



**EARDA**

# **Business Research**

**DR. SHINEY CHIB**





## **EARDA PUBLICATIONS**

All Right reserved. No Part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage or retrieval system, without prior permission in writing from the publishers.

No responsibility for loss caused to any individual or organization action on or refraining from action as a result or the material in this publication can be accepted by EARDA publications or the author/editor.

## **EARDA PUBLICATIONS**

Published by EARDA Publications

International Standard Book Number (ISBN): "978-81-934246-7-4"

Euro Asia Research and Development Association

1C/14, Ramesh Nagar

Delhi Pin Code 110015

**Disclaimer:** The Contents of the paper are written by the authors. The originality, authenticity of the papers, the interpretation and views expressed therein are the sole responsibility of the authors. Although every care has been taken to avoid errors and omission, this compendium is being published on the condition and understanding that the information given in the book is merely for reference and must not be taken as having authority of or binding in any manner on the author (s), editor(s) of publisher.

The publisher believes that the content of this book does not existing copyright/intellectual property of others in any manner whatsoever. However, in case any source has not been duly attributed, the publisher may be notified in writing for necessary action.

**" This book is the outcome of a Research Project sponsored by University Grants  
Commission"**

# **BUSINESS RESEARCH**

**DR. SHINEY CHIB**

B.Tech, MBA,MIRPM,MA(Pub Admn),

MCM,LLB,MA(Psy),M.Phil(IT), MBA (Market Research), PGDAC,

PGTD (ISTD), SHRM (IIM-Ahmedabad), HR-Analytics

(IIM-Rohtak)PhD.



**About the Author :**

Dr.Shiney Chib is working as Director & Research Head with Datta Meghe Institute of Management Studies, Nagpur. She is appointed as adjunct faculty with SEGI-University, Malaysia. She has been conferred with '**Distinguished Educator Award**' by 'Discovery Education Media', '**Rajiv Gandhi Excellence Award in the field of Education**' and '**Mahila Ratan Gold Medal for Academic Excellence**' and '**Eminent Educationist Award**'. She has also received '**Academic Excellence Award**' from NIPM Nagpur chapter. Eight students have been awarded PhD under her guidance. She has industrial experience of 11 years and 15 years of teaching experience. She has more than 100 publications, in the reputed journals, National and International conferences. She has guided 12 Ph.Ds and authored 4 books. She has participated and chaired the conferences held at Hamburg, Germany, Tokyo Japan and Hong Kong. She is a life member of NIPM, NHRD, All India Commerce Association, Lion's Club International.

**Dedicated To My Parents**

**(Late) T. P. Paul (Father) & Lilly Paul (Mother)**

# CONTENTS

<u>Chapters</u>	<u>Page No.</u>
<b>1. Introduction</b>	<b>1-20</b>
<ul style="list-style-type: none"><li>• Meaning,</li><li>• Objectives &amp; Types of research,</li><li>• Research Approach,</li><li>• Research process,</li><li>• Relevance &amp; scope of research in management.</li></ul>	
<b>2. Research in management</b>	<b>21-42</b>
<ul style="list-style-type: none"><li>• General management,</li><li>• Small business innovation research(SBIR),</li><li>• Research in functional areas-marketing,</li><li>• Finance, HR and production.</li></ul>	
<b>3. Research Design</b>	<b>43-60</b>
<ul style="list-style-type: none"><li>• Features of good design,</li><li>• Types of Research Design,</li><li>• Basic Principles of experimental Design,</li><li>• Use of advanced technology in Research Design,</li><li>• Role of Research analyst.</li></ul>	
<b>4. Sampling Design</b>	<b>61-78</b>
<ul style="list-style-type: none"><li>• Steps in sample design,</li><li>• Characteristics of a good sample design,</li><li>• Probability &amp; Non-Probability sampling.</li></ul>	
<b>5. Methods Of Data Collection</b>	<b>79-110</b>
<ul style="list-style-type: none"><li>• Primary data-questionnaire &amp; interviews,</li><li>• Collection of secondary data,</li><li>• Use of computer and information technology in data collection.</li></ul>	

<b>6. Collection &amp; processing data</b>	<b>111-134</b>
• Filed work,	
• Survey errors,	
• Dat coding,	
• Editing & tabulation	
<b>7. Testing of hypothesis</b>	<b>135-151</b>
• Procedure for hypothesis testing,	
• Use of statistical techniques for testing of hypothesis	
<b>8. Interpretation of data</b>	<b>152-166</b>
• Techniques of interpretation,	
• Report writing,	
• Layout of a project report,	
• preparing research reports.	

## Chapter :1

### INTRODUCTION

#### Meaning, Objectives & Types of Research, Research Approach, Research Process, Relevance & Scope of Research in Management.

**Research is defined as –**

“ The process of systematically obtaining accurate answers to significant and pertinent questions by the use of the scientific method of gathering and interpreting information.”

--Clover & Basely

“a scientific undertaking which by means of logical and systematic techniques aims to,

- Discover new facts or verify and test old facts
- Analyse their sequences, interrelations and casual explanations
- Develop new scientific tools, concepts and theories which would facilitate reliable and valid study of human behaviour”---P.V.Young

Research is the process of systematic and in-depth study or search of any particular topic, subject or area of investigation, backed by collection compilation, presentation and interpretation of relevant details d data. It is a careful search or inquiry into any subject or subject matter, which is an endeavour to discover or find out valuable facts which would be useful for further application or utilization.

In total, Research may defined as

- Is any systematic activity carried out in the pursuit of truth?
- It is a purposive investigation
- It is the application of scientific method to add to the present pool of knowledge.
- It is an endeavour to arrive answers to intellectual and practical problems by the application of scientific method.

- It is a way of finding new ways of looking at familiar things in order to explore ways of changing it.
- It is an organized inquiry, designed and carried out to provide information for solving significant and pertinent problems.
- It is an activity that extends, corrects or verifies knowledge.
- It seeks to find explanation to unexplained phenomenon-social and physical, to clarify the doubts and misconceived facts of life.

### **OBJECTIVES OR PURPOSES OF RESEARCH :**

The objective or purpose of research are as follows:

- Research extends knowledge of human beings, social life and environment. Scientist and researches build up the wealth of knowledge through their research findings. They search answers to various questions: What, Where, When, How and Why of various phenomena.
- Research brings the hidden information that might never be discovered fully during the ordinary course of life.
- Research helps in establishing generalization and general laws and contribute to theory building in various fields of knowledge like Law of Demand, theory of consumer behaviour, theories of motivation so on.
- Research verifies and tests existing facts and theories and thus helps in improving our knowledge and ability to handle situations and events.
- General laws developed through research may enable us to make reliable predictions of events yet to happen.
- It aims to analyse inter-relationships between variables and derive to casual explanations; and thus enables us to have a better understanding of the world in which we live.
- Applied research aims at finding solution to socio-economic problems, like social unrest, unemployment, poverty, health problems, human relation problems in organization so on.
- It aims at developing new tools, concepts, and theories for a better study of unknown phenomena.

- It aids in planning and thus contributes to national development.
- Analytical studies of internal and external environment of business and non-business organization provide factual data for rational decision making-formulation of strategies and policies. Studies of their operational problems contribute to an improvement in their performance.

### **NATURE/CHARACTERISTICS OF RESEARCH:**

The characteristic of research are as follows:

- Research means search for truth. Truth means the quality of being in agreement with facts or reality. It also means an established or verified fact. To research is to get nearer to truth, to understand the reality.
- Research is the pursuit of truth with the help of study, observation, experiment and comparison. Thus it is a search for knowledge through objective and systematic method of finding solution to a problem or answer to a question.
- Research refers to a process of identifying the problem, formulating a hypothesis or objective, collecting the facts or data, analysing the same and reaching certain conclusions either in the form of solution to a problem or answer to a question.
- Research is to see “what everybody has seen and think what nobody else has thought.”
- Research is a systematic and critical investigation into a phenomenon.
- It is a purposive investigation, and aims at describing, interpreting and explaining a phenomenon.
- It adopts a scientific method.
- It is objective and logical, applying tests to validate the measuring tools and conclusion reached.
- It is based upon observable experience or empirical evidence.
- Research is directed towards finding answers to pertinent questions and solutions to problems.
- It emphasizes the development of generalization: principles or theories.

- The purpose of research is not to arrive at an answer, which is personally pleasing to the researcher, but rather one, which will stand up the test of criticism.

## **IMPORTANCE OF RESEARCH**

The importance of research is summarised as follows:

- Research extends the frontiers of knowledge.
- It brings the light information that is hidden.
- It establishes generalisations and general laws and contributes to theory building in various fields of knowledge.
- It verifies and test existing theories and help improving our knowledge and ability to handle situations and events.
- It enables us to make reliable predictions/forecasts of events yet to happen.
- It facilities analysis of inter-relation between variables and to derive casual explanations.
- Applied research/action research tries to find solutions to problems.
- It helps in developing new tools, devices, concepts, theories etc for a better study/understanding of unknown phenomena.
- It aids in planning.
- It helps in evaluation of polices and programmes.
- It aids rational decision-making.
- It provides basis for formulation of policies.
- It inculcates critical thinking and promotes the development of logical habits of thinking.
- The inventions and discoveries are the result of research only.
- The research for, research for innovations, new things, new facilities etc is contributing in a big way in the economic development in general and improving quality of human life in particular.

Research is defined as –

“ the process of systematically obtaining accurate answers to significant and pertinent questions by the use of the scientific method of gathering and interpreting information.”

--Clover & Basely

“a scientific undertaking which by means of logical and systematic techniques aims to,

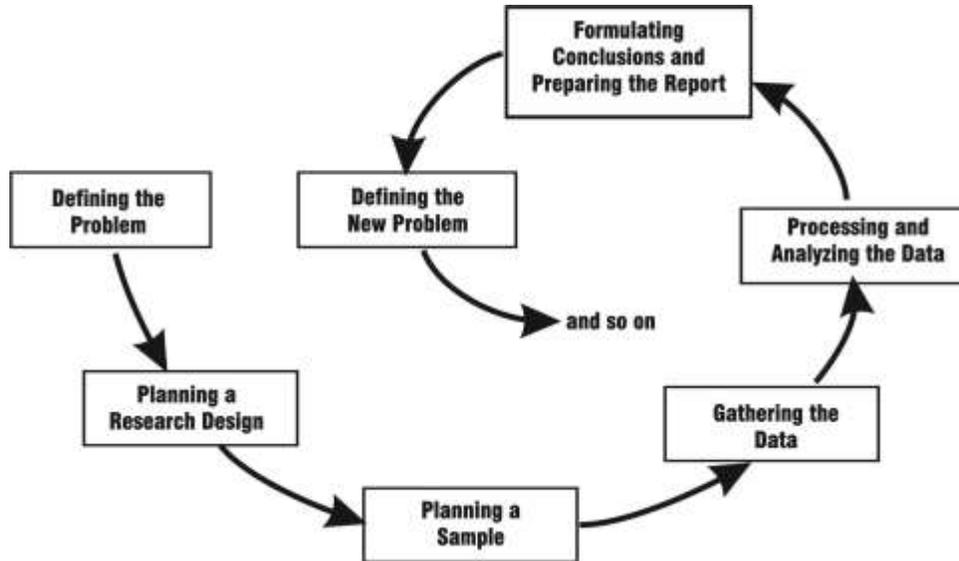
- Discover new facts or verify and test old facts
- Analyse their sequences, interrelations and casual explanations
- Develop new scientific tools, concepts and theories which would facilitate reliable and valid study of human behaviour”---P.V.Young

Research is the process of systematic and in-depth study or search of any particular topic, subject or area of investigation, backed by collection compilation, presentation and interpretation of relevant details d data. It is a careful search or inquiry into any subject or subject matter, which is an endeavour to discover or find out valuable facts which would be useful for further application or utilization.

In total, Research may defined as

- Is any systematic activity carried out in the pursuit of truth?
- It is a purposive investigation
- It is the application of scientific method to add to the present pool of knowledge.
- It is an endeavour to arrive answers to intellectual and practical problems by the application of scientific method.
- It is a way of finding new ways of looking at familiar things in order to explore ways of changing it.
- It is an organized inquiry, designed and carried out to provide information for solving significant and pertinent problems.
- It is an activity that extends, corrects or verifies knowledge.

- It seeks to find explanation to unexplained phenomenon-social and physical, to clarify the doubts and misconceived facts of life.



