications

The managerial implications in research summarize what the results mean in terms of actions. In other words, the management effects compare the results with the action pattern and indicate which action, or even no action, should be taken in response. The effects of management should only be on the data provided in the report and not on how the results are implemented. If you are explicitly asked to do so, you may want to add sections with clear labels that separate your ideas from interpreting documented research for the classroom, the effects of management should flow directly from the objectives, summarising the answers to the questions asked and the evidence supporting the answers.

Managerial Implications should focus only on the evidence provided in the report rather than on how to implement the results. If explicitly asked to do so, you may add sections-clearly labeled-that separate your idea generation from the evidence-based interpretation of the research results. In terms of order, the Managerial Implications should flow directly from the Objectives, summarizing the answers to the questions asked and the evidence in support of the answers. The Managerial Implications may identify additional information needed for greater certainty. It should also remind the reader of limitations or caveats about the using the results, such as amount, quality or timeliness of the data.

17.4 END MATTER

At the end of the report, appendices should be enlisted in respect of all technical data such as questionnaires, sample information, mathematical derivations and the like ones. Bibliography of sources consulted should also be given. Index (an alphabetical listing of names, places and topics along with the numbers of the pages in a book or report on which they are mentioned or discussed) should invariably be given at the end of the report. The value of index lies in the fact that it works as a guide to the reader for the contents in the report

17.4.1 References

A references page is the last page of an essay or research paper that's been written in APA style. It lists all the sources you've used in your project so readers can easily find what you've cited. References can be described as giving credit, with citation, to the source of information used in one's work. Research is a build up on what other people have previously done thus referencing helps to relate your own work to previous work. References are a way to provide evidence to support the assertions and claims in your own assignments. References are also a way to give credit to the writers from whom you have borrowed words and ideas. References should always be accurate, allowing your readers to trace the sources of information you have used. The best way to make sure you reference accurately is to keep a record of all the sources you used when reading and researching for an assignment.

Referencing is important for a number of reasons, some of which include:

It allows for acknowledgement of the use of other people's opinions, ideas, theories and inventions.

Helps readers understand what influenced the writer's thinking and how their ideas were formulated.

Helps the readers evaluate the extent of the writer's reading.

Enables readers to visit source materials for them and verify the information.

What kind of information do I need to reference Printed books are not the only sources that require acknowledgement. Any words, ideas or information taken from any source requires a reference. Reference when you are using words or ideas from:

Books and journal articles newspapers and

magazines pamphlets or brochures

films, documentaries, television programs or advertisements

websites or electronic resources letters, emails, online discussion forums

Personal interviews lecturers or tutors. (Not always necessary but check with your lecturer or tutor about their preferences before you draw on their ideas.) You also need to reference when you reprint any diagrams, illustrations, charts or pictures.

No need to reference when you are, writing your own observations or experiment results, for example, a report on a field trip, writing about your own experiences, for example, a reflective journal writing your own thoughts, comments or conclusions in an assignment evaluating or offering your own analysis, using 'common knowledge' (facts that can be found in numerous places and are likely to be known by a lot of people) or folklore using generally accepted facts or information. This will vary in different disciplines of study.

Reference styles are standardised rules for presenting information about the sources used in a text. Typically, a style will describe how to organise information about author(s), publication year, title and page numbers. There are many different ways to organise the references of a text. Some reference styles follow the author-year format, while others are based on footnotes and/or numerical references. Some of the most commonly used reference styles in academic writing are:

American Psychological Association (APA6th) – author-year, commonly used in psychology, economics, educational sciences and health sciences (see the free guide at Purdue University)

The Chicago styles : Chicago 16 A – footnote system Chicago 16 B – author-year, commonly used in the humanities Harvard – author-year, a general reference system used in many disciplines (and used here at Sok & Skriv). See for example Quote, Unquote.

Modern Language Association (MLA) – author -page number, widely used in linguistics and literature (see, e.g., the description at Cornell University). Vancouver – numbered system, commonly used in medicine, health sciences and natural sciences.

IEEE (Institute of Electrical and Electronics Engineers) – used in, e.g., engineering and computer science (cf. the Citation Compass). Different academic journals use different reference styles.

17.4.2 Bibliography

A bibliography is a list of the entire sources researcher have used (whether referenced or not) in the process of researching research work. In general, a bibliography should include:

the authors' names

the titles of the works

the names and locations of the companies that published your copies of the

sources the dates your copies were published

the page numbers of your sources (if they are part of multi-source

volumes) Bibliography can be defined as,-

A list of reference materials (involving any kind of content; text, music, paintings, video etc.) elucidating the type, nature and other detailed information on the basis of name, date, place and genre of the materials.

A complete categorical compilation of any type of content based on its creator(s), editors and time (of production, distribution).

Bibliography, also known as works cited, reference list is basically an orderly study and referencing of books and source materials used in academic research. It might or might not include any information on the literary analysis or criticism of the materials cited.

There are many types of bibliographies and the leading bibliographies defer slightly in the names they assign to its various branches. There are primarily:

Annotated Bibliography: This provides a brief description or annotation of the cited sources. The annotation comprises of a brief summary of content along with a short analysis or evaluation.

Current Bibliography: It provides a list of published material and sources which are recently published or currently recorded material. The purpose of a current bibliography is to report recent literature as soon as it is published.

Retrospective Bibliography: It provides a lists of documents or parts of documents (articles) published in previous years, as distinct from a current bibliography.

Serial Bibliography: It is published over a period of known and pre-defined time slots. Time intervals for serial bibliographies normally range from weekly to annual basis and informs on the updates of book and research article titles.

National Bibliography: This provides a list of documents and sources published in a particular country and are produced in the national or local language of a country

International Bibliography: This provides a list of works, sources, publications, manuals, books, notes, articles and websites collected from worldwide sources.

Subject Bibliography: This provides a list of works and sources relates to a specific subject.

Period Bibliography: It provides a list of works and publications produced within a specific period range.

Analytical Bibliography: It refer to the collection of sources and material for the purpose of critical study and evaluation.

Descriptive (Physical) Bibliography: This provides a list of detailed facts for a book analysis by listing its size, format, binding, and publication details.

Historical Bibliography: It provides a list of contextual factors related to the production of a book i.e. printing details, publishing, bookselling and binding etc.

Textual Bibliography: It provides a list of literary materials, concerned with identification and editing of transcription errors from manuals, manuscripts, transcripts, scribes and inscriptions.

Enumerative (Systematic) Bibliography: It provides a list of the list of books according to some system, common theme or reference plan and includes information on by author, by subject, or by date. Contrary to a descriptive style, an enumerative one only provides minute details on books and sources.

A mandatory requirement of copyright laws and academic conventions is that whenever a research paper is written, there should be a section at the end of it where you acknowledge the sources used. So, bibliography means listing all the sources which you have consulted while writing your essay or research article. The sources may be in the form of printed and online books, websites, web documents, web blogs, newspaper articles, journals, pod casts, wikis, unpublished material, maps etc. Citation ensures that the information contained in the research paper is based on logic, truth and facts. Absence of references or bibliography indicates that the paper may be a piece of plagiarism.

There are various formats used in the creation of bibliographies such as the American Psychological Association (APA), Modern Language Association of America (MLA) and Chicago Manual of Style and Council of Biology Editors (CBE). The APA style of referencing is common in the papers written on topics of social sciences; MLA style is used in field of humanities; and CBE is a popular citation style in the natural sciences.

APA style referred by American Psychological Association uses both in-text citations and a list of references to document the sources. This style is used in social sciences.

For Example, American Psychological Association. (2010). Publication Manual of the American Psychological Association (6th ed.). Washington, DC: Author. Follow this format while citing book with one author using APA style:

Author: Include author's last name and only the initials of the first name.

Year: Year of publication will be included in parenthesis.

Title: Title should be in italics and only the first word of title should be capitalized or any proper noun.

Place of Publication: Mention city name followed by country name separated by a comma.

Publisher: include the name of publisher. Author's Last name, Initials. (Year) . Title (italicized).

17.4.3 Annexure

An annexure is a set of legal documents, which are added at the end of the report or book, to validate the information provided in the main text. An appendix refers to an extension to the research paper that contains information which is too detailed to put in the main document or report. Annexure or Attachments are a valuable way of removing detail from a report. Annexure provide a place to those items which do not fit in the report being too much detailed or specialize for example, sample design; formula for sample size, detailed statistical tables, maps, diagrams etc. Thus, anything which can help the reader to understand the contents of the report can be included in the annexure.

This gives supplementary information, sometimes of a fairly extensive type. A detailed description of the sample design, questionnaire and the instructions given to interviewers are mentioned. Full statistical tables are given for reference. Where survey information has been drawn from official or other sources, full details of these are listed. In the case of motivation research it is customary to include one or two typical interviews verbatim. Photographs or other illustrative material might also be included featuring. Photographs of packaging or advertising which was of interest to the survey. The types of materials which are to be placed under annexure are:

The work of any past research which is used in the current report as a module;

The format of the questionnaire;

Original data;

Intermediate tables of any algorithm which are used in the report;

Some micro level charts and figures;

Long tables;

Organizational details.

The appendices are used as an area for storing information which is important to the arguments raised in the research but, because of its length, detail or complexity would otherwise interrupt the flow of arguments in the research. We have a chance to make our report more valuable to the reader and to convince even reviewers critical of our argument of the usefulness of our publication. Thus, tables, transcripts of interviews, chronologies, glossaries, diagrams, biographies, lists of "Further Reading." etc., most probably, be appreciated.

Requirement for All Attachments

Lettering or Numbering: They must be allocated letters, Annex A, B, C, etc. Alternatively they may be given Arabic numbers. Letters and Roman numerals have the advantage that they don not get muddled with paragraph numbers. Care should be taken with Roman numbering in reports with numerous attachments, as many junior clerical staff are not familiar with many Roman numbers.

Mentioning: They must be mentioned in the report itself. They are supplementary to the main core of the report. They may, perhaps, deal with specialist or esoteric matters. Thus as they sit on the fringe of affairs, if the reader's attention is not drawn to them he may never see them. A note in brackets may be used.

Listing: They must be listed in or immediately after the table of contents. If there is no list of contents they should be iternized at the foot of the report.

After Note: In exceptional cases information may have to be added as it arises.

17.5 SUMMARY

In this unit, our main focus was on how to write a research report. At the outset, we defined what a research report enumerates the significance of a research report. Mostly, research work is presented in a written form. The practical utility of research study depends heavily on the way it is presented to those who are expected to act on the basis of research findings. Research report is a written document containing key aspects of research project. Research report is a medium to communicate research work with relevant people. It is also a good source of preservation of research work for the future reference. Many times, research findings are not followed because of improper presentation. Preparation of research report is not an easy task. It is an art. It requires a good deal of knowledge, imagination, experience, and expertise. It demands a considerable time and money. A research report is a document prepared by an analyst or strategist who is a part of the investment research team in a stock brokerage or investment bank. A research report may focus on a specific stock or industry sector, a currency, commodity or fixed-income instrument, or on a geographic region or country. Research reports generally, but not always, have actionable recommendations such as investment ideas that investors can act upon. The characteristics are: 1. Simplicity 2. Clarity 3. Brevity 4. Positivity 5. Punctuation 6. Approach 7. Readability

Accuracy 9. Logical Sequence 10. Proper Form 11. Presentation. Research report can vary differently in its length, type and purpose. Kerlinger (2004) states that the results of a research investigation can be presented in number of ways via a technical report, a popular report, a monograph or at times even in the form of oral presentation.

17.6 GLOSSARY

Annexure: An annexure is an addition to something, often to a document. When used generally to simply mean something added, annexure is interchangeable with annex.

Appendix: An appendix contains supplementary material that is not an essential part of the text itself but which may be helpful in providing a more comprehensive understanding of the research problem or it is information that is too cumbersome to be included in the body of the paper.

Bibliography: A bibliography is a list of all of the sources you have used (whether referenced or not) in the process of researching your work.

References: References provide the information necessary for readers to identify and retrieve each work cited in the text.

17.7 SELFASSESSMENT QUESTIONS

Multiple Choice Questions:

Why do you need to review the existing literature?

To make sure you have a long list of references
Because without it, you could never reach the required word-count
To find out what is already known about your area of interest
To help in your general studying

What do you mean by Unit of Analysis?

a. Main parameter b. Variables
c. Sample d. Constructs

A formal document that presents the research objectives, design of achieving these objectives, and the expected outcomes/deliverables of the study is called:

a. Research design b. Research proposal
c. Research hypothesis d. Research report

Final stage in the Research Process is :

a. Problem formulation b. Data collection
c. Data Analysis d. Report Writing

A comprehensive full Report of the research process is called:

a. Thesis b. Summary Report
c. Abstract d. Article

17.8 LESSON END EXERCISE

Would you begin your research report by thoroughly reviewing the literature?

____

____

____

Why do you want to write a report?

____

____

____

What basic format reference book would be appropriate for writing a report?

____

____

____

How much detail do you think you should discuss about methods of study?

____

____

____

Would you consider writing a research report as an on-going task throughout the research study?

____

____

____

Discuss the difference between a bibliography and a reference.

____

____

____

17.9 SUGGESTED READING

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw

Hill, New Delhi.
William G. Z. Business Research Methods. Thomson, India.
Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

UNIT-IV **Lesson No. 18**

EFFECTIVENESS OFRESEARCH REPORT

STRUCTURE

18.1 INTRODUCTION

Just as a research exercise adopts a scientific process, there are scientific ways of evaluating it. Unless research is scientifically evaluated, the implications cannot be assessed properly. Also, the researcher's hard work will remain unacknowledged. The purpose of research can be classified into three categories. These are:

Training in research,

Research for problem solving, and

Research in inquiry of truth or creation of knowledge.

There are definite patterns here. For example, most research projects, especially, projects at the post graduate and doctoral levels are mainly on training in research. The projects that are linked to MA(DE) courses of IGNOU are also intended to train a professional in research in open and distance education. Most of the action research and institutional projects aim at problem solving; the magnitude of the problem can vary from a classroom to the entire educational system. Research that generates knowledge or information is usually characterized by sustained work in a field for years and decades by one person or a group of senior professionals. It should not be difficult to appreciate that the paradigm for the evaluation of research reports, for three different goals cannot be the same. Though research methodology has to be a common interest in all evaluation, it will be the main focus in the evaluation of a report where training is the goal. When it comes to research for knowledge creation, the emphasis on methodology becomes redundant not because methodological sophistication is not needed, but such researches are results of sustained efforts that are usually done meticulously and published in reputed journals. Whatever maybe the purpose of research, the quality of work needs to be assessed scientifically to understand the real value or worth of the research work. As a researcher, you should be able to identify quality research work from others that are poorly planned, conducted and reported. It is also an essential skill of the researcher, as during literature review you should be able to give appropriate weight to the related works based on their quality. It is assumed that in social science research, the methodology is highly significant, and use of right research design lead us to right research results that are reliable, valid and useful. In this lesson, we will discuss how to evaluate the effectiveness of research reports.

18.2 OBJECTIVES

After reading this lesson, you will be able to:

List the major items for evaluating a research report

Explain the scientific criteria for evaluating a research report,

Critically assess the validity of the mechanism of evaluating research reports

18.3 EFFECTIVENESS OFRESEARCH REPORTS

An understanding of the criteria for evaluation of research report is reinforcement of the understanding of the entire process, techniques and tools of research. Evaluation of research reports is carried out for various purposes including award of degree in the case of research leading to a Master degree like MA(DE) or M.Phil/Ph.D. As in any evaluation, we need a set of criteria for evaluating the quality of a research report. It is necessary, because the criteria help us to:

Judge the adequacy of the research report;

Guide the researcher to undertake research in a planned manner;

Identify previous research that may require further validation by using different tool or a different setting; and

Guide the preparation of a good research report.

In this section, we will discuss various components of a research report and identify the issues that need to be considered while evaluating a research report.

Introduction

As mentioned earlier, the first chapter is the introduction. Introduction is the best section to learn about the researcher, particularly, in case of social and behavioural sciences, to which education and distance education belongs. It is important to understand the researcher’s affiliations and the purpose for which the study was undertaken.

Another important point is to understand whether the researcher has a commitment to a particular point of view. Whenever a researcher tries to support an assumption or a hypothesis from a particular point of view, there is a likelihood of bias. At times, the researcher gets emotionally involved and brings in a bias which is easy to detect. In case you come across a study which clearly states that the study is to prove something, it is clear that the researcher already holds a strong point of view and is only trying to prove himself/herself correct. Much of this can be understood from the emotionally charged language of the investigator in the report.

Another important source of a bias is the researcher's strong socio-cultural affiliation. For example, in a multi-cultural or a multi-lingual society, his/her strong personal affiliation may affect the process of research.

An easy way for identifying any biases in a study is to look into the researcher's efforts in collecting research and thematic literature. For example, a particular area of research may have quite a range of references and the studies might contradict one another in terms of findings. Of these, the researcher may choose only such studies that support his or her point of view instead of referring to the contradictions.

As a reviewer, you may first like to identify if there is any bias; also the extent of the bias and its likely impact on the research process and findings.

The second important point for investigation, particularly in the introductory chapter, is the kinds of argument being built up to justify the relevance of the study.

It is expected that the introductory chapter provides a broad overview which indicates the candidate understands of the broad framework in which the educational system operates and in which the specific area of his or her research is located. Further, from this broad spectrum the researcher should be able to narrow down to the problem under investigation. In this process, a good research report will quote information and data from previous research.

Thus, the entire chapter should be seen from several angles.

The broad framework of education or related area of education laid out,

The strength of arguments to justify the study – selection of the problem,

Skill in narrowing down to the identification of the problem, and

The presence of biases issues.

Finally, the test of the quality of an introductory chapter is its ability to convince the reader and the reviewer that the problem chosen is

Relevant,

Important,

Timely,

Researchable, and

Within the competence of the researcher.

Should these criteria be fulfilled, the chapter on introduction should be considered to have done well.

Review of Literature

Researchers follow two alternative paths so far as review of research literature is concerned. Some researchers review literature as a part of the introductory chapter to build up the required rationale. However, many researchers provide a separate chapter on the review of research literature. Research reports/papers comprise three major elements:

an introduction justifying the

research, methodological details and

findings and their implications.

A competent reviewer should touch upon all the three elements, though most reviewers use only findings. Very few researchers look into findings as well as the research methodology in their reviews. From the angle of findings, major objectives of a review are:

To find gaps in research: For example, if one could conceptually map out a broad area of research and see what researches have already been conducted, the exercise will automatically provide the gaps. Within a broad research area, a number of variables are likely to be impacting a particular phenomenon or process in education. Some variables may have been and some others, may not have been studied. Those variables which have not been studied are the indicators of gaps.

To identify the areas of overlap: There will be areas in which several studies have been conducted in one and the same way. The same variables have been used time and again in the research projects that are more or less similar. These are the cases of overlap. Review of literature will allow us to understand and identify such overlaps.

To identify contradictions: It is quite likely that researches conducted in one and the same area provide contradictory results. For example, there are contradictory results as to what happens to students who participate in Personal Contact Programmes in a distance education programme and who do not. A number of studies indicate that it has no impact on the performance of learners. On the other hand, there are certain studies which indicate that those who participate in PCPs perform better. The review of research literature also brings these contradictions to surface. A researcher may then conclude whether or not research on a specific issue is conclusive.

Now, these are the three components (from the angle of findings) which should help in crystallizing the problem of research.

As a reviewer, you need to check out whether the researcher has indeed identified:

Gaps

Overlaps and

Contradictions

If so, he/she has done his/her job. However, within that broad framework, you may have to make a qualitative assessment of how effectively the researcher has argued the cases of gaps, overlaps and contradictions.

The second major purpose of review is to derive guidelines for the methodology of research. From the methodology perspective, the review of research literature should help the researcher to be able to derive implications for.

Sample – sampling technique, sample size, etc.