## **TYPES OF RESEARCH**

Research may be classified as follows:

According to Intent: Pure Research

Applied Research

Exploratory Research

Descriptive Research

Diagnostic Research

Evaluation Studies

Action Research

According to the methods of Study: Experimental Research

Analytical Study

Historical Study

## Survey

According to the nature of the data: Quantitative Research

## Qualitative Research

**PURE RESEARCH:** Pure research is also known as basic or fundamental research. It is undertaken out of intellectual curiosity. It is not necessarily problem oriented. It aims at extension of knowledge. It may lead to either discovery of new theory or refinement of an existing theory. The development of various sciences owes much too pure research. The findings of pure research enrich the storehouse of knowledge that can be drawn upon in the future to formulate significant practical researches. Pure research lays the foundation for applied *research*. The findings of pure research formed the basis for innumerable scientific and technological inventions like steam engine, machines, automobiles, electronic gadgets etc which have revolutionized and enriched our human life.

Pure research offers solution to many practical problems, for example Maslow's theory of motivation serves as a guideline for formulating incentive schemes and approaches to motivating employees in organization. Pure research helps to find the critical factors in a practical problem. For example a common sense approach to problem like communal disharmony or ethnic conflict fail to abstract the key factors. On the other hand, by deeper study social study maladies can be better understood and it may be possible to find a solution such problems. Pure research develops many alternatives solutions and thus enables to choose the best solution.

**APPLIED RESEARCH :** Applied research is carried out in life to find solution to areal-life problem requiring an action or policy decision. It is thus problem oriented and action directed. It seeks an immediate and practical result example marketing research carried on for developing a new market of for studying the post purchase experience of customers. There is a vast scope for applied research in the fields of technology, management, commerce, economics and other social sciences. Applied research can contribute new facts. A practical study designed to improve productivity in an agriculture farm may stimulate theoretical analysis of extension technology, land tenure system, price parity between agriculture inputs and outputs etc. Applied research can put theory to test. Applied research is also a scientific endeavour. The researcher has to design it scientifically. From his knowledge he has to develop a conceptual framework for his study and formulate hypothesis. Thus his study offers an opportunity to test the validity of existing theory.

Applied research may aid in conceptual clarification. It helps to integrate the previously existing theories and may help in giving new dimensions to it.

**EXPLORATORY RESEARCH / FORMULATIVE RESEARCH** : Exploratory research is preliminary study of an unfamiliar problem about which the researcher has little knowledge. It is similar to a doctor's initial investigation of a patient suffering from an unfamiliar malady for getting some clues for identifying it.

Purpose of exploratory study may be:

- To generate new ideas
- To increase the researcher's familiarity with the problem
- To make a precise formulation of the problem
- To gather information for clarifying concepts
- To determine whether it is feasible to attempt the study.

Katz conceptualizes two levels of exploratory studies, "at the first level is the discovery of the significant variables in the situation; at the second the discovery of relationship between variables". Exploratory research is necessary to get initial insight into the problems for the purpose of formulating them for more precise investigation. Hence it is also known as formularize research.

**DESCRIPTIVE RESEARCH:** Descriptive study is a fact-finding investigation with adequate interpretation. It is the simplest type of research. It is more specific than an exploratory study, as it has focus on particular aspects or dimensions of the problem studied. It is designed to gather descriptive information and provides information for formulating more sophisticated studies. Data are collected by using one or more appropriate methods: observation, interviewing and mail questionnaire. A descriptive study aims at identifying various characteristics of a community or institution or problem under study, but it does not deal with testing of proposition or hypothesis.

Descriptive study can focus directly on a theoretical point. It may be useful in verifying focal concepts through empirical observation. It highlights the important methodological aspects of data collection and interpretation. The collection of factual data increases our awareness of the relative accuracy of our measuring devices. Descriptive information obtained in research may be useful

for prediction about areas of social life outside the boundaries of the research. Descriptive studies are valuable in providing facts needed for planning social action programmes.

**DIAGNOSTIC STUDY:** It focus on discovering what is happening, why is it happening and what can be done about it. It aims at identifying the causes of a problem and the possible solution for it. It involves prior knowledge of the problem, its thorough formulation, clear-cut definition of the given population, adequate methods for collecting accurate information, precise measurement of variables, statistical analysis and test of significance. The aim *is* to obtain complete and accurate information about a given situation/phenomenon.

**EVALUATION STUDIES :** Evaluation study is one type of applied research. It is made for assessing the effectiveness of social or economic programmes implemented or for assessing the impact of developmental projects on the development of the project area. Evaluation research is thus, directed to assess or appraise the quality and quantity of an activity and its performance and to specify its attributes and conditions required for it success. It is also concerned with change over time. As Suchman puts it, "evaluative research asks about the kind of change the program views as desirable, the means by which the change is to be brought about, and the signs according to which such change can be recognized." First evaluation study is conducted for a client who intends to use the findings as a basis of decision making. This is quite different from basic research, which aims knowledge for its own sake. Second the evaluation researcher deals with his client's questions relating to the latter's programme, while basic researcher formulates his own research questions. Third the evaluation researcher measures with 'what is' rather than with comparisons of 'what is' with 'what ought to be'. Fourth unlike basic researcher who normally control over research work , the evaluation researcher works in a setting where priority goes to the programme as opposed to the evaluation. Fifth researcher-programme personnel conflicts are inherent in evaluation study. While the researcher is interested in objective evaluation and public dissemination of results, the project personnel expect that the evaluation results should mean for in-house use only.

**ACTION RESEARCH :** Action research is a type of evaluation study. It is a concurrent evaluation study of an action programme launched for solving a problem for improving an existing situation. It includes the following steps:

- Diagnosis

- Sharing of diagnostic information.
- Planning: developing change programme.
- Initiations of organizational change.
- Implementation of participation and communication process.
- Post experimental evaluation.

**EXPERIMENTAL RESEARCH** : Experimental research is designed to assess the effects of particular variables on a phenomenon by keeping other variables constant or controlled. It aims at determining whether and in what manner variables are related to each other. The factor which is influenced by other factors is called a dependent variable and the other factors, which influence it are known as independent variables. The nature of relationship between independent variables and dependent variables is perceived and stated in the form of casual hypothesis. A closely controlled procedure is adopted to test them.

**ANALYTICAL STUDY OR STATISTICAL METHOD** : Analytical study is a system of procedures and techniques of analysis applied to quantitative data. It may consist of a system of mathematical models or statistical techniques applicable to numerical data. Hence it is also called as Statistical Method. This study aims at testing hypothesis and specifying and interpreting relationships. It concentrates on analyzing data in depth and examining relationships from various angles by bringing in as many relevant variables as possible in the analysis of plan. This method is extensively used in business and other fields in which quantitative numerical data are generated. It is used for measuring variables, comparing groups and examining association between factors.

**HISTORICAL RESEARCH** : Historical research is a study of past records and other information sources with a view to reconstructing the origin and development of an institution or a movement or a system and discovering the trends in the past. Its objective is to draw explanation and generalization from the past trends in order to understand the present and to anticipate the future. It enables us to grasp our relationship with the past and to plan more intelligently for the future. The past contains the key to the present and the present influences the future. It includes the following steps:

- Feasibility of the study should be examined.

- Selection of the problem
- Data collection
- Analysis of data
- Interpretation

**SURVEYS** : Survey is a fact finding study. It is a method of research involving collection of data directly from a population or a sample thereof at particular time. It requires expert and imaginative planning, careful analysis and rational interpretation of the findings. Data may be collected by observation, or interviewing or mailing questionnaires. The analysis of data may be made by using simple or complex statistical techniques depending upon the objectives of the study. It includes the following steps:

- *Selection of a problem and its formulation*
- *Preparation of the research design.*
- *Operationalisation of concepts and construction of measuring index and scales.*
- *Sampling.*
- *Construction of tools for collection of data and their pre-test.*
- *Field work and collection of data.*
- *Processing of data and tabulation.*
- *Analysis of data*
- *Reporting.*

**CASESTUDY** : A case-study is an in depth comprehensive study of a person, a social group , an episode, a process, a situation, a programme, a community, an institution or any other social unit. A case study helps to secure a wealth of information about the unit of study, which may provide clues and ideas for further research. It provides an intensive analysis of many specific details that are overlooked in other methods. Case study may be conducted as an independent study or a supplementary investigation to a survey. The primary distinction between a case study and survey lie in the intensity and depth of investigation and its coverage.

**FIELD STUDIES** : Field studies are scientific enquiries aimed at discovering the relations and interactions among sociological, psychological and educational variables in social institutions and actual life situations like communities, schools, factories, organizations and institutions.

Steps involved in field study are as follows:

- *Preliminary planning*
- *Anthropological study*
- *Formulation of research design*
- *The pretesting of research instruments and procedures*
- *The full-scale field operations*
- *The analysis of materials.*

**QUANTITATIVE RESEARCH** : Quantitative research is based on quantitative data. The phenomena under study can be measured in terms of some quantity or amount. For example sales can be measured in terms of rupees. Quantitative research attempts precise measurement of something.

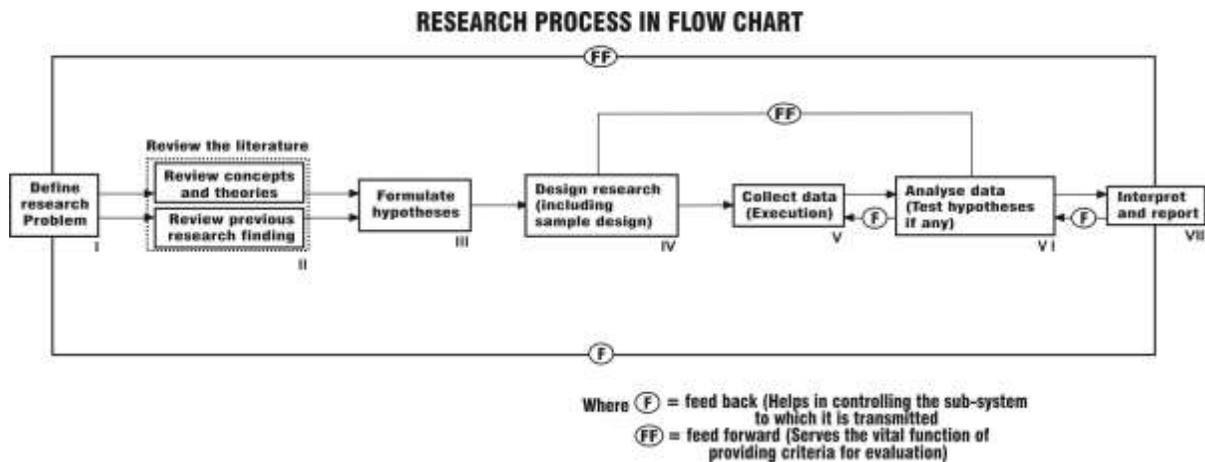
**QUALITATIVE RESEARCH** : Qualitative research is based on attributes. An attribute is a quality or characteristic which cannot be precisely measured, but whose presence or absence can be identified and counted. Qualitative research aims at discovering the underlying motives, desires, opinions, preferences etc. Using in depth interviews, direct observation, content analysis etc. Qualitative research in practice is a very difficult job.

Research process consist of series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps.

According to Nachmias, the research process has seven stages as follows:

- *Identification & selection of the research problem.*
- *Choice of a theoretical framework for the research problem.*
- *Formulation of the research problem.*
- *Design of the experiment or inquiry.*

- *Definition and measurement of variables.*
- *Sampling procedures.*
- *Tools and techniques for gathering data.*
- *Coding, editing and processing of data.*
- *Analysis of the data*
- *Reporting research*



Source : *Research Methodology, C.R.Kothari, New Age International Publishers (second revised edition), Pg. 11, Fig. 1.1*

These steps are not mutually exclusive, neither they are separate and distinct. They do not necessarily follow each other in any special order.

The very purpose of research design is, therefore to set up research project which must result in logical conclusions. For this purpose,

Research process may adopt the following steps :

- *Identifying and stating the management problem and problem area*
- *Planning the project in such a way that the observation empirically demonstrates the probability or non probability of the relationship between the phenomena*
- *Formulating the hypothesis*
- *Crystallizing, the objectives, purposes, rationale, scope and expected limitation of the research*

- *Planning to formulate the research project, identifying the resources, including financial and human resources.*
- *Identifying the types of data to be collected and its sources*
- *Specifying the methods of data collection and analysis*
- *Estimating the expected result and comparing it with the company's expectation for problem solving and decision making*
- *Finalizing a systematic scheme for proceeding with the project including sampling, survey analytical frame work and report writing.*

Identification of the management problem and making observation of the problem situation make the first step in designing a research objective. Various problems like long term problem, short term problem, policy problem, general management problem, functional area problem and so on are generally to be tackled. A proper statement of the research problem is the primary step in any research design. An academic research problem is generally based on academic considerations, while a managerial research problem is based on management practices, needs, objectives and goals.

Research analysis, *“is the systematic investigation, compilation, manipulation and presentation of information to a decision-maker in order to aid in the decision making process.”* It is basically a process rather than a specific tool or model. It may be accomplished by using specific analytical tools and models, but it is the end product, more than the analytical tools and models which have been used in its conduct, which is an aid in decision-making.

The researcher must ensure project planning as a part of design since it is planning which takes him through proper path to accomplish the goals. Incorporating the project planning with the research design facilitates the researcher observation, data collection, analytical skill, formation of hypothesis, induction and deduction skills and final conclusions. In fact it streamlines the whole research process and the researcher is enabled to foresee the result of the project. It enables the researcher to make observation which empirically demonstrates the probability and non-probability of the relationship between the phenomena.

Hypothesis forms the integral part of the research design, since it is hypothesis that provides direction for a research. Though it is a skeptical proposition which needs to be tested, it

sharpens the observation and increases skill for analysis. Hence hypothesis is properly positioned in a research design.

Objectives of research, scope of research, purpose of study, rationale and limitations must be identified when the research is designed. Objectives of the research depend on problem identification, while scope of the study is based on its utility. Rationale portrays the reasoning of study i.e. why should be undertaken. Limitations which may stand on the way of the study may also be foreseen in the designing stage itself, so that the researcher will be able to effectively tackle weaknesses and threats, and make the full use of the opportunities and strengths. Research design should include all these factors. Resources required to carry out the research project, including monetary and human resources, must be identified once the project is crystallized. This would enable the researcher to economize the resources and maximize the result.

---

Data collection is the most essential aspect of any research because the whole result of the research depends on the data and information. How much primary and secondary data must be collected, and which is the source these and many related aspects must be decided well in advance.

Methodology of data collection, analysis and testing of hypothesis must also form a part of research design. Sample size with confidence limit from specific population, how, when and from where to collect the data, and how much sampling error can be tolerated, so on can be foreseen in research design.

Expected result must be foreseen, and it should be determined whether the results will conform with the researchers needs, objectives and goals. Only when the researcher has expectations to produce a result in accordance with the expectations of the organization he may proceed further because no company can use its resources without adequate result. This aspect must be the prime in the mind of every researcher and hence it has an important place in the research design. The observations and conclusions helps to support the decision making regarding the objective of the research.

The most important task of a researcher prior to plunging into active research is to finalize a systematic scheme of research which enables the researcher to move in a systematic manner to complete the project. It includes the sequence of steps including conducting the survey, analyzing the data and the report writing and the people time and money involved in it. A complete scheme of research is prepared at this point Research in management is basically meant for a specific

purpose; and that is why management researchers are by and large result oriented. Research provides an analytical framework for the subject matter of investigation. It establishes the relationship between the different variables, especially the relationship of dependent variables with the independent variables. The cause-effect relationship between different variables can also be identified, leading to valuable observations, generalizations and conclusions. Thus research enables one to develop theories and principles, on the one hand and to arrive at generalizations, on the other. Both are aids to problem solving. Gathering primary data for analytical purposes or using secondary data for first-hand investigation should be involved in research. It stimulates the process of understanding, on the one hand and deepens the insight on the other. Obviously managerial efficiency increases.

Research in management helps to achieve the following objective of the management:

- *Decision-making objective*
- *Project objective*
- *Policy objective*
- *Controlling objective*
- *Economic and business environmental objective*
- *Market objective.*
- *Product development objective.*
- *Innovation objective.*
- *Customer satisfaction objective.*
- *Innovation objective.*
- *Profit objective.*
- *Promotional objective and*
- *Corporate image objective*

Research in management has unlimited scope in business organization. It has already been pointed out that decision-making is considerably influenced by research in the relevant areas, while the

project objective stands for the role played by research in project identification, feasibility and project implementation. There is a corporate policy for any organization, which is linked with the corporate objectives and organizational philosophy, culture and climate. Research finding influences corporate policy, research plays a vital role in shaping organizational philosophy, culture and climate. Research helps in identifying the risk and uncertainty, which in turn helps in formulating the strategies, which will help the organization to perform under the critical situations. Research helps the formulation of a standard formula, enabling the executives to rely moderately on personal judgment, especially at the middle and lower levels. Decision making calls for alternative courses of action and in identifying alternatives. Every manager is expected to be a controller and every management function has an element of controlling function in it.

Economic use of resources is one of the most important managerial functions. Optimum resource utilization has to take into account at least five major function: namely,

- *Investment problems*
- *Pricing problems*
- *Allocation problems*
- *Queuing and inventory problems and*
- *Human resource problems.*

Optimization decisions are necessarily to be based on adequate investigation and analysis of all the forgoing areas. The market objective of managerial research may be defined as research objective for the sake of controllable positioning in the market. It may include market share, penetration, profit margin, sales volume, business growth etc. Marketing research may come under this category. In connection with product development too, extensive researches are carried out. The constraints that generally affect new product development are shortage of important new product ideas, fragmented market, social and governmental restrictions and shorter life span of successful products. Research has a great role to play in order to minimize the magnitude of these problems. For the sake of profit maximization, many companies do conduct researches, while some engage consultants. As a matter of fact, both profit maximization and promotion are assisted by research findings, though they are not solely dependent on research. Shaping of corporate image is inextricably tied in with the company's relations with external groups. The manner and

method it employs to attain its desired place in industry will have a profound effect on its employees, customers, competitors and the general public. The business entity develops an institutional social philosophy which guides its actions in personnel relations, community relations, the relation with business associates and competitors, the government, shareholders etc. research has vital role to play in this respect. To sum up, research has a very vital role to play in the realm of management, and hence its scope is tremendous as far as business executive is concerned. Performing calculations almost at the speed of light the computer has become one of the most useful research tools in modern times. Computers are ideally suited for data analysis concerning large research projects. Researchers are essentially concerned with huge storage of data, their faster retrieval when required and processing of data with the aid of various techniques. In all these operations, computers are of great help. Their use, apart expediting the research work, has reduced human errors and added to the quality of research activity.

The computers can perform many statistical calculations easily and quickly. Software packages are readily available for the various simple and complicated analytical and quantitative techniques of which researchers generally make use of. The only work a researcher as to do is to feed in the data he gathered after loading the operating system and particular software package on the computer. Techniques involving trail and error process are quite frequently employed in research methodology. This involves lot of calculations and work of repetitive nature. Computer is best suited for such techniques, thus reducing the human errors on the one hand and producing the final result rapidly on the other. Innumerable data can be processed and analysed with greater ease and speed. Moreover, the results obtained are generally correct and reliable. Not only this, even the design, pictorial graphing and report are being developed with the help of computers. Researchers interested in developing skills in computer data analysis, must aware of the following steps:

- *Data organization and coding*
- *Storing the data in the computer.*
- *Selection of appropriate statistical measures/techniques.*
- *Selection of appropriate software package.*
- *Execution of computer program.*

#### **Limitations of computers.**

- Computerised analysis requires setting up of an elaborate system of monitoring, collection and feeding of data. All these require time, effort and money. Hence computer based analysis may not prove economical in case of small projects.
- Various types of detail which are not being specifically fed to computer may get lost sight of.
- The computer does not think, it can only execute the instruction of a thinking person. If poor data or faulty programs are introduced into the computer, the data analysis would not be worthwhile. The expression “garbage in, garbage out” describes this limitation very well.

A researcher can make various types errors at the different stages of the entire research.

**Defining research problem**-this error occurs when the researcher is unable to define and identify the actual problem. Many a times the problem identified can be ambiguous or can have various dimensions. More clarity and specifications are needed at this stage.

**Incorporating irrelevant variables:** these errors are called as type I and type II error which occurs due to the selection of inappropriate or insignificant variable which are totally irrelevant and ignoring the influential or important variable. Looking at the problem, the researcher needs to examine the relevant variable from prior understanding, from theory or from literature review.

**Surrogate information error:** the respondent introduces this error while answering the researcher’s question. The information required is different from the information sought.

**Measurement error:** The difference between the researcher’s requirement of the information and what the instrument provides is a measurement error. This occurs due to the validity issue of the scales used.

**Experimental error:** These error occurs due to extraneous variable in an experiment and the actual impact of the independent variables on the dependent variable are different from the impact attributable in the experiment to independent variable.

**Errors in subject selection:** This error is due to two reasons. (a) the population required is different from population actually selected. (b) the population specified is different from population listed in a frame. It is also called frame error.

**Sampling error:** this is the difference between a truly representative sample and a probability sample. This is one of the common mistakes researcher make.

**Selecting error:** This indicates errors due to the differences between a truly representative sample and a non-probability sample. It largely constitutes bias errors.

**Non-responsive error:** Errors introduced by the lack of response of certain respondents is a sample.

<b>QUESTIONS:</b>
-------------------

1. Do you agree with the statement ‘research is much concerned with proper fact finding, analysis and evaluation?’ give reasons in support of your answer?
2. Define research and examine its characteristics.
3. Discuss the purpose of research.
4. “Search for facts should be made by scientific method rather than by arbitrary method. Discuss. “
5. What do you mean by research? Explain its significance in modern times.
6. What is research? What are the broad categories in which it can be classified?
7. What is research? Explain the types of research.
8. Define research process and explain the different stages in research process.
9. What are the steps required in planning a research project?
10. Explain the relevance of research in management.
11. Describe the importance and Scope of research in management
12. Write short note on the types of errors in research.
13. Explain the role of computers in research.

\*\*\*\*\*

## Chapter :2

### RESEARCH IN MANAGEMENT

**General management, small business innovation research (SBIR), Research in functional areas-marketing, finance, HR and production.**

The Small Business Innovation Research (SBIR) program is a set-aside program (2.5% of an agency's extramural budget) for domestic small business concerns to engage in Research/Research and Development (R/R&D) that has the potential for commercialization. The SBIR program was established under the Small Business Innovation Development Act of 1982 (P.L.97-219), reauthorized until September 30,2000 by the Small Business Research and Development Enhancement Act (P.L.102-562), and reauthorized again until September 30,2008 by the Small Business Reauthorization Act of 2000 (P.L.106-554).

SBIR programs using an annual set-aside of 2.5% for small companies to conduct innovative research or research and development (R/R&D) that has potential for commercialization and public benefit. Currently, eleven Federal agencies participate in the SBIR program: the Departments of Health and human Services (DHHS), Agriculture (USDA), Commerce (DOC), Defense (DOD), Education (EoED), Energy (DOE), Homeland Security (DHS), and Transportation (DOT); the Environmental Protection Agency (EPA), the National Aeronautics and Space Administration (NASA), and the National Science Foundation (NSF). To date, over \$12 billion has been awarded by the SBIR program to various small businesses.

**Objectives.** The SBIR Program includes the following objectives: using small businesses to stimulate technological innovation, strengthening the role of small business in meeting Federal R/R&D needs, increasing private sector commercialization of innovations developed through Federal SBIR R&D, increasing small business participation in Federal R/R&D, and fostering and encouraging participation by socially and economically disadvantaged small business concerns and women-owned business concerns in the SBIR program. The STTR and SBIR programs are similar in that both programs seek to increase the participation of small businesses in Federal R&D and to increase private sector commercialization of technology developed through Federal R&D.

The unique feature of the STTR program is the requirement for the small business concern applicant organization to formally collaborate with a research institution in Phase I and Phase II.

**Differences between SBIR and STTR.** The SBIR and STTR programs differ in two major ways. First, under SBIR Program, the Principal Investigator must have his/her primary employment with the small business concern at the time of award and for the duration of the project period, however, under the STTR Program, primary employment is not stipulated. Second, the STTR Program requires research partners at universities and other non-profit research institutions to have a formal collaborative relationship with the small business concern. At least 40 percent of the STTR research project is to be conducted by the small business concern and at least 30 percent of the work is to be conducted by the single, “partnering” research institution.

### **SPSS**

Statistical Package of Social Sciences (SPSS) is a versatile and inexpensive program to use, but the strength of the program is in its simplicity. Programs generating reliable information can be managed by virtually anyone with limited computer experience, simply by using the SPSS system manual. The resultant data printouts are straightforward and easy to interpret. The intention of this report is to describe the adaptation of SPSS for use as a library management tool. The SPSS program was initially written at Stanford University in 1965 through the close cooperation of social science researchers, computer scientists, and statisticians, and statisticians. Today it has evolved into a comprehensive statistical package currently available in many computer centres. SPSS has a wide range of statistical analysis procedures including descriptive (calculator type) and inferential (cause and effect) statistics, and graphic and comparative features that enable researchers to manipulate data in useful ways. SPSS has been utilized for both inferential and descriptive analysis of data. However, the descriptive component alone can be a beneficial tool in practical library management and decision making. In a recent published report, the descriptive aspects of SPSS have been promoted as an alternative to preparing manual on-line service reports. The intention of the study conducted by the University of Nebraska Medical Centre library was to investigate the application of SPSS as an aid in calculating computer-assisted instruction (CAI) usage data.