Research designs,

Variables to be studied,

Scaling technique,

Research instruments,

Data collection,

Statistical or qualitative techniques for data analysis

Now, while you evaluate a report, you may like to check whether the researcher has adequately analyzed and reflected on the previous research studies from the various aspects of research methodology listed above. Further, does he/she indicate his/her decision to choose the research designs, variables, etc. backed up by previous studies. If so, the purpose of review has been well achieved. If not, it is merely a ritual.

There is no one way of reviewing literature. There are at least four basic patterns.

One and the most elementary pattern is where a researcher presents the findings of a study against the name of the author in one paragraph. The second paragraph refers to another, and third to yet another researcher and so on. In such a case, the researcher does not interlink one study with another.

The second type of review is when a reviewer refers to a particular set of findings and provides a few references of those who contributed to that particular finding in brackets. In this case, he/she basically clusters studies around a finding (say, relationship between two variables) or around a common mission. Still, the research does not compare any two sets of findings or any two sets of researchers.

The third pattern is when a reviewer describes a phenomenon investigating various researches as a support.

The fourth and better approach is when a researcher develops a conceptual framework of his/her research in a particular field. The conceptual framework is built on the theoretical literature and creative argumentation. Within this framework, the researcher maps out the previous research. Fitting it into a conceptual framework allows him/her to compare and contrast issues and findings, identify the gaps, overlaps, contradictions and also derive methodological implications.

Obviously, the four patterns are in a taxonomic structure. As such, you, as an evaluator of a report, will award higher credit to the researcher who uses pattern four and least to the one who resorts to pattern one for reviewing research literature.

The other quality of a good literature review is its exhaustiveness. What is the scope of the review in terms of the period and extent of journals and databases covered? How the literature search was conducted to ensure that the review is comprehensive and not significant work has been missed? The research should mention these in the literature search to explain the inclusion and exclusion criteria and their rationale. While the quality of a literature review can be judged from its presentation and nature of argumentation in a conceptual framework, the earliest and latest references adds to the reputation of the researcher as to have followed rigour in the research work.

So, as an evaluator, you will assess whether any significant related work has been missed by the researcher. If yes, then the review is not of high quality.

Objectives and Hypotheses

All research studies have a section on objectives and hypothesis. It is important to examine whether the researcher has raised very clearly the questions to which he/she is looking for a solution. These questions should be explicit — the researcher should categorically put down the questions on paper. This set of questions can be converted into objectives.

Objectives are the foundations of a research project. Eventually the objectives guide the entire process of research. The major attributes of well written objectives are –

Clarity of expression and direction-The objectives must have been stated clearly enough to indicate what the researcher is trying to investigate. It is equally important to avoid overlaps in stating objectives.

Measurability-The objectives must be stated in a manner that they are measurable; in case of qualitative research it should be possible to at least codify the data and information so that assessment can be made whether the objectives have been achieved or not.

Comprehensiveness-The objectives provide the guiding framework for a research project. Hence, the statement of objectives should be comprehensive enough to cover each and every aspect of the research study. Stating differently, nothing should be outside the purview of the stated objectives.

Judiciousness-Another important attribute is the judiciousness in and justifiability of choosing and stating objectives. For example, many young scholars, in their post graduate dissertations and doctoral theses mention "recommending future research" as one of the objectives. In all fairness, this is not feasible. Similarly, in a short time-bound project, a research objective that actually calls for sustained and long-term study becomes less feasible.

Thus, an evaluator, while evaluating the research objectives needs to examine clarity of expression, measurability, and comprehensiveness of the objectives and judiciousness in choosing and stating them.

Hypotheses, as you have read, "is a statement of causal or non-causal relationship of two or more variables under study." The proposition of a hypothesis is derived from theoretical constructs, previous research and logical analysis. More often than not, the researcher mentions the literature that leads to the formulation of hypothesis. One important task of the evaluator is to check whether the researcher has provided sound back-up from the previous research and findings and important theoretical analyses to justify his/ her formulation of hypotheses.

Hypotheses are stated either in null or directional form. Null form, does not presuppose any specific relationship, e.g. 'there will be no relationship between academic achievement and intelligence of the students'. On the contrary, directional hypotheses presuppose relationship, e.g. 'distance learners who study self-instructional material will perform better than those who study through conventional textual material'. When a null hypothesis is tested, it may point to a positive, neutral or negative relationship that can be used to derive conclusion. When a directional hypothesis is tested, it produces one of the two results – true or false. If it is false, it does not automatically show that the reverse is true.

For the purpose of evaluation, it is important to examine whether

The choice of hypotheses – null or directional, was logical and whether the researcher has adequately argued out his/her case;

The hypotheses are testable;

The hypotheses are stated clearly indicating one to one relationship between two (or more) variables; and

In case of a multivariate situation, the relationship of the cluster of independent variables vis-à-vis the criterion variable is well defined.

IV. Choice of Research Design

There are several research methods and designs that can be chosen by a researcher to achieve the objectives and test the hypotheses. While evaluating a research report, it is important to assess whether the chosen design is competent to respond to the research objectives and questions laid down. For example, if the objective is to test the impact of a broad treatment to a group of learners, it has to follow an experimental design. Similarly, if the objective is to assess the status of certain psycho-social variables in a given sample of population, it would require survey methodology. Within a survey, if the purpose is simply to describe their status and not to compare them with any standard norm or not even develop a norm, the design can be descriptive.

Hence, in evaluating research, it is necessary to check the choice of appropriate research design against the objectives. Another means of evaluating the applicability and be fittingness of the research design is to check it against the hypothesis. If the hypothesis to be tested, is formulated in terms of relationships, the study has to adopt a methodology, e.g. survey, by which relationships can be tested. Compared to it, if the hypothesis is to test the performance of two different groups against a particular type of treatment, the research design has to provide for that opportunity by adopting an experimental design. Depending on the nature of the groups, the treatment, the size of the sample and also the nature of that experiment, one would adopt a pre-experimental, quasi experimental or true experimental design, or even post-experimental research.

Thus an evaluator of a research report needs to examine the appropriateness of the choice of research design vis-a-vis the research objectives. The details of the design, e.g. type of experimental design etc. too have to be evaluated. Equally important is the argument put forward by the researcher in deciding the research design.

Choice of Variables

Choice of variables is an important step in a research project. There can be at least three sets of variables, namely, independent, dependent and intervening variables. There are

also other ways of classifying variables like socio-economic, demographic, psychological, organizational, etc. The later classification is relevant with regard to basic content of research whereas the former is directly linked to research methodology – how you deal with them while analysing research data. Here, we shall concentrate on the first set. The important point to evaluate in the choice of variables is the formulation of the dependent or the criteria variables. This is particularly important in experimental research where the impact of other variables on the criterion variable is assessed. In order that the research makes a meaningful contribution, it is important to choose the independent variables as meticulously as possible. The choice of independent variables depends upon more than one consideration. One of the considerations is the existing knowledge on the basis of previous research which shows that certain types of variables are indeed related and predict the variation of the criterion variable. The second important consideration is the assumption of the researcher
– that there are particular sets of variables that are likely to be related to the dependent variables. The third set of variables is the intervening variables. These are often ignored in research, although these actually intervene and influence the relationship between the independent and the criterion variables. On the basis of the research literature, the researcher is expected to identify such variables that are likely to influence the relationship under test. Besides the identification and classification of the variables, it is important that variables are measurable. Further, all variables may not have standard definitions. In such a case, it is expected that the researcher shall provide operational definitions and also indications of their measurability. Hence, an important consideration in evaluation is how meticulously the variables have been identified and classified under the three categories mentioned above. The second important consideration in this case is whether the researcher has provided operational definitions of at least such variables as do not have a standard meaning in the literature. The third important consideration is whether there are clear indications of the measurability of variables

VI. Research Instrumentation

Several types of research instruments have been discussed in the blocks and units of this course. These are psychological tests, achievement tests, questionnaires, opinionnaires, information blanks, inventories, interview schedules, etc. For the purpose of evaluating a research report, the important consideration is whether the instrument chosen or developed is appropriate for measuring the variables or not.

It is important to note that a research instrument is for the measurement of variables. Every variable has certain attributes of its own, amenable to measurement by different types of scaling, namely, nominal, ordinal, ratio and interval. Similarly, there are variables which are amenable only to rigorous standardized tests, like those of intelligence, reasoning ability, etc. There are others which can be measured through inventories or questionnaires. Then there are variables which necessitate the use of interviews with probing questions to be able to go into the details of a process. The common mistake in this area is the use of incompatible instruments vis-à-vis the variables being measured; for example, researchers may use a questionnaire to measure attitude. Sometimes researchers use questionnaires for conducting interviews as if a questionnaire is no different from an interview schedule. More often than not, interviewing is called for when a lead question leads to 'Yes' and/or 'If No' kind of situation.

The points to be borne in mind while evaluating research instruments are the following:

Whether the researcher has chosen an instrument that can actually measure the variables.

Whether the research instrument has been picked up from an existing stock or has been constructed by the researcher. In case of the former, whether the researcher has checked its validity, and reliability and the sample on which the original study was conducted. Whether the standardisation on the original sample is valid for the sample on which the researcher has used the instrument and drawn inferences. In case the researcher has developed the instrument on his/her own, has care been taken to check the attributes of the tool, a dependable research instrument, be it a questionnaire, inventory or an interview schedule. As an evaluator, you may like to check the reliability and validity of the instruments used to ascertain the appropriateness of the instruments.

Whether the researcher has tested the feasibility of the use of instrument. For example, a questionnaire is not a feasible instrument for the illiterates unless the researcher himself/herself records the responses. Similarly, a research instrument that requires considerable time to respond is unlikely to be responded by those who run short of time e.g. executives.

Besides these three criteria, research instruments can be examined from the angle of language, communication, provision of recording response by the respondents, etc

VII. Sample

There are two major issues that need to be considered with respect to samples, namely,

Sample Size

Sampling Technique and Type of the Sample.

The size of the sample depends on the nature of objectives of a research project and the research design. For example, in case of rigorous experimentation, it is not only difficult to handle large samples, but also not necessary. Similarly, for surveys and such other status studies, samples have to be large. The main consideration here is that there has to be an optimum size of the sample beyond which it is waste of research resources. For this purpose use of sample size determination formulae and standard tables help the researchers to be objective.