The medical library is the central facility supporting faculty and student use of CAI programs and is responsible for the service’s financial management and administration. In

addition, the library manages a satellite CAI facility located in the Physician Assistant Program offices.

## **CAI USAGE DATA**

SPSS performs its various functions by means of a sequence of control cards that the user must prepare. These control cards contain the numerical code and instruct the system on the processing of the data. To avoid on-line storage fees, keypunch cards were chosen as the most cost-effective input mode.

Usage of SPSS's descriptive capabilities requires employing three of its procedural subprograms: Frequencies, Cross tabulation, and Breakdown. The subprogram, Frequencies, is the computation of the number of times a variable is encountered. For CAI monthly data tabulation, the frequencies requested are the number of times a program was used, the number of programs accessed by user status, the total number of programs accessed on each host system, the number of programs accessed at a particular terminal, and the number of programs accessed at a particular time of day. The subprogram, Cross tabulations, generates frequency distributions comparing two or more variables, such as the number of host system programs used compared to the time of day the program was accessed, or the status of the program user compared to the terminal location. The subprogram, Breakdown, calculates and prints the sums, means, standard deviations, and variances of one field against another. For example, minutes of program use can be "broken down" by program name alternatively, minutes of program use can be "broken down" by

- (1) The status of the user,
- (2) The program name, and
- (3) The terminal used.

This breakdown procedure is also used with a compute card to determine anticipated costs for each host system. The library executes the statistical runs in a batch mode at the university computer centre during open access periods, times provided for use of the mainframe computer when it is not otherwise engaged in university business. The library incurs no expense in utilizing the SPSS program when statistics are tabulated during these open access periods.

## **ADVANTAGES**

The advantages of adapting the SPSS program are :

- (1) The speed of calculation and analysis of data by the computer.
- (2) the many combinations of data comparisons,
- (3) the resultant ease in verification of monthly bills and preparation of monthly and fiscal year reports, and
- (4) The ability to use the data to assess growth of the service program and to identify usage trends for future budgetary planning.

In a record-keeping situation, SPSS has proven to be very efficient in analyzing CAI usage data. This suggests that SPSS may be applicable to other library record-keeping systems such as on-line usage statistics, circulation figures, or collection analysis. However, SPSS is only one of many commercially available programs potentially useful in library management. Consultation with local computer analysis will help to determine the appropriate software package adaptable to the library's administrative needs. In times of financial restraint, utilization of commercially-produced statistical software packages for administrative record keeping can save personnel time and aid in long-range planning. SPSS is a versatile and simple program to modify for library record keeping. While the majority of SPSS usage has been directed at research analysis, it has proven to be a valuable management planning tool for the library.

## **MARKETING RESEARCH**

**Definition and meaning of marketing research** : Marketing research, according to the American Marketing Association, is the systematic and objective identification, collection, analysis, dissemination and use of information for the purpose of improving decision making related to the identification and solution of problems and opportunities in marketing. First, marketing research is systematic, i.e. systematic planning is required in all the stages of the marketing research process. The procedure followed at each stage should be methodologically sound, well documented and as much as possible, planned in advance. Marketing research should be objective and should be conducted impartially. Therefore, in the marketing research process, marketing information is formally gathered, stored, analysed and distributed to managers in accordance with their informational needs at regular intervals on a planned basis.

The system is built upon an understanding of the information needs of marketing, and ability to supply that information when, where and how the managers require it. Data is derived

from the marketing environment and transferred into information that marketing managers can use in their decision making.

**Marketing Information System (MIS)** is a formalized set of procedures for generating, analyzing, storing, and distributing information to marketing decision makers on an ongoing basis. The definition of MIS is similar to marketing research, except that MIS provides information continuously rather than on the basis of ad hoc research studies.

- MIS comprises four elements :
- Internal continuous data
- Internal ad-hoc data
- Environmental scanning
- Marketing research

### **Internal Continuous Data**

MIS can convert financial data like profitability of a particular product, customer or a distribution channel into a form usable by the marketing department. This is done by means of disaggregating the database of sales of products to customers. Information like allocation of discounts, promotional and transport costs of products, etc. are stored in the MLS. A detailed description of transactions with the customers and the associated costs allow marketers to carry out analysis of their marketing activities.

Sales forces are monitored by means of recording the sales achieved, the number of new accounts opened, size of orders, number of new accounts opened, size of orders, number of calls made, etc. This can be recorded in total or broken down by product or customer. This data can provide information on sales force effectiveness.

The data of customer transactions and associated costs can also be used for specific purposes. Management may look at how sales of any has reacted to a price increase or change in advertising copy. Capturing data on MIS allow specific analysis to be conducted when needed.

**Environmental Scanning :** Environmental analysis whereby economic, social, legal, and technological forces are monitored should be considered part of the MIS. These are the forces that shape the context within which suppliers, companies, distributors and the competition do business.

Environmental scanning provides an early warning system for the forces which may impact a company's products and markets in the future. Scanning enables an organization to act upon, rather than react to, opportunities and threats. The focus is on the longer term perspective allowing a company to be in a position to plan ahead. It is a major input into strategic decisions.

**MARKETING RESEARCH :** While environmental scanning focuses on the longer term, marketing research considers the more immediate situation. It is concerned with the provision of information about markets and reaction of these to various product, price, distribution and promotion actions. Marketing research contributes heavily to marketing mix planning.

There are two types of marketing research:

- External continuous data includes television audience monitoring and consumer panels where house hold purchases are recorded over time.
- External ad hoc data which is gathered by means of surveys into specific marketing issues including usage and attitude studies, advertising, product testing etc.

### **Applications of marketing research**

Organizations conduct marketing research primarily for two purposes:

- (i) To identify marketing problems, and
- (ii) To solve marketing problems

Companies conduct researches to identify marketing problems. Typical marketing problems relate to market share, shifts in consumer tastes and preferences, competition, sales forecasting, short and long range planning, and company / brand. Those researches that aim at solving marketing problems focus on devising solutions for the above problems. Research findings in this case are used to identify solutions to marketing problems. Research findings in this case are used to identify solutions to marketing problems. Various methods are used to devise solutions for problems such as identifying attractive segments, formulating the right pricing policies, new product development and planning effective and efficient distribution solutions.

**Continuous Research Interview :** In this method the same respondents are interviewed repeatedly. Respondents are enrolled by the research agency. Information is gathered from these

respondents on a periodic basis. Thus, it is possible to track changes among the same set of audience over a period of time.

**Consumer panels:** Consumer panels are formed by recruiting a large number of households which provide information about their purchases over time. By using the same households and tracking the same variables over a period of time, measures of brand loyalty and switching can be determined. The demographic profile of the type of person who buys a particular brand can also be found out. Changes in market share can also be examined over time. Thus, it is possible to track even small behavioural changes in response to changes in marketing variables.

**Retail audits:** Sales of brands can be measured by means of laser scans of barcodes on packages which are read at the checkout counter. Brand loyalty and switching cannot be measured, but accurate assessments of sales achieved by the store and competitive activity is provided. For identifying geographic areas or type of outlets where new products may be introduced, such audits can be particularly useful. Potential problems related to distribution, in-store promotions or layout can also be assessed by using retail audits. Sales potential and sales forecasts can also be planned with such data.

**Television Viewership Panel:** The audience size of a programme is measured minute by minute. Commercial breaks can be allocated rating points according to the proportion of the target audience watching the programme. This is the currency by which television advertising is bought and judged. People meters record whether the set is on or off and which channel is being watched, a hand console records who is watching. The accuracy of the data depends on the representativeness of the panel and the extent of unbiased audience responses.

**Marketing research suppliers and services:** Marketing research suppliers and services provide most of the information needed for making marketing decisions. Marketing research suppliers can be broadly classified as internal or external.

**An internal supplier** is a marketing research department within the firm. The marketing research department's place in the organizational structure may vary considerably. At the extreme, the research department may be centralized and located at the corporate headquarters. At the other extreme is a decentralized structure in which the marketing research function is organized along divisional lines. In a decentralized structure, the company may be organized into divisions by

products, customers, or geographical regions, with marketing research personnel assigned to various divisions.

**External suppliers** are external firms hired to supply marketing research data. These external suppliers, which collectively comprise the marketing research industry, range from the small, one of few persons operations to vary large global corporations. External suppliers can be further classified as full service suppliers or limited service suppliers. Full service suppliers offer the entire range of marketing research services, from problem definition, approach development, questionnaire development, sampling, data collection, data analysis, and interpretation, to report preparation and presentation. The services provided can be further broken down the syndicated services, standardized services, customized services and internet services.

**Syndicated services** collect information of known commercial value that they provide to multiple clients on a subscription basis. Surveys, diary panels, scanners, and audits are the main means by which this data is collected. For example, in India the television audience rating measurements provided by TAM is a syndicated research that is available at a price to all media buying houses.

**Standardized services** are research studies conducted for different client firms but in a standard way. For example, procedures for measuring effectiveness of advertising have been standardized so that the results can be compared across studies and evaluative norms can be established.

**Customized services** offer a wide variety of marketing services customized to suit a specific client's needs. Each marketing research project is treated uniquely.

**Internet services** are offered by several marketing research firms including some who have specialized in conducting marketing research on the internet.

**Limited service providers** specialize in one or a few phases of the marketing research process. Services of such suppliers are classified as field services (data collection), coding and data entry, analytical services (designing questionnaires and drafting the sampling plan), data analysis, and branded products (that are proprietary models of these companies).

Researchers are not rare in the production area. Product development research, innovation research, cost reduction research, performance improvement research, work simplification research, profitability improvement research, inventory control research, product design analysis, process improvement studies, manufacturing process research – these are some of the prominent

areas of research in the production function. In fact, there are numerous areas for research in production and materials, while research and development activities mainly concentrate on product development and the production process.

Manufacturing research identifies new and better ways of producing goods, invents new technologies, reduces costs and improves product quality. Capacitors for electrical circuits, for example, were once made by interleaving metal in alternate layers with paper or plastic. This process was costly because it used a great deal of metal. Now capacitors are made by evaporating thin films of metal on paper or plastic, which is then wound into a roll. This process, by which less metal is used and better capacitors are produced, is the result of manufacturing research. Electronics, microminiaturization, solid state devices, computers and magnetic alloys, and the concomitant tying together of manufacturing research and product design are the result of constant research in production.

The sources of manufacturing research projects or topics may consist of :

- i. Suggestions from the top management ;*
- ii. Proposals from production engineers;*
- iii. Suggestions from the research staff;*
- iv. Hints from product research;*
- v. Suggestions from the marketing department;*
- vi. Interest generated by the general technological scene, etc.*

The information from the production system may include details of raw material inventory, finished goods inventors, work-in-process inventories and component parts inventories. The data on maintenance and the characteristics of production equipment may be used for equipment replacement analysis. The data on the composition of the work force, the availability of work force skills, turnover, absenteeism, shifts, labour-turnover, etc. also are relevant to the studies in production management, while cost information is useful for cost control studies, cost benefits analysis, and such other areas.

**Selection of a Manufacturing Research Project:** It is often a difficult task to select a worthwhile research project based on definite criteria. An adequate rate of return should be the

basic characteristic of any production research, No organization can afford to go in for a trial and error method. For example, in a cost reduction engineering project, the rate of return may be worked out by bearing in mind;

- i. *The cost of the engineering effort required to carry out the project.*
- ii. *The investment to be made in the plant for the purpose of development and modernization,*
- iii. *The contribution of the new project to savings over a period of years,*
- iv. *The average rate of return as a percentage and ROI*
- v. *A comparison among the alternative technologies*

It is not easy to calculate the rate of return accurately. But the effort required and the potential benefits should be examined to make certain that the game is worth the candle. Long term profitability is another important factor to be taken into account before a manufacturing project is selected. There should be reasonable probability for success if any manufacturing research should be undertaken. The time factor is another important aspect. One would not ordinarily engage in manufacturing research if the project is so long range that the benefits of success will be in the far distant future, because further technological change may take place in the long run.

The outcome of the study should have a wide scope of application, which is one of the decisive criteria for the selection of research projects in production. However, certain projects of limited applicability would initially be acceptable, if there is a possibility of extending the results to several similar field of activity. For example, in the research on the development of the powder metallurgy technique in making bearings, the porosity of the resulting product was a feature that made the process particularly desirable. Only when the result of the investigation is of utility value and the organization concerned intends to make use of it, does the project have viability. Manufacturing researchers, therefore, make it a point to select the projects which are backed by the need for satisfying factors, so that the pressure in favour of adoption may make their utilization more likely. The probability of acquiring patents is another positive factor aiding the selection of a production research. However, in certain cases, there may not be scope for patents, though they are of great utility. For example a research that leads to statistical quality control and quality

assurance may not have any scope for patent development, though it is of a very high utility. A very prominent factor to be taken into account in the selection of a production research project is that it should be compatible with existing staff skills. If, however, certain people do not develop enthusiasm another group may be entrusted with the responsibility, especially in a large organization.

The scientific method is the basis of any research. However, in manufacturing research, the researcher is expected to undertake more result oriented research projects than method oriented theoretical studies. Some people, who have the training, background and intellectual capacity to undertake research in manufacturing technology have a strong bias toward pure or basic research. They avoid work which is justified by practical needs. One of the most important criteria in selecting a project is the extent to which it contributes to the knowledge and application of the company in current fields or a field which has utility value for the company. Manufacturing research should not only provide answers to definite problems but build up a stock of knowledge on how to make things. The production planner should be in a position to draw vital information from the fund of knowledge contributed by manufacturing research. Unfortunately, many researchers conveniently ignore this need. Such research studies then become intellectual gimmicks. It is, however, unwise on the part of the researcher to select gigantic projects which tie up huge resources of the organization even if all the other factors are favourable. The selection of a research project in manufacturing management should not be based on hit-or-miss estimates or cursory considerations. The choice of a research project should be a prudent action which takes in to account all the relevant aspects, including utility, outcome, time, money, manpower and obviously the corporate objectives.

**RESEARCH IN PRODUCTION** : A considerable percentage of industrial development work is done in manufacturing and engineering cost forms a substantial portion of the production cost. Research in production, therefore, is capable of bringing about cost reduction. The industrial and production processes have become increasingly complex nowadays, especially in the context of dynamic technological development, complex labour problems backed by politically motivated and militant trade unions, conspicuous socio-economic changes, changes in human values, tastes and preferences, and government policies. In such a dynamic business environment, the need for manufacturing research is vital for the purpose of establishing a manufacturing climate which is in harmony with environmental conditions.

Research in production aids innovation, product development, product differentiation and diversification, on the one hand, and industrial development, on the other. Research and development encourage the continuity of activity level in the manufacturing organization, which creates a climate for the development and growth for production activity. Research in production not only reduces the cost of material and manpower, but lowers the inventory carrying cost as well. It leads to a reduction in the volume of unskilled labour. Moreover, it increases the efficiency and skill of the work force and shop floor supervisors, and enables the organization to update itself. It reduces the possibility of obsolescence.

**RESEARCH IN OPERATION :** Any research result in the area of production should have a practical application; and; there should be an effective system for moving technology from research to operations. Technology is a knowledge of the physical and life sciences as applied to practical purposes. This knowledge is effectively developed by a number of technical activities, and in scientific ways as in fundamental research, applied research, etc. While fundamental research seeks to establish the principles and relations underlying this knowledge, applied research crystallises the true nature of the knowledge and demonstrates its potential utility by the employment of bench scale apparatus. The development of the production process puts this knowledge to practical use in a workable prototype form, while engineering refines the knowledge for commercial exploitation, or other practical uses. In other words, this knowledge or technology is the result of research, which is put to use by the operational management. Many companies have faced the problems of moving research results effectively on the shop floor. It is mainly because no single solution is universally applicable that there cannot be a very acceptable procedure to reduce the research results to practice in all situations.

Quinn and Mueller have, however, suggested a general framework for a four step programme which has been derived from the accumulated experience of over 200 operating and research managers of leading US companies.

- **Step 1.** Examine technological transfer points. Analyse the critical points across which technology must flow if it is to be successfully exploited. Recognise the potential resistances to the flow at each of these interfaces.
- **Step 2.** Provide information to target research. Generate adequate information so that research can be targeted towards company goals and needs. Develop a comprehensive long

range planning programme to determine what technology is relevant to the company's future and to serve as a focal point for information flows. Establish special organizations, where needs, to seek specific new technological opportunities for the company, to provide commercial intelligence information about competitive activities and to make careful economic evaluations / technologically based new ventures.

- **Step 3.** Foster a positive motivational environment. Establish a motivational environment which actively stimulates technological progress and its associate organizational change. Develop tough-minded top-management attitudes, policies, and long term controls which foster rather than hinder the production and use of new technologies.
- **Step 4.** Plan and control exploitation of research and development results. Design special exploitation organizations and procedures to ensure:
  - a) That competent groups have both the authority and the obligation to develop new technologies at each of their critical stages :
  - b) That each major technological transfer is planned and monitored to control effort expenditures, cash flows and timing ; and
  - c) That the entire transfer system implements the critical strategy, which determines success or failure, at each major technological transfer point.

Despite serious attempts, many companies have their own constraints to target H&D and create a positive motivational environment for moving the research results to the act additional field. When the technologies and products are totally new to the company the magnitude of constraints widens, and the difficulties to reduce the research result to operational areas are serious. Then entirely new skills may be required at each technical stage i.e. scientific inquiry, reduction to practice, entrepreneuring a new enterprise and operating a full scale facility. The synchronization of these skills to successfully exploit a new technology calls for a strong co-ordinating authority which cuts across the usual formal lines of research, development, engineering, marketing, and manufacturing. In the absence of such a prudent co-ordination, each functional group may tend to look after its own interests and thus jeopardize the common and organization objectives. Then there is the possibility of unnecessary delays, dilution of

effectiveness of the new technology and cost escalation sometimes, even total abandonment. In order to avoid such an eventuality, the company should have an action plan consisting of :

- i. The creation of specific formal organizations of experts and competent people.
- ii. The establishment of thorough procedures for planning and monitoring of the effect and
- iii. The integration of these organizations and procedures to implement the critical strategy that determines the success of each specific technological exploitation.

If a large company wants to diversify effectively through research or to make radical technical advances, the stimulus must come from the top corporate level. The top management must be willing to underwrite long term fundamental and new product research, backed by corporate-level opportunity seeking; and economic evaluation groups should help target research activities and refine the commercial potential of research results. The company should then develop a procedure which efficiently links together the series of sequential steps, leading from the initial technical inquiries to the eventual goal of exploitation. It must ensure that the programme is adequately planned – from research to exploitation. It must provide a basis for controlling the transfer of research throughout its cycle. The procedure must monitor the three factors related to technical progress, viz. technical and commercial effort versus results, rupee costs and returns and the time interval. For the purpose of planning and controlling such techniques as PERT / CPM, budget reviews, venture analyses, rational , decision models, etc. are widely used. As each major programme comes through the management is in a position to evaluate and incorporate changes , if any, from time to time. order to promote the smooth and efficient running of its affairs. It may be held before the production cycle. In the form of raw material inventory, at an intermediate stage in the production cycle, as in process inventory; or as the end of the production cycle, as finished goods inventory”. There may be a spare parts inventory, too, depending on the actual need for these parts. However, the finished goods inventory is considered to be the most prominent, through the procedure developed can be adapted with minor changes in other types of inventory as well. There can be inventory control in uncertainly without reordering, inventory control with ordering with a certain amount of demand, and inventory with recording the uncertain demand.

A certain amount of inventory is essential for efficient functioning in any company. It, however, depends on the actual of various factors, such as economic order quantity (EOQ), lead time, the magnitude of certainly, determinant forces, and so on. Control over inventory can be

exercised by changing the timing of production runs, by changing the size of the runs and by changes in promotional effort or sales inducements. Inventory control becomes more effective when the marketing activity is constant, or demand is determined by various methods. It is in this context that the marketing and market researches assume significance: inventory studies enable the manager to determine the exact size of the inventory to be maintained. Of course there are advantages and disadvantages associated with increased inventory. The advantages include economies of production with large run sizes: faster supply of orders to customers: stabilize workloads : more than proportionate profits and windfalls from speculation in a market with inflationary tendencies.

The disadvantages of a large inventory are : greater inventory carrying cost (examples, ware house rent, depreciation and deterioration, interest on additional cost, physical handling and accounting), blocking up of funds and increased uncertainty, In normal circumstances, increased inventory is desirable only when the marginal profits considerably exceeds the marginal inventory carrying cost. Systematic research has a great role to play in determining the economic inventory level (EIL).

To sum up, it may be asserted that research is inseparable from the functional area of production and materials management.

**RESEARCH IN FINANCE DEPARTMENT :** The growth and development of a business organization is directly linked with the success of its finance function, which is concerned with determining long term and short term financial objectives, formulating and formulating and promulgating financial policies, and developing procedures that aid in the promulgation of the firm's policies. The finance function includes financial control, which consists of two steps, viz.,

- i. Developing standards of performance ; And
- ii. Comparing activities with these standards.

Financial managers cannot evade the responsibility of replanting either. In the pursuit of managing finance, the firm's primary goals have to be converted into the immediate goals of the finance department. Usually, there are two approaches of financial goals. These are

- i. Profit Risk Approach to Financial Goals ; And
- ii. Liquidity – profitability Approach to Financial Goals.

This does not, however, signify that all other approaches are ruled out.

The profit risk approach recognizes that finance deals with creating a manner framework to maximize profits while minimizing risk. For the achievement of this balance, controls over fund flows are ensured and adequate flexibility to respond to changes in the operating environment is maintained. This approach generally takes care of four aspects – maximizing profit, minimizing risk, maintaining control and ensuring flexibility. For profit maximization a high level of long term corporate profit is planned by the financial management. This is possible only if the finance department is enriched with sufficient information inflow on the internal and external forces that influence corporate goals, decisions, and actions. The Management information System (MIS) aids in this process, provided it is backed by adequate research and analysis. The finance executive anticipates the problem areas, and considers ways and means of tackling the difficulties in order to plan the steps that would enable him to minimize risk. Anticipating problem areas and alternative courses of action, as well as choosing the best course – these depend on a systematic analysis of the situation, which is the work of a research analyst. Maintaining control over flows of funds requires a financial reporting system which can provide a timely and accurate picture of the firm's activities. It enables one to locate and correct errors or weaknesses, if any, from time to time. That is why, the collection, computation, presentation and interpretation of up to date are the basic tools of any financial reporting system.

Uncertainty is a part of any business, and therefore, sufficient provision is made for financial flexibility to tide over the uncertainty. Flexibility refers to an appropriate relationship between the finance function and other functional areas, which are interdependent. An adequate information flow between them is prerequisite. And here, too, research plays an important role.

The liquidity-profitability approach emphasizes the two variables of liquidity and profitability. Liquidity indicates that a firm had adequate cash on hand to meet its obligations at all times, including unforeseen situations. It calls for the maintenance of a cash reserve for emergencies as well. Profitability is the rate of profit on investment. The achievement of liquidity in the liquidity-profitability approach requires the minimization of risk and an adequate control over the firm's activities. Thus, both the approaches are interlinked. Profitability analysis itself is a sort of research, while the maintenance of liquidity for emergencies and unforeseen situations is

a result of the studies of the future (forecasting and adaptive planning). That is why, systematic studies occupy a central place in the liquidity profitability analysis.

**FUNCTIONS OF FINANCIAL MANAGEMENT :** Liquidity and profitability are essential elements for financial management, though assets management and funds management are not unimportant . The financial manager's functions in relation to liquidity include forecasting cash flows, and managing the flow of internal funds. The success of any funds flow management depends on the matching of cash inflows with cash outflows. Cash flow forecasting, based on an analysis and interpretation of the appropriate data, is the basis of the matching process. The financial manager is expected to identify the various sources of funds available for the company from time to time. It is his duty to analyze each source of funds in comparison with respective needs at the given point of time before choosing the best and the most economic source of funds. Any decision on raising funds should be based on an adequate analytical framework. For example, a company may have a number of alternative wrecks. Viz... new shares, debentures, market borrowing, bank loans, etc. The finance department has to analyse all the sources in comparison with the nature of the company's requirements, and the availability of funds, and to choose the best from among them. Even in managing the funds flows effectively, a constant vigil should be maintained on the various bank accounts and funds. All these postulate some sort of research and analysis.

The functions that lead to probability include cost control, pricing, forecasting future profits, and measuring the cost of capital, among other things. A detailed cost accounting system is a part of large-scale business operation, which requires a operation, which requires a large quantity of processed data, and its systematic presentation and interpretation. Pricing decisions are joint decisions of the marketing and finance managers; and both cost and marketing data and analytical framework support them In order to forecast future profiles, an adequate analysis of current costs, likely cost escalations, likely changes in the ability of the firm to sell lilt products, market changes, internal and external forces- all these are essential. Both debts and equity make relative contributions to the capital requirement of the firm. Each source of capital involves the cost of capital. Short-term debt may be more expensive than long-term debt. Preferred or common stock yields different returns to the holders. All these factors have to be studied in detail with a view to measuring the cost of capital. The cost component of each source is investigated in detail prior to determining the profit margin that will pay for the borrowed funds and yield a satisfactory

return on investment (ROI) as well as a share of the factors of production. Profitability analysis, therefore, is a sort of research conducted by the finance department.