As an evaluator what is to be considered is: whether the sample size is large enough for the study and the sample size has been determined scientifically?

Another important aspect is the technique of choosing a sample. There are several techniques of choosing a sample, namely, randomization, stratified randomization, clustering, etc. In purposes of exploring a new phenomenon primarily for understanding and learning, one may use purposive sampling. Since it is a purposive sample, it is obviously not randomized. It does not have the value of generalization but can be used for generating the first set of information. The important point is to check whether there is a case for using purposive sampling. A random sample is often quoted as the ideal sample; it is necessary for generalization and the creation of new knowledge. In practical terms, strict randomization is more often used in statistical quality control in production industry. However, a modified version like stratified randomization is used in educational and social research. It provides a sound basis for generalization. Many a times, it has been observed that researchers do not use randomization, but state that random sampling method has been used. For example, distribution of a questionnaire in a gathering of 500 people and getting a return of 350 responses is not a random sampling.

The important point to check here is whether the researcher has identified the right and relevant criteria for stratification and sub-stratification of the population; and then developed a sampling frame to choose a stratified random sample, using appropriate randomization.

It is also important to look into the argument that is put forward by the researcher justifying the method of selection of the sample. There are studies where the researcher started with a particular sample size, but ended up with a considerably reduced one. In such events, although the sampling technique may have been technically correct, the researcher miscalculated the feasibility of involving the sample in the research. As a result, beginning with a stratified large whole sample, the researcher ends up with a small sample by default or with a residual sample. This may require a change in the statistical analysis of the data.

An evaluator needs to carefully evaluate the explanation provided by the researcher and the way he/she proposes to cope with it.

VIII. Data Collection and Analysis

Along with the quality of research instruments and the sampling technique, the quality of the outcome of research also depends on the quality of data itself. In turn, the quality of data is determined by the procedure of data collection. The indication of the quality of data lies in the dependability of the information collected from the sample. A normally observed problem associated with data collection is on the spot collection that provides a definite return of responses, but not necessarily quality responses because the respondent is likely to respond mechanically. When the researcher sends the instrument by mail or leaves it to the respondent to respond, there is a large amount of loss because only a small percentage of the prospective respondents respond. The data thus generated are not of the pre-determined sample but of the positive type of respondents in the sample, that does not include data from the ‘non-respondents’.

Generally speaking, in a research exercise that requires primary data, it is better to collect the data personally. In case of secondary data, it is important to check the sources of data and their trustworthiness.

The points to check here are whether the researcher has categorically recorded the details of data collection:

research instruments administered personally or by a representative or by mail,

sources and authenticity of secondary data, and

the kind of problems that the data might have in terms of quality

Data analysis can be either qualitative or quantitative. Although qualitative research and qualitative analyses are gaining momentum, a large majority of research depends on quantitative methods and statistical procedures. A major point in evaluating research is to check whether the researcher has chosen qualitative methods where the data are qualitative and objectives and hypothesis do not demand a quantitative analysis. Similarly, whether the researcher has chosen a quantitative technique where a qualitative answer is not required.

Within quantitative research, statistics can be parametric or non-parametric. The point for evaluation is whether the researcher has justified the choice of the broad option between parametric and non-parametric tests. The size of the sample is another determinant in choosing the statistical tests as a small sample often warrants non-parametric tests. Within the parametric and nonparametric tests there is a wide range of options. Whether the researcher has chosen the right test is another important point to consider. For example, within the application of simple central tendencies, it has to be seen whether the researcher has chosen mean where it is indeed the appropriate application or whether he/she has mechanically used a mean where a median or a mode would have been the more appropriate choice. If the choice of statistics has been correct, the next point to look into is the correctness of calculations. Of course, with the increasing use of computer, this problem has reduced. Given the development of the use of computers and availability of ready-made statistical computing software and other packages, it is important to examine whether the researcher has used a computer for data analysis. If not, the evaluator may have to calculate the values himself or herself.

IX. Findings and Implications

Having dealt with the issues pertaining to introduction, review of related literature, methodology, we now concentrate on results. The final outcome and hence, the value of a research project lies in its findings, i.e., results and their interpretation. The findings are presented not only in the descriptive form but also in the form of tables and graphic representations. The evaluator needs to examine whether a table or a graph has been used wherever it was required. Similarly, if a table or a graphical presentation has been provided, it has to be seen whether it has been given a title and followed up with an explanation. The second important point to check is whether the researcher has related the graphic presentation with the corresponding tabular presentation and ensured that there is no contradiction between the two. Along with the results, it is necessary to provide an interpretation and implications of the results. These are usually drawn for the purpose of policy formulation, planning and execution in education. In order to do so, a researcher often refers to previous studies to derive support for his or her findings. He/she is also expected to reflect previous studies that contradict his/her findings. The important issue of evaluation here is to critically examine the way a researcher interprets the findings:

Whether the interpretation has been related to the introductory chapter where the researcher built up the rationale for his/her research.

Whether the argument built up in each chapter is adequately reflected in the interpretation of the results.

In other words, this is where an evaluator examines the analytical skill of a researcher, the skill of extrapolation, the skill of observing and explaining relationships between two or more variables, etc.

Summary and Conclusions

This is a common chapter in all research studies. It provides a quick glimpse of the entire research project. Brevity with comprehensiveness is the rule of the game for this chapter. An evaluator often examines within the small space of a summary how a researcher has built the rationale, how he/she has documented the objectives, hypotheses, research methodology and findings.

XI. Referencing

Referencing is an important skill. Most research students make mistakes as they lack both the skill and the seriousness of purpose vis-à-vis referencing. In fact, many evaluators go straight into the section on references. It is easy to check because there are some definite internationally recognised forms. One comes across many research theses where a researcher mentions only the authors and the names of the books. He/she skips the place, publisher and even the year of publication, little realizing that an evaluator or a future reader of the thesis would not be able to refer to any book or study without the details about the publisher and the year of publication. At the very elementary level, it is necessary to check whether all the information has been provided. It has also to be checked whether the information provided is in one of the standard forms. It has to be seen if the references are indexed, in case they are not numbered in the text. Similarly, if the references are numbered in the text, it has to be seen whether the references appear serially according to their appearance in the text. As a matter of practice, as an evaluator goes through the text of the report and comes across a reference in the text, he/she should immediately check whether there is a corresponding entry in the reference. This helps him/her to identify the missing references and common mistakes in the spellings of names and years of publications.

There is a difference between a bibliography and references. Research reports require references and not a bibliography. Many students commit the mistake of providing references under the heading of bibliography. In text references not listed at the end of the chapter or thesis is a strong negative point of a quality research work.

XII. Annexure

Last but not the least, an important component of a research report is annexure. Annexure usually contain the research instruments, sampling frame, instructional material that might have been used for experimentation, etc. Annexures are also serialized. The main purpose of the annexure is that an evaluator can check the actual quality of the research instrument and material that have been used by the researcher. Annexure need to be fully documented and also serialized as indicated in the content of the research report. Thus the check points are the comprehensiveness and serialization of the annexure.

XIII. General Indicators

Besides the important points we discussed in the preceding pages, there are issues that need to be considered such as language, typographical errors, presentation, etc. Some of these issues are:

Language and expression including correctness of syntax, spelling, etc.

Typing, word processing and printing of the report – readability including margins, line spacing, type font and size, placement of tables, diagrams, illustrations and graphs.

Binding and overall get up.

18.4 FORMAT FOR EVALUATION OFRESEARCH REPORTS

In this section, we shall present a format for evaluation of research report to help you assess the quality of the report using a set of criterion, and calculate score as 'quality index'. Though this is a more quantitative way of evaluating research reports, we also recommend you to consider the issues discussed in the preceding section, and write a qualitative report. The qualitative description and the statistical 'quality index' should match.

In order to use the format, we suggest you to go through each criterion, and give relevant score in the box. If a particular criterion is not applicable for a research report, you must mention the same. While calculating the maximum score for evaluation, you must subtract the scores for the criteria that are not applicable for the specific research report. Add the total scores given by you for the different criteria, and than use the following formula for calculating "Quality Index".

Research Report Quality Index= (Score obtained) x 100/ Maximum Actual Score.

Suppose all the criteria were included in evaluating a research report, then the Maximum Actual Score would be 65. If the report gets a grand total score of 48, then the Research Report Quality Index score is 73.84.

If the Maximum Actual Score taken into consideration is 55, and the grand total score in all the categories is 48, the Research Report Quality Index score is 87.27.

18.5 LIMITATIONS IN RESEARCH REPORT

Sample size/sample bias. For example, let's say 100 hundred people should participate in your survey. Each person may give you individual results, but it does not mean that the same results belong to the whole population.

Access to data. You will not always be able to go through all the resources. You can't gather all the data you want for your research since it will take a lot of time. Because of it, your work might not cover each aspect.

Lack of time. Often deadlines are the reason why your study and research might not be complete. When we get a task, we have a limited amount of time to do it. To get a good grade, we need to submit the assignment prior to the deadline.

Financial resources. Sometimes we need some equipment or additional software to conduct the research. This might be a problem since we don't always have the sum we need.

Data collection. There are different ways to collect data: interviews, surveys, questionnaire, etc. The way you collect data might be a real limitation since the answers and the results vary.

Method. When you are finding new information, you use a specific research method. Different methods give you various opportunities. Quality of the datum you get often depends on the method you choose.

18.6 SUMMARY

In this lesson we have discussed about the criteria used for evaluating the effectiveness of research report as well as the limitations in research report. We classified research reports into three categories – research for training, research for problem solving and research for knowledge generation. We have also mentioned that there are no water-tight compartments these categories of research. Variation is on emphasis, e.g. in the case of research for training, emphasis is laid on proper use and application of research methodology. It is also anticipated that as a researcher moves from the maiden research effort to the second, third and so on, his/her methodological skills will improve. But for any research, importance of methodology cannot be underestimated. We conclude this Unit with a Research Report Evaluation Performa which by and large covers all the major points enumerated in the text of this lesson.

18.7 GLOSSARY

Research report: A research report is a well-crafted document that outlines the processes, data, and findings of a systematic investigation.

Report format: Report writing is a formal style of writing elaborately on a topic. The tone of a report is always formal.

Research follow-up: Follow-up procedures are an important component of all research. They are most often conducted during the actual research but can also be conducted afterward. Follow-up is generally done to increase the overall effectiveness of the research effort.

Oral presentation: Oral presentations, also known as public speaking or simply presentations consist of an individual or group verbally addressing an audience on a particular topic.

Communication process: The communication process has five steps: idea formation, encoding, channel selection, decoding and feedback.

18.8 SELFASSESSMENT QUESTIONS

Multiple Choice Questions:

Clarity of expression and direction:

a. Measurability b. Comprehensiveness
Judiciousness

The points which a researcher should keep in mind while evaluating the variables in research report are:

a. Formulation of variables b. Classification of variables
Whether the variables are operationally defined
Measurability of variables.

The first page of the research report is:

a. Appendix b. Bibliography

c.	Index	d.	Title page

The last page of the research report is:

a.	Appendix	b.	Bibliography
c.	Index	d.	Title page

The content of the research report is can be classified into 3 parts which are:

a.	Prefatory items	b.	Text
c.	Terminal/References	d.	All of the above

Bibliography means:

a.	Foot Note	b.	Quotations
c.	List of Books referred	d.	Biography

Interpretation should be:

a.	Subjective	b.	Objective
c.	Integrity	d.	None of the above

The last state of research proces s is:

a.	Review of literature	b.	Report writing
c.	Research design	d.	Analysis of data

In a thesis, figures and tables are included in:

a.	the appendix	b.	a separate chapter
c.	the concluding chapter	d.	the text itself

An abstract contains:

A brief summary of research problems.
A brief summary of the findings of the report.
A brief analysis of data.
A brief interpretation of data.

18.9 LESSON END EXERCISE

Examine any research report/dissertation or a thesis available to you. Read the chapter on Review of Literature. Try to evaluate the chapter and write your comments.

List the points for evaluating the chapter on Data Collection and Analysis.

State the points on which the Annexure of a research report are evaluated.

Discuss the importance of Internet reporting and research follow-up.

Explain how to use tables for presenting numerical information.

Discuss the research report from the perspective of the communications process.

18.10 SUGGESTED READING

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

William G. Z. Business Research Methods. Thomson, India.

Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

RESEARCH REPORT WRITING

UNIT-IV **Lesson No. 19**

RESEARCH ETHICS: MEANINGAND RELEVANCE OF ETHICS IN BUSINESS RESEARCH; RESEARCHERS' ETHICAL CODE

STRUCTURE

19.1 INTRODUCTION

Ethics are norms or standards of behaviour that guide moral choices about our behaviour and our relationship with others. As in other aspects of business, all parties in research should exhibit ethical behaviour. Research ethics are moral principles that guide researchers to conduct and report research without deception or intention to harm the participants of the study or members of the society as a whole, whether knowingly or unknowingly. Practising ethical guidelines while conducting and reporting research is essential to establish the validity of research. As a researcher we must follow ethical guidelines issued by regulatory committees in order to ensure the safety of the participants of a study, the public at large and that of the researcher himself/herself. Following ethical guidelines will ensure that our research is authentic and error-free, and will allow us to gain credibility and support from the public. Researcher must adhere to ethical guidelines also while presenting the findings in a manuscript. This will ensure that the article is plagiarism-free and also no unverified data reaches the readers of the article. Apart from that, research ethics fill in a sense of responsibility among researchers and make it easy to fix responsibility in case of misconduct.

19.2 OBJECTIVES

On successful completion of this lesson, you will be able to:

Understand the meaning and importance of ethics in research. Explain the significance of ethical values in business research. Explain the researchers' ethical code.

19.3 RESEARCH ETHICS

Ethical issues arise at a variety of stages in business and management research. This lesson is concerned with the concerns about ethics that might arise in the course of conducting research. The professional bodies concerned with the social sciences have been keen to spell out the ethical issues. Issues cannot be ignored, in that they relate directly to the integrity of a piece of research and of the disciplines that are involved. Discussions about the ethics of business and management research bring us into a realm in which the role of values in the research process becomes a topic of concern. Ethical issues revolve around such concerns as the following:

How should we treat the people on whom we conduct research?

Are there activities in which we should or should not engage in our relations with them?

19.3.1 Meaning

Ethics in business research refers to a code of conduct or expected societal norms of behaviour while conducting research. Ethical conduct applies to the organisation and the members that sponsor the research, the researchers who undertake the research, and the respondents who provide them with the necessary data. The observance of ethics begins with the person instituting the research, who should do so in good faith, pay attention to what the results indicate, and surrendering the ego, pursue organisational rather than self interests. Moreover, ethical conduct should also be reflected in the behaviour of the researchers who conduct the investigation, the participants who provide the data, the analysts who provide the results, and the entire research team that presents the interpretation of the results and suggests alternative solutions. Thus, ethical behaviour pervades each step of the research process, viz., data collection, data analysis, reporting, and dissemination of information on the Internet, if such an activity is undertaken. How the subjects are treated and how confidential information is safeguarded is all guided by business ethics.

There are business journals such as the Journal of Business Ethics and the Business Ethics Quarterly that are mainly devoted to the issue of ethics in business. The American Psychological Association has established certain guidelines for conducting research, to ensure that organisational research is conducted in an ethical manner and the interests of all concerned are safeguarded. Further, several ethical issues should be addressed while collecting primary data. As previously noted, these pertain to those who sponsor the research, those who collect the data, and those who offer them. The sponsors should ask for the study to be done to better the purpose of the organisation, and not for any other self serving reason. They should respect the confidentiality of the data obtained by the researcher, and not ask for the individual or group responses to be disclosed to them, or ask to see the questionnaires. They should have an open mind in accepting the results and recommendations in the report presented by the researchers.

Ethical behaviour of respondents includes:

The subject, once having exercised the choice to participate in a study, should cooperate fully in the tasks ahead, such as responding to a survey.

The respondent also has an obligation to be truthful and honest in the responses.

Misrepresentation or giving information, knowing it to be untrue, should be avoided

Further, ethics in experiments refers to the correct rules of conduct necessary when carrying out experimental research. Researchers have a duty to respect the rights and dignity of research participants. This means that they should take certain rules of conduct into account. For example, a student may face an ethical dilemma when taking a test. Another student may arrange to exchange multiple choice responses to a test via electronic text messages. This represents an ethical dilemma because there are alternative courses of action each with differing moral implications. An ethical idealist may apply a rule that cheating is always wrong and therefore would not be likely to participate in the behaviour. An ethical relativist may instead argue that the behaviour is acceptable because a lot of the other students will be doing the same. In other words, the consensus is that this sort of cheating is acceptable, so this student would be likely to go ahead and participate in the behaviour. Researchers and business stakeholders face ethical dilemmas practically every day.

19.3.2 Relevance of Ethics in Business Research

There are several reasons why it is important to adhere to ethical norms in research.

First, norms promote the aims of research, such as knowledge, truth, and avoidance of error. For example, prohibitions against fabricating, falsifying, or misrepresenting research data promote the truth and minimise error.

Second, since research often involves a great deal of cooperation and coordination among many different people in different disciplines and institutions, ethical standards promote the values that are essential to collaborative work, such as trust, accountability, mutual respect, and fairness. For example, many ethical norms in research, such as guidelines for authorship, copyright and patenting policies, data sharing policies, and confidentiality rules in peer review, are designed to protect intellectual property interests while encouraging collaboration. Most researchers want to receive credit for their contributions and do not want to have their ideas stolen or disclosed prematurely.

Third, many of the ethical norms help to ensure that researchers can be held accountable to the public. For instance, federal policies on research misconduct, conflicts of interest, the human subjects protections, and animal care and use are necessary in order to make sure that researchers who are funded by public money can be held accountable to the public.

Fourth, ethical norms in research also help to build public support for research. People are more likely to fund a research project if they can trust the quality and integrity of research.

Finally, many of the norms of research promote a variety of other important moral and social values, such as social responsibility, human rights, and animal welfare, compliance with the law, and public health and safety. Ethical lapses in research can significantly harm human and animal subjects, students, and the public. For example, a researcher who fabricates data in a clinical trial may harm or even kill patients and a researcher who fails to abide by regulations and guidelines relating to radiation or biological safety may jeopardize his health and safety or the health and safety of staff and students.

Therefore, given the importance of ethics for the conduct of research, it should come as no surprise that many different professional associations, government agencies, and universities have adopted specific codes, rules, and policies relating to research ethics. Many government agencies have ethics rules for funded researchers.

19.4 RESEARCHERS' ETHICAL CODE

Research staff and research support firms should practice good business ethics. Researchers are often the focus of discussions of business ethics because of the necessity that they interact with the public. Several professional organisations have written and adopted codes of ethics for their researchers, including the American Marketing Association, the European Society for Opinion and Market Research, and the Marketing Research Society. In addition, the researchers have rights. In particular, once a research consulting firm is hired to conduct some research, they have the right to cooperation from the sponsoring client. Also, the researchers also have the right to be paid for the work they do as long as it is done professionally. Sometimes, the client may not like the results. But not liking the results is no basis for not paying. In addition, the client should pay the researcher in full and in a timely manner.

Further, questions about ethics in business and management research also bring in the role of professional associations, such as the American Academy of Management (AAM) and the Market Research Society (MRS), which have formulated codes of ethics on behalf of their members. APA's Ethics Code mandates that psychologists who conduct research should inform participants about: The purpose of the research, expected duration and procedures. Participants' rights to decline to participate and to withdraw from the research once it has started, as well as the anticipated consequences of doing so.