In the management of funds, and the overall finance function, there is an element of research, though it may not be identified as a formal research activity. Hampton has observed: “Financial tools should be applied in a logical, overall problem-solving process. It is the manager’s job to find a systematic manner of developing valid information from available data.” The operation of financial tools can be explained in certain steps in sequence. These sequential steps are almost identified to the steps in research. These are:

- i. Gather the relevant data;
- ii. Process the data;
- iii. Examine the information;
- iv. Select the appropriate tools;
- v. Apply the tools to decision-making;
- vi. Identify alternative courses of action; and
- vii. Select the best alternative.

As a matter of fact, research is the basic ingredient of financial decision-making. The tools of financial analysis, viz., funds flow analysis, ratio analysis, cost-volume profit analysis (break-even analysis), an analysis of financial leverages, etc. Extensively make use of research tools in the process of analysis. It may, therefore, be conducted that research tools are closely linked with the tools of financial management’s and that the finance function largely depends on research tools, either directly or indirectly.

**RESEARCH IN HUMAN RESOURCE MANAGEMENT** : The personnel function is not an exception in the application of research as an aid to effective decision-making. It make use of research more than any other functional area, because human behavior is prone to constant changes; and the personnel management’s task is to deal with human attitudes to work life differ substantially from those to social life. Individual approaches and motives are not identical; and heredity’s influence and environment’s impact on human behavior affect organizational behavior

and productivity. Research in personnel management is, therefore, more vital than in any other functional area because of the greater importance of the human element in the enterprise.

An individual's motivation to work is associated with his needs and their satisfaction. These needs may be physical needs, security needs, social needs, egoistic needs, and so on. An effective personnel manager determines the respective needs of his work force, and formulates personnel policies to satisfy those needs, so that they can be motivated to contribute their best to the attainment of organizational goals. A systematic research plays a great role in the determination of the individual needs of the work force. That is why personnel policies have to be guided by research.

Job satisfaction enables workmen to derive a sense of accomplishment which leads to creativity. It creates a sense of belonging in the minds of the people at work. Management approaches to the people make a great impact on their accomplishments. These approaches may be classified into;

- i. Traditional approach;
- ii. Human relations approach;
- iii. Collective bargaining approach;
- iv. Motivation approach;
- v. Stick to rule approach; and
- vi. Neglecting approach.

Though the managerial approach to manpower is largely influenced by organizational philosophy, the individual and group behaviour of the workers towards the organisation plays no less an important role. In fact, organizational approach to the people is not an isolated event. Many factors are to be evaluated and taken into account by the personnel department prior to formulating the approach to their people. Any personnel policy or decision should be analysis based and result oriented.

The modern age is characterized by a low profile of worker logic and constantly declining trend in job satisfaction and work motivation. In order to counteract these trends, a number of techniques have been applied, which however, have proved to be futile. Goal setting, participative

management, job redesign, carrot-and-stick motivational tools”, etc. are prominent among them. While job redesign seeks to make the job more challenging, the goal setting approach ignores the job itself but makes it easier for workers to leave it. Nevertheless, the two approaches share a common interest in increasing workers’ discretion and self-control. Participative management attempts to ensure workers’ participation in the management of the enterprise and the carrot-and-stick approach makes the use of both the negative and positive means of motivation. Each approach is the result of research. But, in course of time, human behaviour changes and fresh researches become necessary to tackle fresh problems. The factory system, technological development, large scale operations, political changes, politically motivated trade unionism, recent trends in managerial outlook etc – all these bring about substantial changes in the worker’s behaviour. Obviously, personnel policies should be relevant to the changes and based on fresh empirical studies. There is, therefore, great scope for studies in personnel management.

Recent developments in the process of production necessitate job enrichment programmes, which consist of the following elements

1. Workers should be given full task:
2. New and more difficult tasks should be introduced :
3. The accountability of individual workers should be increased :
4. Elements of planning and co-ordination should be introduced :
5. Workers should be given additional authority and freedom:
6. Workers should be empowered to deal directly with clients or suppliers.

All these elements are incorporated only after a well-defined analysis. One of the most advanced forms of job redesign is self-managing work teams (autonomous work groups), which meet periodically to determine job assignments, scheduled work breaks, and lay down the rate of production.

In fact every decision or action of the work group is based on analysis and investigation. In this context, it can be observed that research is a part and parcel of effective management of all the functional areas. Production function extensively makes use of research and development for new product development, product differentiation, diversification, innovation, technological advancement, etc. leave aside the other areas : while the marketing man makes use of marketing

research, forecasting, market research, motivation research, evaluation research, promotion research and so on for formulating effective marketing strategies. The financial executive is able to guide the company to the path of success only with the help of research and analysis. Research and analysis play no less a role in the personal function also – say “selection, training and development, placement and induction, compensation, job evaluation, motivation, manpower planning, labour relations, labour welfare, industrial disputes and their settlements, collective bargaining – especially in the modern age of industrial disharmony and their settlements, collective bargaining – especially in the modern age of industrial disharmony and strife. Management information system, at the same time, is closely linked with research and analysis. To conclude, it can be observed that research is an essential aid not only to functional areas, but also to the general and top management. However, on a close observation, one would be able to realize that it would take many more years for the research to receive its due emphasis in the realm of Indian business.

Every action taken by every functional area manager should contribute to the overall objective of the organization. Hence the process of establishing organizational objectives and objectives for all its sub-units helps the organization in its basic direction. Any action of any of the functional areas should be measured by comparing it with the organisational objectives. The formulation of a hierarchy of contributions emphasizing the priorities would possibly facilitate a meaningful comparison. Any managerial research would be of interest, if it is able to aid in this process.

The objectives and activities of any organization are identified while forming the organization. The identification of any of the activities follows the formulation of objectives. For example, a business organization identifies its objectives first; the kind of product or service to be produced and distributed comes next. In some cases, it is the other way round. However, it is only after the identification of the product that the way of distribution, the production process, personnel policies, etc. are evolved. The functional areas obviously derive their objectives and activities from the overall objectives and goals of the organisation, while the contributions of the various functional areas may be aggregated as the sum total of the organizational output. A managerial research should necessarily take into account this inter-relationship between objectives and functions.

In a profit oriented organization, profitability studies are result oriented (A short run profitability measurement is easier than a long run profitability determination) But there are many organizations where profits may not be the sole or even a partial goal. In such situations, the researcher finds it difficult to measure or identify the effect of certain fulfilling objectives, policies and procedure in quantitative terms. Similar is the position in the case of quality consciousness. Since quality is intangible, it cannot be directly measured. However, there are certain characteristics variables and determinants in the functional areas like statistical quality control and various quality control techniques which can be subjected to research. In spite of their limitations, many managerial studies in the functional areas like production, marketing and finance have become very common nowadays.

## **QUESTIONS**

1. Explain Small Business Innovation Research (SBIR)
2. Write Short note On SPSS
3. Explain The Role Of Research In Marketing Department
4. Explain The Role Of Research In Production Department
5. Describe The Importance Of Research In Operation
6. Explain The Role Of Research In Finance Department
7. Explain The Role Of Research In Human Resource Management

\*\*\*\*\*

## Chapter :3

### RESEARCH DESIGN

**Features of good design, Types of Research Design, Basic Principles of experimental Design, Use of advanced technology in Research Design, Role of Research analyst.**

Research design is a purposeful scheme of action proposed to be carried out in a sequence during the process of research focusing on the management problem to be tackled. It must be a scheme for problem solving through proper analysis, for which a systematic arrangement of managerial efforts of investigate the problem is necessary. It defines the task of a researcher from identifying a managerial problem and problem area to report writing with the help of collection, tabulation, analyzing and interpretation of data. A research methodology is the arrangement of conditions for the collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure.

A research design indicates the observations that are to be made, how to make them and how to make quantitative analysis of such observations, which hypothesis should be tested, how the study should be carried out, what type of data should be collected and from which source, how to make an analysis, how to interpret and presents, how much money to spend , and within how much time the project should be completed.

A research methodology is the plan of action to be carried out in connection with a research project. It is only a guideline for the researcher to enable him to keep track of his actions and to know whether he was moving in the right direction in order to achieve his goal. The design may be a specific presentation of the various steps in the process of research. These include the selection of research problem, the presentation of the problem, the formulation of hypothesis, conceptual clarity, methodology, data collection, survey of literature and documentation, the testing of hypothesis, interpretation, bibliography, presentation and report writing.

#### **STAGES**

For the systematic presentation, the process of research may be classified under stages –

#### ➤ *Primary Stage*

- *Secondary Stage*
- *Tertiary Stage*

The Primary Stage includes :

- *Observation*
- *Interest*
- *Crystallization & identification of research problem*
- *Formulation of hypothesis*
- *Primary synopsis*
- *Conceptual clarity*
- *Documentation*
- *Preparation of Bibliography and*
- *Research designs*

The secondary stage includes :

- *Project Planning*
- *Project formulation*
- *Questionnaire preparation*
- *Investigation and data collection*
- *Preparation of final synopsis*
- *Compilation of data*
- *Classification*
- *Tabulation & presentation of data*
- *Experimentation*
- *Analysis*
- *Testing of Hypothesis and*

- ***Interpretation***

**The tertiary stage includes :**

- ***Report writing***
- ***Observation, suggestions and conclusions.***

“A research design is the arrangement of conditions for collections and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure.” In fact, the research design is the conceptual structure within which research is conducted; it constitutes the blueprint for the collections, measurement and analysis of data. As such the design includes an outline of what the researcher will do from writing the hypothesis and its operational implications to the final analysis of data. More explicitly, the design decisions happens to be in respects of:

- (i) *What is the study about?*
- (ii) *Why is the study being made ?*
- (iii) *Where will the study be carried out ?*
- (iv) *What type of data is required ?*
- (v) *Where can the required data be found ?*
- (vi) *What periods of time will the study include ?*
- (vii) *What will be the sample design ?*
- (viii) *What techniques of data collections will be used ?*
- (ix) *How will the data be analysed ?*
- (x) *In what style will the report be prepared ?*

Keeping in views the above stated design decisions; one may split the overall research design into the following parts:

- (a) ***the sampling design*** which deals with the method of selecting items to be observed for the given study;
- (b) ***the observational designs*** which relates to the conditions under the observations are to be made;

- (c) *the statistical design* which concerns with the questions of how many items are to be observed and how the information and data gathered are to be analysed ; and
- (d) *the operational design* which deals with the techniques by which the procedure specified in the samplings, statistical and observational designs can be carried out.

From what has been stated above , we can state the important features of a research designs as under:

- (i) It is a plans that specifies the sources and types of information relevant to the research problem.
- (ii) It is a strategy specifying which approach will be used for gathering and analyzing the data.
- (iii) It also includes the time and cost budgets since most studies are done under these two constraints.

In brief, research design must, at least, contain – (a) a clear statement of the research problem; (b) procedures and techniques to be used for gathering information; (c) the population to be studied; and (d) methods to be used in processing and analysing data.

**NEED FOR RESEARCH DESIGN :** 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 expenditure of effort, time and money. Just as for better, economical and attractive construction of a house, we need a blueprint (or what is commonly called the map of the house) well thought out and prepared by an expert architect, similarly we need a research design or a plan in advance of data collection and analysis for our research project. Research design stands for advance planning of the methods to be adopted for collecting the relevant data and the techniques to be used in their analysis, keeping in view the objective of the research and the availability of staff, time and money. Preparation of the research design should be done with great care as any error in it may upset the entire project. Research design, in fact, has a great bearing on the reliability of the results arrived at and as such constitutes the firm foundation of the entire edifice of the research work.

## **FEATURES OF A GOOD DESIGN**

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

- (i) The means of obtaining information;
- (ii) The availability and skills of the researcher and his staff, if any;
- (iii) The objective of the problem to be studied;
- (iv) The nature of the problem to be studied; and
- (v) The availability of time and money for the research work.

### **IMPORTANT CONCEPTS RELATING TO RESEARCH DESIGN**

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

**1. Dependent and independent variables:** A concept which can take on different quantitative values is called a variable. As such the concepts like weight, height, income are all examples of variables. Qualitative phenomena ( or the attributes ) are also quantified on the basis of the presence or absence of the concerning attribute(s).

**2. Extraneous variable:** Independent variables that are not related to the purpose of the study , but may affect the dependent variable are termed as extraneous variables. Suppose the researcher wants to test the hypothesis that there is a relationship between children gains in social studies achievement and their self – concept is an independent variable and social studies achievement is a dependent variable. Intelligence may as well affect the social studies achievement, but since it is not related to the purpose of the study undertaken by the researcher, it will be termed as an extraneous variable. Whatever effect is noticed on dependent variable as a result of extraneous variable(s) is technically described as an “experimental error”. A study must always be so designed that the effect upon the dependent variable is attributed entirely to the independent variable(s), and not to some extraneous variable or variables.