The American Marketing Association, in furtherance of its central objective of the advancement of science in marketing and in recognition of its obligations to the public, has established some important principles of ethical practice of marketing research for the guidance of its members. In an increasingly complex society, marketing research is more and more dependent upon marketing information intelligently and systematically obtained. The consumer is the source of much of this information. Seeking the cooperation of the consumer in the development of information, marketing management must acknowledge its obligation to protect the public from misrepresentation and exploitation under the guise of research. Similarly, the research practitioner has an obligation to the discipline he practices and to those who provide support for his practice- an obligation to adhere to basic and commonly accepted standards of scientific investigation as they apply to the domain of marketing research. It is the intent of this code to define ethical standards required of marketing research in satisfying these obligations. Adherence to this code will assure the user of marketing research that the research was done in accordance with acceptable ethical practices. Those engaged in research will find in this code an affirmation of sound and honest basic principles that have developed over the years as the profession has grown. The field interviewers who are the points of contact between the profession and the consumer will also find guidance in fulfilling their vitally important role. The codes' of ethics for researchers are:

> No individual or organization will undertake any activity that is directly or indirectly represented to be marketing research, but that has as its real purpose the attempted sale of merchandise or services to some or all of the respondents interviewed in the course of the research.
>
> If a respondent has been led to believe, directly or indirectly, that he or she is

participating in a marketing research survey and that his or her anonymity will be protected, the respondent's name shall not be made known to anyone outside the research organization or research department, or used for anything other than research purposes.

There will be no intentional or deliberate misrepresentation of research methods or results. An adequate description of methods employed will be made available upon request to the sponsor of the research. Evidence that fieldwork has been completed according to specifications will, upon request, be made available to buyers of research.

The identity of the survey sponsor and/or the ultimate client for whom a survey is being done will be held in confidence at all times, unless this identity is to be revealed as part of the research design. Research information shall be held in confidence by the research organization or department and not used for personal gain or made available to any outside party unless the client specifically authorizes such release.

A research organisation shall not undertake studies for competitive clients when such studies would jeopardize the confidential nature of client-agency relationships.

For Users of Marketing Research

A user of research shall not knowingly disseminate conclusions from a given research project or service that are inconsistent with or not warranted by the data. 2. To the extent that there is involved in a research project a unique design involving techniques, approaches, or concepts not commonly available to research practitioners, the prospective user of research shall not solicit such a design from one practitioner and deliver it to another for execution without the approval of the design originator.

For Field Interviewers

Research assignments and materials received, as well as information obtained from respondents, shall be held in confidence by the interviewer and revealed to no one except the research organization conducting the marketing study.

No information gained through a marketing research activity shall be used, directly or indirectly, for the personal gain or advantage of the interviewer.

Interviews shall be conducted in strict accordance with specifications and instructions received.

An interviewer shall not carry out two or more interviewing assignments simultaneously unless authorized by all contractors or employers concerned. Members of the American Marketing Association will be expected to conduct themselves in accordance with provisions of this code in all of their marketing research activities.

Although codes, policies, and principles are very important and useful, like any set of rules, they do not cover every situation, they often conflict, and they require considerable interpretation. It is therefore important for researchers to learn how to interpret, assess, and apply various research rules and how to make decisions and to act ethically in various situations. The vast majority of decisions involve the straightforward application of ethical rules.

Other steps researchers should take include:

Discuss the limits of confidentiality. Give participants information about how their data will be used, what will be done with case materials, photos and audio and video recordings, and secure their consent.

Know federal and state law. Know the ins and outs of state and federal law that might apply to your research. For instance, the Goals 2000: Education Act of 1994 prohibits asking children about religion, sex or family life without parental permission. Another example is that, while most states only require licensed psychologists to comply with mandatory reporting laws, some laws also require researchers to report abuse and neglect. That’s why it’s important for researchers to plan for situations in which they may learn of such reportable offenses. Generally, research psychologists can consult with a clinician or their institution’s legal department to decide the best course of action.

Take practical security measures. Be sure confidential records are stored in a secure area with limited access, and consider stripping them of identifying information, if feasible. Also, be aware of situations where confidentiality could inadvertently be breached, such as having confidential conversations in a room that’s not soundproof or putting participants’ names on bills paid by accounting departments.

Think about data sharing before research begins. If researchers plan to share their data with others, they should note that in the consent process, specifying how they will be shared and whether data will be anonymous. For example, researchers could have difficulty sharing sensitive data they've collected in a study of adults with serious mental illnesses because they failed to ask participants for permission to share the data. Or developmental data collected on videotape may be a valuable resource for sharing, but unless a researcher asked permission back then to share videotapes; it would be unethical to do so. When sharing, psychologists should use established techniques when possible to protect confidentiality, such as coding data to hide identities. "But be aware that it may be almost impossible to entirely cloak identity, especially if your data include video or audio recordings or can be linked to larger databases," says Merry Bullock, PhD, associate executive director in APA's Science Directorate.

Understand the limits of the Internet. Since Web technology is constantly evolving, psychologists need to be technologically savvy to conduct research online and cautious when exchanging confidential information electronically. If you're not a Internet whiz, get the help of someone who is. Otherwise, it may be possible for others to tap into data that you thought was properly protected.

19.5 SUMMARY

Research is the process of finding solutions to a problem after a thorough study and analysis of the situational factors. Business research is an organised, systematic, data-based, critical, objective, inquiry or investigation into a specific problem, undertaken with the purpose of finding answers or solutions to it. In essence, business research provides the necessary information that guides managers to make informed decisions to successfully deal with problems. In one form or another, both theory and information play an important role in research. All professions are guided by a code of ethics that has evolved over the years to accommodate the changing ethos, values, needs and expectations of those who hold a stake in the professions. Some professions are more advanced than others in terms of the level of development of their code of ethics. Some have very strict guidelines, monitor conduct effectively and take appropriate steps against those who do not abide by the guidelines. Most professions have an overall code of conduct that also governs the way they carry out research. In addition, many research bodies have evolved a code of ethics separately for research. Medicine, epidemiology, business, law, education, psychology and other social sciences have wellestablished codes of ethics for research.

Ethics in business research refers to a code of conduct or expected societal norms of behaviour while conducting research. Ethical conduct applies to the organisation and the members that sponsor the research, the researchers who undertake the research, and the respondents who provide them with the necessary data. Ethical behaviour pervades each step of the research process. Certainly, there are researchers who would twist results for a client or who would fabricate results for personal gain. However, these are not professionals. When one is professional, one realises that one's actions not only have implications for oneself but also for one's field. Indeed, just a few unscrupulous researchers can give the field a bad name. Thus, researchers should maintain the highest integrity in their work to protect our industry. Research participants should also play. Their role, or else the data they provide will not lead to better products for all consumers. Finally, the research users must also follow good professional ethics in their treatment of researchers and research results.

19.6 GLOSSARY

Ethics: Code of conduct or expected societal norms of behaviour.

Business Ethics: The application of morals to behaviour related to the exchange environment.

Moral Standards: Principles that reflect beliefs about what is ethical and what is unethical

Ethical Dilemma : Refers to a situation in which one chooses from alternative courses of actions, each with different ethical implications

Relativism: A term that reflects the degree to which one rejects moral standards in favour of the acceptability of some action. This way of thinking rejects absolute principles in favour of situation-based evaluations

Idealism: A term that reflects the degree to which one bases one's morality on moral standards.

19.7 SELFASSESSMENT QUESTIONS/CHECKYOUR PROGRESS

Multiple Choice Questions:

Which term below refers to norms or standards of behavior that guide moral choices about research behaviour?

a. Politics b. Ethics
c. Standards d. Philosophy
Integrity

The goal of ethics in research is to _____.
to ensure that no one is harmed
to protect researchers from legal liability
to identify what behaviour is appropriate
to ensure research sponsors fulfill their legal obligations
none of the above

All of the following are unethical activities except _____.

violating nondisclosure agreements
revealing the names of research participants
stating that a survey will take 10 minutes to complete when the actual time is 15 minutes
invoicing for time that was not spent on the project.
providing incentives for respondents

Research should be designed such that participants do not experience _____.

a. physical harm b. discomfort
c. embarrassment d. loss of privacy
all of the above

Essential in ethics and ethical standards is:

A good grasp of research methods.

The capacity to produce good research.

A good understanding of business.

The capacity to distinguish between right and wrong.

The ethically reflective practitioner:

Engages in research on ethics.

Thinks critically about the standard of their research and their code of conduct and behavior as a researcher.

Is particularly bound by rules and standards.

Wastes a lot of time just thinking.

The two important components of research responsibility are: sincerity in work and avoiding _________________.

a. Plagiarism   b. Writing the thesis

c. Research techniques   d. Confidentiality

19.8 LESSON END EXERCISE

Why are ethical issues important in relation to the conduct of business research?

____

____

____

What are some of the difficulties of following this ethical principles?

____

____

____

Why is the privacy principle important?

____

____

____

Describe and define business research.

Demonstrate awareness of the role of ethics in business research.

Define ethics and understand how it applies to business research.

19.12 SUGGESTED READING

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

William G. Z. Business Research Methods. Thomson, India.

Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

UNIT-IV **Lesson No. 20**

NEED FOR PLAGIARISM; PUBLICATIONAND PRESENTATION OF BUSINESS RESEARCH

STRUCTURE

20.1 INTRODUCTION

Plagiarism is any unauthorised use of parts or the whole of any article without giving proper credit to the original writer. Any unethical copying of any writing is basically considered theft, and therefore it takes away the originality and trustworthiness of the content. Plagiarism is presenting someone else's work or ideas as your own, with or without their consent, by incorporating it into your work without full acknowledgement. All published and unpublished material, whether in manuscript, printed or electronic form, is covered under this definition. Plagiarism may be intentional or reckless, or unintentional. Under the regulations for examinations, intentional or reckless plagiarism is a disciplinary offence. The necessity to acknowledge others' work or ideas applies not only to text, but also to other media, such as computer code, illustrations, graphs etc. It applies equally to published text and data drawn from books and journals, and to unpublished text and data, whether from lectures, theses or other students' essays. We must also attribute text, data, or other resources downloaded from websites. The best way of avoiding plagiarism is to learn and employ the principles of good academic practice from the beginning. Avoiding plagiarism is not simply a matter of making sure your references are all correct, or changing enough words so the examiner will not notice your paraphrase; it is about deploying your academic skills to make your work as good as it can be.

20.2 OBJECTIVES

After successful completion of this lesson, students will be able to:

Understand the concept and need of plagiarism check in research.

Explain the process of presentation of research reports.

Discuss the process of publication of research reports.

20.3 NEED FOR PLAGIARISM CHECK

Plagiarism is using the words or ideas of others and passing them off as your own. Plagiarism is a type of intellectual theft. Plagiarism can take many forms, from deliberate cheating to accidentally copying from a source without acknowledgement. Consequently, whenever you use the words or ideas of another person in your work, you must acknowledge where they came from.

Figure 20.1: Steps to learning about plagiarism

It is important to know what plagiarism is, and what form it takes (some common types of plagiarism are listed here). It's also important to know how plagiarism happens. The final step is to develop effective academic skills. Many students who plagiarise do so unintentionally, often because they don't have the academic skills to avoid over-reliance on the work of others or because they aren't sure what constitutes plagiarism. In research writing, avoiding plagiarism is important because once you submit a research paper that is flagged as plagiarised, it will cost you significantly. For students, this could mean a failing grade, and for professional researchers, this could adversely affect their reputation and credibility. There are different types of plagiarism; Direct Plagiarism, Mosaic Plagiarism, Paraphrasing Plagiarism, and Self Plagiarism. Plagiarism is one of the worst offenses a scholar can commit. Even if you're a graduate student who's trying to finish a paper quickly before a deadline, you never want to resort to simply regurgitating information from your sources. In order to properly learn your field of study and develop your own views, it's crucial to do more than reproduce what other academics have said, without attribution. Avoiding plagiarism can benefit both students and professional researchers. Here are some of the reasons why it is best to avoid plagiarism in research, no matter what level of academia you belong to:

It helps you improve the quality of your work by not solely relying on and copying the ideas of others.

It saves you from the embarrassment and shame if (or when) you get caught plagiarising.

It enhances your capability to paraphrase meaning of established ideas and values in the proper way.

It can prevent grave and negative consequences in your career.

You can master conveying your ideas without having to worry about unintentionally copying someone else’s work (there are a lot of plagiarism fixers available online that can help you do this).

The three ways must be follow to cite sources correctly and avoiding plagiarism in professional or academic research. These are:

Direct Quotation: If you want to avoid plagiarism, you can use direct quotations when you use the words of the author verbatim. When doing this, you should also put indicators, such as “according to” or “in (Author)’s view” to inform your readers that you are integrating the ideas of others in your paper. However, it’s best to limit the number of times you use a direct quote and only use them when necessary. Otherwise, you run the risk of sounding like an encyclopaedia of existing knowledge, not an academic with your own insights.

That being said, here are some of the times when it is best to use a direct quote:

When the thought and meaning of the source will be different when you

reword it. When the original statement uses powerful language.

· When you are quoting a coined term or phrase that originated from the author.

Here’s an example:

Original: From 1984, by George Orwell: “Power is in tearing human minds to pieces and putting them together again in new shapes of your own choosing.”

Direct quote: As George Orwell said, “Power is in tearing human minds to pieces and putting them together again in new shapes of your own choosing.”

To properly cite and quote authors, it is also important to note that distinct citation styles require different formatting.

For example, in APA style, when you direct a quote, you must include the author's last name, the year, and the page number (Orwell, 1949, p. 266). But, in MLA, you only need to write the author's last name and the page number (Orwell 266).

Paraphrasing: Paraphrasing is when you keep the original meaning of the source, but write it in your own words. You will avoid plagiarism when you only get inspiration from the ideas of others and cite them in your reference, rather than copying their words to the letter. However, you can cite an author and paraphrase the original statement, but still be plagiarising. Therefore, in avoiding plagiarism, you should refrain from rewriting the original text. Instead, use your own words, and retain just the most significant phrases or terms.

Here's an example:

Original: "There is no doubt that China today is a superpower in every respect. According to several scholars, China has already become a global hegemon in recent years, threatening the position of the United States and the entire Western world." (Source: Tonon, A. (2021) Is China the world's next global hegemon? Roar News)

Plagiarism: According to Tonon (2021), China today is a superpower in every sense of the word. Their power and influence as a country are threatening the position of the entire Western world.

Paraphrasing: Andrea Tonon (2021) argues that with the influence and power that China holds today, they are destabilising the authority and control that the United States and the West hold over the global economy.

Use Plagiarism Checkers: Another way to ensure that your research is plagiarism-free is by using plagiarism checkers. These applications or websites help you weed out any instances of unintentional citing with plagiarism fixer tools that identify parts in your paper that need a quotation, citation, or paraphrased statements almost similar to the original text. But how do you know that you can trust these plagiarism fixer tools to help you submit research that is 100% authentic and free of plagiarism?

Flowcite: An integrated academic research platform that helps you avoid plagiarism.

Flowcite offers a suite of academic writing tools and services, including:

A similarity checker that provides and analyses your research within minutes by comparing it to external sources. It also detects citations in all citation styles – MLA, APA, Chicago, etc.

Integrated proofreading, peer-review, and publishing services, enabling you to complete the research process from writing to publishing on one platform.

A reference management tool that helps you to search, save and organise all your citations smoothly.

An AI-driven article summariser service, in partnership with Scholarly, saves you time (by 70%) in finding relevant sources. It provides a good summary of an article in less than 60 seconds, enabling you to decide immediately whether an article is relevant to your study.

Real-time collaboration tools, allowing you to edit, share, and proofread your partner's work.

The plagiarism can happen because they often find hard to write papers on their own. Thus, when a student/researcher is assigned to prepare a scientific paper that can influence his final grade significantly and he has no time, desire or ability to fulfil a task on his own, the solution comes by itself. In such situation, many students decide to take a risk and copy someone's work, neglecting the requirements of their teachers, which often lead to very unpleasant consequences. This is a wrong approach to your assignments.

To get the most accurate answer to the question "Why is plagiarism important?" or "How does plagiarism checker work" we should try to understand what reasons usually force people to use it. One of the main reasons why researchers of all levels do this is time. Due to a terrible lack of time, a large academic load, and close deadlines, many young people decide to copy a completed essay or another paper and submit it as their own, which can be understood. However, modern educational institutions have no tolerance to this form of cheating. Thus, a student/researcher dooms himself to failure.

Although it can sound a bit controversial, despite the fact that plagiarism is strictly prohibited in all universities or colleges and plagiarism consequences are dire, there is no reason to avoid it because sometimes it can be quite handy and useful. Of course, from the moral point of view, plagiarism is pure theft, which manifests itself in the form copying or use of someone else's knowledge and attribution of your authorship to it. Therefore, most often people think of it as of something bad.

However, there is also another side of the coin. If you look at this matter from other perspectives, you will not argue that without some "borrowing", the humanity would not reach neither scientific, technical, nor cultural progress. From the scientific point of view, the development of the science would be simply impossible if there were no chance to borrow, interpret or use the knowledge, works or findings of other people. Otherwise, each researcher or scientist would waste lots of time in order to study the same issue all over again. That is why many famous scientists were accused of plagiarism. And this is natural because, when it comes to indeed major problems, you can't do everything from the very beginning, the whole science is developing thanks to the shared knowledge of many people. This is why some borrowing is always allowed. However, only in the form of quotes or references. Remember that only citing your sources in a proper manner will be able to create an excellent paper using all necessary materials.

Writing a Research Paper without Plagiarising follow some rules or tips and these rules or tips are:

Never forget using quotation marks when it is needed.

Include only the most important quotes of phrases in your text.

Paraphrase the information that you want to use but do not provide in the exact formulation.

Summarise the key ideas from every source instead of copying the whole thing. Watch out for pasting too long phrases.

Always refer to the author where it is needed.

Form the list of cited sources in accordance with all rules and include all important bibliographic information in it.

20.4 PUBLICATION OF BUSINESS RESEARCH

Data publication and reporting is the process of preparing and disseminating research findings to the scientific community. Scholarly disciplines can only advance through dissemination and review of research findings at professional meetings and publications in discipline-related journals. The tacit assumption in publishing is one of trust between the author(s) and readers regarding the accuracy and truthfulness of any submission.

The practice of ensuring research integrity is relevant at all stages of research investigation, from early conceptualisation, design, implementation, to analysis. This practice also extends to the stage of documenting and preparing results for publication. In this process, researchers may experience many more challenges to preserving research integrity.

There are some considerations/issues which must be followed when reporting and publishing a research or data. There are often factors in research settings that can result in compromises to data integrity. These factors may facilitate conditions where the goal of conducting research in as objective a manner as possible can sometimes be challenged. These can be categorised as either external or internal factors. The External factors comprise:

Publication pressure

  Professional competition

  Job security

  Lack of formal mentoring

  Unclear guidelines

Lack of penalties

  Little chance of getting

caught The Internal Factors include:

  Individual ego or vanity

  Personal financial gain

Psychiatric illness

Incompetence

Sloppy writing/reporting

Investigators demonstrating lapses of integrity while engaged in data reporting and publishing can have a negative influence in the direction of future research efforts, threaten to compromise the credibility of a particular field of study, and may ultimately risk the well-being and safety of the public in general, as well as research subjects in particular. Sources of guidance promoting good data reporting practices and publishing include faculty advisors who carefully instruct graduate students, departmental chairpersons mentoring researchers new to the field, regular review of published university policies, existing codes of professional ethics, or established government rules and regulations. Deficiencies in training or a lack of awareness of existing policies, codes, or rules may increase the likelihood of a deviation from the acceptable standards of practice in reporting and publishing. There are some issues related to integrity of data reporting and publication. These are:

Misrepresentation: Due to problems data in collection, researchers may omit data that is not supportive of the research hypothesis. Alternately, data may be fabricated if the data collection process was somehow interrupted or data was lost, and the researchers believe the invented data would have been similar to what was anticipated. In either case, the true scope of the data findings remains hidden from readers who are unable to accurately assess the validity of the findings.

Plagiarism: Plagiarism is the act of taking credit for ideas or data that rightfully belongs to others. Related to this is the theft of ideas from grants and drafts of papers that a researcher has reviewed. This harms the researcher(s) from which the idea(s) or data was appropriated improperly acknowledged.