**3. Control:** One important characteristic of a good research design is to minimize the influence or effect of extraneous variable(s). The technical term “control” is used when we design the study minimizing the effects of extraneous independent variables. In experimental researches, the term “control” is used to refer to restrain experimental conditions.

**4. Confounded relationship:** When the dependent variable is not free from the influence of extraneous variable(s), the relationship between the dependent and independent variables is said to be confounded by an extraneous variable(s).

**5. Research hypothesis:** When a prediction or a hypothesized relationship is to be tested by scientific methods, it is termed as research hypothesis. The research hypothesis is a predictive statement that relates an independent variable to a dependent variable. Usually a research hypothesis must contain, at least, one dependent and one independent variable. Predictive statements which are not to be objectively verified or the relationships that are assumed but not to be tested, are not termed research hypothesis.

**6. Experimental and non – experimental hypothesis- testing research:** When the purpose of research is to test a research hypothesis, it is termed as hypothesis – testing research. It can be of the experimental design or of the non – experimental design. Research in which the independent variable is manipulated is termed “experimental hypothesis – testing research” and a research in which an independent variable is not manipulated hypothesis – testing research”.

**7. Experimental and control groups:** In an experimental hypothesis-testing research when a group is exposed to usual conditions, it is termed a ‘control group’, but when the group is exposed to some novel or special condition, it is termed an ‘experimental group’. In the above illustration, the Group A can be called a control group and the Group B an experimental group. If both groups A and B are exposed to special studies programmes, then both groups would be termed ‘experimental groups.’ It is possible to design studies which include only experimental groups or studies which include both experimental and control groups.

**8. Treatments :** The different conditions under which experimental and control groups are put are usually referred to as ‘treatments’. In the illustration taken above, the two treatments are the usual studies programme and the special studies programme. Similarly, if we want to determine through an experiment 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.

**9. Experiment :** The process of examining the truth of a statistical hypothesis, relating to some research problem, is known as an experiment. For example, we can conduct an experiment to examine the usefulness of a certain newly developed drug. Experiments can be of two types viz., absolute experiment and comparative experiment. If we want to determine the impact of a fertilizer

on the yield of a crop, it is a case of absolute experiment; but if we want to determine the impact of one fertilizer as compared to the impact of some other fertilizer, our experiment then will be termed as a comparative experiment. Often, we undertake comparative experiments when we talk of designs of experiments.

**10. Experimental unit(s)** : The pre-determined plots or the blocks, where different treatments are used, are known as experimental units. Such experimental units must be selected (defined) very carefully.

**DIFFERENT RESEARCH DESIGNS** : Different research designs can be conveniently described if we categorize them as :

- (1) Research design in case of exploratory research studies;
- (2) Research design in case of descriptive and diagnostic research studies, and
- (3) Research design in case of hypothesis-testing research studies.

We take up each category separately.

**1. Research design in case of exploratory research studies:** Exploratory research studies are also termed as formative research studies. The main purpose of such studies is that of formulating a problem for more precise investigation or of developing the working hypotheses from an operational point of view. The major emphasis in such studies is on the discovery of ideas and insights. As such the research design appropriate for such studies must be flexible enough to provide opportunity for considering different aspects of a problem under study. Inbuilt flexibility in research design is needed because the research problem, broadly defined initially, is transformed into one with more precise meaning in exploratory studies, which fact may necessitate changes in the research procedure for gathering relevant data. Generally, the following three methods in the context of research design for such studies are talked about:

- (a) The survey of concerning literature;
- (b) The experience survey and
- (c) The analysis of ‘insight-stimulating’ examples.

*The survey of concerning literature* happens to be the most simple and fruitful method of formulating precisely the research problem or developing hypothesis. Hypotheses stated by earlier

workers may be reviewed and their usefulness be evaluated as a basis for further research. It may also be considered whether the already stated hypotheses suggest new hypothesis. In this way the researcher should review and build upon the work already done by others, but in cases where hypotheses have not yet been formulated, his task is to review the available material for deriving the relevant hypotheses from it.

Thus, in an exploratory or formative research study which merely leads to insights or hypotheses, whatever method or research design outlined above is adopted, the only thing essential is that it must continue to remain flexible so that many different facets of a problem may be considered as and when they arise and come to the notice of the researcher.

**2. Research design in case of descriptive and diagnostic research studies:** Descriptive research studies are those studies which are concerned with describing the characteristics of a particular individual, or of a group, whereas diagnostic research studies determine the frequency with which something occurs or its association with something else. The studies concerning whether certain variables are associated are examples of diagnostic research studies. As against this, studies concerned with specific predictions, with narration of facts and characteristics concerning individual, group or situation are all examples of descriptive research studies. Most of the social research comes under this category. From the point of view of the research design, the descriptive as well as diagnostic studies share common requirements and as such we may group together these two types of research studies. In descriptive as well as in diagnostic studies, the researcher must be able to define clearly, what he wants to measure and must find adequate methods for measuring it along with a clear cut definition of 'population' he wants to study. Since the aim is to obtain complete and accurate information in the said studies, the procedure to be used must be carefully planned. The research design must make enough provision for protection against bias and must maximize reliability, with due concern for the economical completion of the research study. The design in such studies must be rigid and not flexible and must focus attention on the following:

- (a) Formulating the objective of the study (what the study is about and why is it being made?)
- (b) Designing the methods of data collection (what techniques of gathering data will be adopted?)
- (c) Selecting the sample (how much material will be needed?)

- (d) Collecting the data (where can the required data be found and with what Time period should the data be related?)
- (e) Processing and analyzing the data.
- (F) Reporting the findings.

In a descriptive/diagnostic study the first step is to specify the objectives with sufficient precision to ensure that the data collected are relevant. If this is not done carefully, the study may not provide the desired information.

Then comes the question of selecting the methods by which the data are to be obtained. In other words techniques for collecting the information must be devised. Several methods (viz., observation, questionnaires, interviewing, examination of records, etc.), with their merits and limitations, are available for the purpose and the researcher may use one or more of these methods. While designing data-collection procedure, adequate safeguards against bias and unreliability must be ensured. Whichever method is selected, questions must be well examined and be made unambiguous; interviewers must be instructed not to express their own opinion, observers must be trained so that they uniformly record a given item of behaviour. It is always desirable to pre-test the data collection instruments before they are finally used for the study purposes. In other words, we can say that ‘*structured instruments*’ are used in such studies.

In most of the descriptive/diagnostic studies the researcher takes out sample(s) and then wishes to make statements about the population on the basis of the sample analysis or analyses. More often than not, sample has to be designed. “As data are collected, they should be examined for completeness, comprehensibility, consistency and reliability”

The data collected must be processed and analysed. This includes steps like coding the interview replies, observations, etc.; tabulating the data; and performing several statistical computations. To the extent possible, the processing and analyzing procedure should be planned in detail before actual work is started. This will prove economical in the sense that the researcher may avoid unnecessary labour such as preparing tables for which he later finds he has no use or on the other hand, re-doing some tables because he failed to include relevant data. Coding should be done carefully to avoid error in coding and for this purpose the reliability of coders needs to be checked. Similarly, the accuracy of tabulation may be checked by having a sample of the tables

re-done. In case of mechanical tabulation the material (i.e. the collected data or information) must be entered on appropriate cards which is usually done by punching holes corresponding to a given code. The accuracy of punching is to be checked and ensured. Finally, statistical computations are needed and as such averages percentages and various coefficients must be worked out. Probability and sampling analysis may as well be used. The appropriate statistical operations, along with the use of appropriate tests of significance should be carried out to safeguard the drawing of conclusions concerning the study.

Last of all comes the question of reporting the findings. This is the task of communicating the findings to others and the researcher must do it in an efficient manner. The layout of the report needs to be well planned so that all things relating to the research study may be well presented in simple and effective style.

Thus, the research design in case of descriptive/diagnostic studies is a comparative design throwing light on all points narrated above and must be prepared keeping in view the objective(s) of the study and resources available. However, it must ensure the minimization of bias and maximization of reliability of the evidence collected. The said design can be appropriately referred to as a survey design since it takes into account all the steps involved in a survey concerning a phenomenon to be studied.

The difference between research designs in respect of the above two types of research studies can be conveniently summarized in tabular form as under :

**Table**

Research Design	Type of Study	
	Exploratory of Formulative	Descriptive/Diagnostic
Overall Design	Flexible design (design must provide opportunity for considering different aspects of the problem)	Rigid design (design must make enough provision for protection against bias and must maximize reliability)
(i) Sampling design	Non-probability sampling design (purposive or judgement sampling)	Probability sampling design (random sampling)
(ii) Statistical Design	No pre-planned design for analysis	Pre-planned design for analysis
(iii) Observational design	Unstructured instruments for collection of data	Structured or well thought out instruments for collection of data
(iv) Operational design	No fixed decisions about the operational procedures	Advanced decisions about operational procedures.

**3. Research design in case of hypothesis-testing research studies :** Hypothesis-testing research studies (generally known as experimental studies) are those where the researcher tests the hypotheses of casual relationships between variables. Such studies require procedures that will not only reduce bias and increase reliability, but will permit drawing inferences about causality. Usually experiments meet this requirement. Hence, when we talk of research design in such studies, we often mean the design of experiments.

Professor R.A.Fisher's name is associated with experimental designs. Beginning of such designs was made by him when he was working at Rothamsted Experimental Station (Centre for Agricultural Research in England). As such the study of experimental designs has its origin in agricultural research. Professor Fisher found that by dividing agricultural fields or plots into different blocks and then by conducting experiments in each of these blocks, whatever information is collected and inferences drawn from them, happens to be more reliable. This fact inspired him to develop certain experimental designs for testing hypotheses concerning scientific investigations.

Today, the experimental designs are being used in researches relating to phenomena of several disciplines. Since experimental designs originated in the context of agricultural operations, we still use, though in a technical sense, several terms of agriculture (such as treatment, yield, plot, block etc.) in experimental designs.

**BASIC PRINCIPLES OF EXPERIMENTAL DESIGNS** : Professor Fisher has enumerated three principles of experimental designs : (1) the Principle of Replication; (2) the Principle of Randomization; and the (3) Principle of Local Control.

According to the Principle of Replication, the experiment should be repeated more than once. Thus, each treatment is applied in many experimental units instead of one. By doing so the statistical accuracy of the experiments is increased. For example, suppose we are to examine the effect of two varieties of rice. For this purpose we may divide the field into two parts and grow one variety in one part and the other variety in the other part. We can then compare the yield of the two parts and draw conclusion on that basis. But if we are to apply the principle of replication to this experiment, then we first divide the field into several parts, grow one variety in half of these parts and the other variety in the remaining parts. We can then collect the data of yield of the two varieties and draw conclusion by comparing the same. The result so obtained will be more reliable in comparison to the conclusion we draw without applying the principle of replication. The entire experiment can even be repeated several times for better results. Conceptually replication does not present any difficulty, but computationally it does. For example, if an experiment requiring a two-way analysis of variance is replicated, it will then require a three-way analysis of variance since replication itself may be a source of variation in the data. However, it should be remembered that replication is introduced in order to increase the precision of a study; that is to say, to increase the accuracy with which the main effects and interactions can be estimated.

The *Principle of Randomization* provides protection, when we conduct an experiment, against the effect of extraneous factors by randomization. In other words, this principle indicates that we should design or plan the experiment in such a way that the variations caused by extraneous factors can all be combined under the general heading of “chance.” For instance, if we grow one variety of rice, say, in the first half of the parts of a field and the other variety is grown in the other half, then it is just possible that the soil fertility may be different in the first half in comparison to the other half. If this is so, our results would not be realistic. In such a situation, we may assign the

variety of rice to be grown in different parts of the field on the basis of some random sampling technique i.e., we may apply randomization principle and protect ourselves against the effects of the extraneous factors (soil fertility differences in the given case). As such, through the application of the principle of randomization, we can have a better estimate of the experimental error.

The *Principle of Local Control* is another important principle of experimental designs. Under it the extraneous factor, the known source of variability, is made to vary deliberately over as wide a range as necessary and these needs to be done in such a way that the variability it causes can be measured and hence eliminated from the experimental error. This means that we should plan the experiment in a manner that we can perform a two-way analysis of variance, in which the total variability of the data is divided into three components attributed to treatments (varieties of rice in our case), the extraneous factor (soil fertility in our case) and experimental error.” In other words, according to the principle of local control, we first divide the field into several homogeneous parts, known as blocks, and then each such block is divided into parts equal to the number of treatments. Then the treatments are randomly assigned to these parts of a block. Dividing the field into several homogenous parts is known as ‘blocking’. In general, blocks are the levels at which we hold an extraneous factor fixed, so that we can measure its contribution to the total variability of the data by means of a two-way analysis of variance. In brief, through the principle of local control we can eliminate the variability due to extraneous factor(s) from the experimental error.