Selectivity of reporting / failure to report all pertinent data: This is the practice of only using data that supports one's research hypothesis and ignoring or omitting data that does not. A related practice is inaccurate reporting of missing data points. As explained under "Misrepresentation" earlier, the true scope of the data findings remains hidden from readers who are unable to accurately assess the validity of the findings.

Failure to disclose conflicts of interest: Editors, reviewers, or readers who are not aware of possible conflicts of interest (financial and otherwise) may not have an opportunity to adequately assess the validity of research findings without being aware of possible undue influences from the sponsors of an investigation. These conflicts may compromise researchers' credibility in their fields.

Publication bias / neglecting negative results: Since the vast majority of research findings submitted to professional journals tend to be 'positive' in nature, the literature in most scientific fields demonstrates a negative bias. This in part reflects the reluctance of journal editors to publish articles with negative findings. Thus, researchers are less willing to report findings that fail to demonstrate an intended effect or yield an expected result. The value of these publications could be substantial in that other investigators would not needlessly pursue a fruitless path of research.

Analysis of data by several methods to find a significant result: This is also known as 'milking' or 'dredging the data' and involves researchers utilising a variety of statistical tests in the hopes of yielding a significant result. The proper procedure would be to base the selection of desired tests on a theory or theoretical framework rather than selecting tests a priori. Other related statistical issues include reporting percentages rather than absolute numbers due to small sample size, reporting differences when statistical significance is not reached suggesting a certain trend exists, reporting no difference when statistical power is inadequate, and failure to include the total number of eligible participants. The importance of this last point is the difficulty for readers to be able to determine whether a dismal non-respondent rate might compromise the representativeness of respondents.

Inadequate evaluation of prior research: This refers to an insufficient review of available literature that presents an incomplete picture of the current status of a particular research area. A critique of the included citations may lack the required depth of analysis and fail to justify the need for proposed research.

Ignoring citations or prior work that challenge stated conclusions or call current findings into question: Selective inclusion of citations that minimize threats to the justification for the present study can compromise the integrity of the study. Whether done intentionally or not, omissions can have the untoward consequence of providing support for an author's position.

Misleading discussion of observations: This may result from using inappropriate statistical tests, neglecting negative results, omitting missing data points, failing to report actual numbers of eligible subjects, using inappropriate graph labels or terminology, and data dredging. These can result in readers becoming less able to objectively critique the findings.

Reporting conclusions that are not supported: Faulty data collection, inappropriate analyses, gaps in logic, and unexplained deviation from conventionally accepted methods of interpretation can result in conclusions that are not valid. Readers cannot assess the validity of the conclusions for themselves unless all the necessary information is honestly reported.

Breaking down of a single piece of research into multiple overlapping reports: This can occur when the distinction and differences in findings between reports is negligible and the focus is publishing for quantity versus quality. A related practice is submission of duplicate publications in journal from different disciplines or in different languages. The expectation is that investigators would not read journal from different fields of study or languages. Literature reviews or meta-analyses that are conducted may lead to an inaccurate assessment of findings from a particular research area due to duplicate publications of the same study in different journals.

Inappropriate use of terminology without precise definitions: A potential barrier to successful cross-disciplinary investigations is the use of field-specific terminology. Encouraging the use of precise definitions can reduce confusion and promote understanding of research conducted.

Inflation of research results for the media: This involves providing statements for public and not professional consumption that are insufficiently supported by data for the purpose of publishing un-reviewed or untested results in a non-scientific or non-scholarly magazine/media. Premature reporting of results that turn out to be unsubstantiated may compromise the credibility of a particular field.

Publishing in peer-reviewed journals or presenting in scholarly meetings is the primary mechanism for investigators to disseminate their findings to the research community. This community relies on authors(s) to report the events of a study honestly and accurately. All researchers should be aware of the issues that compromise the integrity of data reporting and publishing. Ensuring integrity is essential to promoting the credibility of all fields of study

20.5 PRESENTATION OF BUSINESS RESEARCH

Research reports are presented in writing and in oral presentations. Written presentations are far more detailed than oral presentations. Oral presentations are often done with presentation software like PowerPoint. Oral presentations cannot cover all of the details covered in the written presentations. There is simply not enough time to read the entire report at a presentation. Reading slides with multiple sentences is a sure way of boring your audience. PowerPoint slides should have few words. Each slide should focus on a single idea supported with a picture, table, or chart. The presentation is a group project. The presentation should include: a short intro, your hypotheses, a brief description of the methods, tables and/or graphs related to your findings, and an interpretation of your data. The presentations should be no more than 10 minutes long. The contents of oral presentations should include:

Title slide: Title of the talk (probably the same as your paper), the names of all group members, the class and university names, and the date the talk is given.

Materials and Methods: Clearly summarise the design. Show a picture of your organisms and justify why they are appropriate for addressing the questions mentioned above. Show a picture of your lab setup and/or of a person doing some of the lab work. Show a diorama of your experimental design (with sample sizes, number of replicates, sampling frequency, etc.). Mention what parameters you measured but do not go into detail on exact procedures used. Do state what statistical tests you used to analyze your data.

Results : First show a photograph (or sketch) that shows an interesting qualitative results (e.g., trays of plants in which one set is noticeably bigger than the other, a drawing of a happy Daphnia) and state that result. Then display the results in graphical form, reminding the audience of your hypothesis and stating whether it was supported as you do so. Use simple, clean, clearly labeled graphs with proper axis labels (no extraneous 3-D effects please). Do not use light colors (yellow, light green, or pink) in your figures, they do not show up well when projected. Indicate the results of the statistical tests on the slides by including pvalues (or asterisks/letters that indicate the significance level) on the same slides with the graphs. If you have multiple results, state them in a logical order.

Implications and Conclusions: Correctly interpret your results. Constructively address sources of error and methodological difficulties. Place your results in context and draw implications from them.

Acknowledgments: Thank anyone who provided advice or assistance. Verbally thank your audience for their attention and tell them you would be happy to answer any questions.

Moreover, presenting your research is an excellent way to get feedback on your work. Professional sociologists often make presentations to their peers to prepare for more formally writing up and eventually publishing their work. Presentations might be formal talks, either as part of a panel at a professional conference or to some other group of peers or other interested parties; less formal roundtable discussions, another common professional conference format; or posters that are displayed in some specially designated area.

20.6 SUMMARY

In this lesson, we have studied the concept of plagiarism, presentation and publication of research. Plagiarism involves the act of using others' work and trying to pass it off as your own original work. Essentially, plagiarism is stealing another person's work or ideas. It can occur intentionally and unintentionally. Students should know how to document their sources and that is called placing a citation in their work. When you understand how to cite sources and when you paraphrase information and cite the source, these steps help reduce plagiarism. Students should complete their own work in class and keep a record of sources to avoid plagiarism. When writing essays and citing sources, you should put the citation in the essay and on the reference page to avoid plagiarism.

20.7 GLOSSARY

Plagiarism: Plagiarism is presenting someone else's work or ideas as your own, with or without their consent, by incorporating it into your work without full acknowledgement.

Publication: The action of making something generally known.

Presentation: A presentation is a means of communication that can be adapted to various speaking situations, such as talking to a group, addressing a meeting.

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

20.8 SELFASSESSMENT QUESTIONS

Multiple Choice Questions:

Which among the following is not a kind of plagiarism?

a. Patch Writing
b. Untainted Plagiarism
c. Indirect Plagiarism
d. Direct Plagiarism

The two important components of research responsibility are: sincerity in work and avoiding ________________.

a. Plagiarism
b. Writing the thesis
c. Research techniques
d. Confidentiality

Plagiarism is against the principles of morality, but no legal action can be taken against the plagiariser.

a. True
b. False

Plagiarism is against the principles of morality, but no legal action can be taken against the plagiariser.

a. True b. False

20.9 LESSON END EXERCISE

How does plagiarism affect the quality of research paper?

How would plagiarism affect the researchers and the research community?

What is wrong with copying from other people's research publications?

What are the basic ways for data/research presentation?

What are some challenges to publication?

20.10 SUGGESTED READING

Cooper, D. R. and Schindler, P. S., Business Research Methods. Tata Mc Graw Hill, New Delhi.

Sekaran U. and Bougie R. Research Methods for Business: A Skill Building Approach. Wiley, India.

Srivastava, T. N. and Rego, S. Business Research Methodology. Tata Mc Graw Hill, New Delhi.

William G. Z. Business Research Methods. Thomson, India.

Harper, W. B., Ralph W. and Stanley F. S. Marketing Research: Text and Cases, Homewood, Irwin.

DIRECTORATE OF DISTANCE & ONLINE EDUCATION

UNIVERSITY OF JAMMU
JAMMU

SELF LEARNING MATERIAL
OF
BUSINESS RESEARCH METHODS
FOR
M. COM- IIIrd Semester

For the Examination to be held in 2023 onwards

Course No. **MCOMC351** **Unit : I - IV**
Subject : Business Research Methods **Lesson No. 1 to 20**

Prof. Sandeep Kour Tandon
Co-ordinator, M.Com
RoomNo. 111, 1st Floor,
Directorate of Distance & Online Education*, University of Jammu*

http:/www.distanceeducationju.in
Printed and published on behalf of Directorate of Distance & Online Education, University of Jammu, by the Director, DD&OE, University of Jammu, Jammu

BUSINESS RESEARCH METHODS

Written by :

I & II Unit

Mr. Gourav Arora
Lecturer, Computer Software Application, at National Skill Training Institute, Jammu

III & IV Unit

Dr. Ritika Sambyal
Lecturer, Commerce Udhampur Campus, University of Jammu, Jammu.

Editing and Proof reading by :

Prof. Sandeep Kour Tandon
Co-ordinator, M.Com
RoomNo. 111, 1st Floor,
DD&OE, University of Jammu

Ms. Shriya Gupta
Teacher Incharge M.Com., Room No. 205, Ist Floor DD&OE, University of Jammu, Jammu.

Ms. Avantika Bakshi
Lesson No. 10
Room No. 205
IInd Floor, DD&OE,
University of Jammu, Jammu.

Printed at : Jandiyal Printing Press / 2023/650