**Important Experimental Designs :** Experimental design refers to the framework or structure of an experiment and as such there are several experimental designs. We can classify experimental designs into two broad categories, viz., informal experimental designs and formal experimental designs. Informal experimental designs are those designs that normally use a less sophisticated form of analysis based on differences in magnitudes, whereas formal experimental designs offer relatively more control and use precise statistical procedures for analysis. Important experiment designs are as follows :

- (a) Informal experimental designs :
  - (i) Before-and-after without control design.
  - (ii) After-only with control design.

(iii) Before-and-after with control design.

(b) Formal experimental designs :

(i) Completely randomized design (C.R.Design).

(ii) Randomized block design (R.B.Design).

(iii) Latin square design (L.S.Design).

(iv) Factorial designs.

We may briefly deal with each of the above stated informal as well as formal experimental designs.

**1. Before-and-after without control design :** In such a design a single test group or area is selected and the dependent variable is measured before the introduction of the treatment. The treatment is then introduced and the dependent variable is measured again after the treatment has been introduced. The effect of the treatment would be equal to the level of the phenomenon after the treatment minus the level of the phenomenon before the treatment. The main difficulty of such a design is that with the passage of time considerable extraneous variations may be there in its treatment effect.

**2. After-only with control design :** In this design two groups or areas (test area and control area) are selected and the treatment is introduced into the test area only. The dependent variable is then measured in both the areas at the same time. Treatment impact is assessed by subtracting the value of the dependent variable in the control area from its value in the test area. This can be exhibited in the following form. The basic assumption in such a design is that the two areas are identical with respect to their behaviour towards the phenomenon considered. If this assumption is not true, there is the possibility of extraneous variation entering into the treatment effect. However, data can be collected in such a design without the introduction of problems with the passage of time. In this respect the design is superior to before-and-after without control design.

**3. Before-and-after with control design:** In this design two areas are selected and the dependent variable is measured in both the areas for an identical time period before the treatment. The treatment is then introduced into the test area only, and the dependent variable is measured in both an identical time-period after introduction of the treatment. The treatment effect is determined by subtracting the change in the dependent variable in the control area from the change in the dependent variable in test area. This design is superior to the above two designs for the simple

reason that it avoids extraneous variation resulting both from the passage of time and from non-comparability of the test and control areas. But at times, due to lack of historical data, time or a comparable control area, we should prefer to select one of the first two informal designs stated above.

**4. Completely randomized design (C.R.design) :** Involves only two principles viz., the principle of replication and the principle of randomization of experimental designs. It is the simplest possible design and its procedure of analysis is also easier. The essential characteristic of the design is that subjects are randomly assigned to experimental treatments (or vice-versa).

(i) **Two-group simple randomized design:** In a two-group randomized design, first of all the population is defined and then from the population a sample is selected randomly. Further, requirement of this design is that items, after being selected randomly from the population, be randomly assigned to the experimental and control groups (Such random assignment of items to two groups is technically described as principle of randomization). Thus, this design yields two groups as representatives of the population. Since in the sample randomized design the elements constituting the sample are randomly drawn from the same population and randomly assigned to the experimental and control groups, it becomes possible to draw conclusions on the basis of samples applicable for the population. The two groups (experimental and control groups) of such a design are given different treatments of the independent variable. This design of experiment is quite common in research studies concerning behavioural sciences.

(ii) **Random replications design:** The limitation of the two-group randomized design is usually eliminated within the random replications design. But in a random replications design, the effect of such differences are minimized (or reduced) by providing a number of repetitions for each treatment. Each repetition is technically called a 'replication'. Random replication design serves two purposes viz., it provides controls for the differential effects of the extraneous independent variables and secondly, it randomizes any individual differences among those conducting the treatments.

**5. Randomized block design (R.B.design)** is an improvement over the C.R.design. In the R.B.design the principle of local control can be applied along with the other two principles of experimental designs. In the R.B.design, subjects are first divided into groups, known as blocks, such that within each group the subjects are relatively homogeneous in respect to some selected

variable. The variable selected for grouping the subjects is one that is believed to be related to the measures to be obtained in respect of the dependent variable. The main feature of the R.B.design is that in this each treatment appears the same number of times in each block. The R.B.design is analysed by the two-way analysis of variance (two-way ANOVA) technique.

6. **Latin square design (L.S.design)** is an experimental design very frequently used in agricultural research. The conditions under which agricultural investigations are carried out are different from those in other studies for nature plays an important role in agriculture. For instance, an experiment has to be made through which the effects of five different varieties of fertilizers on the yield of a certain crop, say wheat, it to be judged. In such a case the varying fertility of the soil in different blocks in which the experiment has to be performed must be taken into consideration; otherwise the results obtained may not be very dependable because the output happens to be the effect not only of fertilizers, but it may also be the effect of fertility of soil. Similarly, there may be impact of varying seeds on the yield. To overcome such difficulties, the L.S. design is used when there are two major extraneous factors such as the varying soil fertility and varying seeds.

7. **Factorial designs:** Factorial designs are used in experiments where the effects of varying more than one factor are to be determined. They are specially important in several economic and social phenomena where usually a large number of factors affect a particular problem. Factorial designs can be of two types (i) simple factorial designs and (ii) complex factorial designs.

Research analyst support evaluation, research ,analysis and writing for projects that inform public policy in child welfare, health and other human services programs.

Specific task are likely to include:

- Develop policy briefs and other short pieces from new or existing research.
- Assist in drafting reports and grant proposals, including
- Researching background on relevant policy debate, legislative proposals and academic literature.
- Translating research finding for policy audiences.

- Interpret and analyze quantitative analyses and outcomes data for practice and policy audience.
- Collect data on program participants and impacts through interviews and focus groups.
- Conduct site visits.
- Work with government agencies and other organizations to access data and program information.
- Deliver technical assistance to help clients understand policy briefs and research reports.
- Staff workgroups and committees which plan program improvement, develop and implement program tools and direct policy.
- Help develop training materials on how to use data to manage programs.
- Project coordination.
- Manage communications among diverse group of clients and constitutes at state and local levels.
- Assist in planning workgroups, committees and trainings.
- Research analyst work in project team, but may take lead responsibility for subtasks.

In nutshell a research analyst plays the following multidimensional role.

- **Change Agent**-As research is defined as defining and redefine of a problem and its results in some change. Hence the research analyst plays the role of a change agent.
- **Investigator**-A research analyst investigates various facts and figures, in order to collect, the data and then to derive the conclusion out of it. Hence he plays the role of an investigator.
- **Monitor**-Research is an individual as well as a team activity. In case of a group research, the research analyst monitors and keeps a track of the proceedings of the research work. Hence he plays the role of a monitor.

- **Psychologist-** A research analyst has to play the role of psychologist, as he has to handle a lot of people in his team, and he has to handle the problems, not only of the research team, but also the problems of respondents.
- **Motivator-** Motivator is a person who stimulates his group mates, in order to handle the adverse situation and the various constraints during the work. A research analyst plays this role .

## QUESTIONS

1. What is research design ?what is the nature and utility of a research design?
2. What are the types of research design ?
3. What do you mean by research design or research plan? What are the contents of it?
4. Why should a research design prepared? How it is prepared?
5. What are the contents of a research design?
6. Explain the role of research analyst.
7. Explain the functions of research analyst.

## **Chapter :4**

### **SAMPLING DESIGN**

#### **Steps in sample design, characteristics of a good sample design, Probability & Non-Probability sampling.**

The first and foremost task in carrying out a survey is to select the sample. Sample selection is undertaken for practical impossibility to survey the population. By applying rationality in selection of samples, we generalize the findings of our research. There are different types of sampling. We may categorize those in three major heads as follows:

1. Random Sampling
2. Purposive Sampling
3. Stratified Sampling

**OBJECTIVES OF SAMPLING :** Main objectives are as follows :

- 1 To obtain information about the population on the basis of sample drawn from such population
- 2 To set up limits of accuracy of the estimates of the population parameters completed on the basis of sample statistic.
- 3 To test the significance about the population characteristics on the basis of sample statistic.

Apart from the above sampling procedures, there are other types of sampling like:

- Quota sampling (a special type of stratified sampling).
- Multi-stage sampling (where samples are selected from a very large area).
- Convenience sampling (where population is not clearly defined and complete source of list is not available).
- Self selected sampling, etc.

After deciding over the samples to be surveyed, the next task is to go ahead with the survey matter.

Survey may be carried out either by directly interviewing the samples or by sending questionnaire to the samples or by mere observation of the characteristics of samples. We have discussed these in earlier chapter.

- **The Sampling Process**

There are seven steps involved in this process.

**Step 1 : Define the population:** It is the aggregate of all elements, usually defined prior to the selection of the sample. The population is said to be completely defined if atleast the following terms are specified:

- i. Elements
- ii. Sampling Units
- iii. Extent
- iv. Time

For example, for monitoring the sales of our product, the population might be

- Element – Our Product.
- Sampling Units – Retail outlets, super markets.
- Extent – Mumbai.
- Time – December 1-31<sup>st</sup>, 2009

**Step 2 : Identify the sampling frame:** The sampling frame should be so selected which consists of almost all the sampling units. Though it is not possible to have one-to-one correspondence between frame units and sampling units, however, we should choose a sampling frame which yields unbiased estimates with a variance as low as possible. Popularly known sampling frame are :

Census reports, electoral registers, lists of member units of trade and industry associations, lists of members of professional bodies, lists of dwelling units maintained by local bodies, returns from an earlier survey and large scale maps etc.

**Step 3: Specify the sampling unit:** The sampling unit is the basic unit containing the elements of the target population.

**Step 4: Specify the sampling method:** The sampling method indicates how the sample units are selected. The most important decision in this regard is to determine which of the two-probability or non-probability samples is to be chosen.

**Step 5: Determine the sample size (n):** The decision about the number of elements to be chosen, i.e., number of observations in each sample (n) of the target population.

**Step 6: Specify the sampling plan:** This means that one should indicate, how decisions made so far are to be implemented. All expected pertinent issues in a sampling survey must be answered by the sampling plan.

**Step 7: Select the sample:** This is the final step in the sampling process. A good deal of field work and office work is introduced in the actual selection of the sample elements. However it depends mainly upon the sampling plan and the sample size required.

- **Advantages of Sampling**

There are various advantages of sampling, as given below:

- i. The ideal solution to know the true or actual values of the different parameters of the population would be to take into account the entire population. However, that is not feasible due to cost, time, labour and other constraints, hence sampling is more economical.
- ii. As the magnitude of operations involved in a sample survey is small, both the execution of the field work and the analysis of results can be carried out much faster and hence is very less time consuming. On the other hand, if we have to gather information about the whole population, in certain cases, we may not be able to have the information collected timely and the entire study may become redundant.
- iii. In case of destructive testing, for the entire population, we may not be left with anything after such an enumeration, i.e., testing the life of a tube or bulb.
- iv. Relatively very small staff is required for gathering information, analyzing the same and preparation of the report.
- v. A researcher can collect more detailed information in much less time than otherwise by possible in a census survey. Moreover, we can afford to have a few specialists

for collection of specialized type of information which otherwise is not possible for a census study because of cost and non-availability of specialists.

- vi. As the scale of operations involved in a sample survey is small, the quality of the interviewing, supervision and other related activities can be better than the quality in a census survey.
- vii. In many cases, sampling provides adequate information needed for the purpose and sufficiently reliable. Also, the sampling techniques make it possible to quantify the magnitude of possible error.

### **Disadvantage of Sampling Technique**

Even though sampling has lot of advantages, it is no free from disadvantage. The disadvantage of sampling technique is as follows:

- i. **Less accuracy:** in comparison to census technique, the conclusion derived using sample techniques are more liable to error. Therefore, sampling techniques are less accurate than the census techniques.
- ii. **Changeability of units:** if the units in the fields of survey are liable to change or if they are not homogeneous, the use of sampling techniques will be very hazardous. These techniques are not scientific to extend the conclusions derived from one set of sample to other sets which are dissimilar or are changeable.
- iii. **Misleading conclusions:** if the sample selection is not done scientifically and due care is not taken, then the conclusion derived from sampling technique, if extended to entire population, will become misleading. For example, if a researcher while studying the expenditure behaviour of MBA students, selects rich students from the entire population, then the conclusion drawn will be applied to all the MBA student which will be wrong.
- iv. **Need for special knowledge:** the sampling technique can be successful only if a competent and able scientist makes the selection. If this is done by non-skilled researcher, the sample selection can be incorrect leading to sampling error.

### **Characteristics of Ideal Sampling**

To adopt an appropriate and unbiased sampling techniques, a researcher has to maintain certain qualities of sampling techniques. They are as follows:

**Representativeness:** An ideal sample is the sample which represents the characteristics of entire population. Thus the selection procedure for sampling should be such that the sample selected has all those qualities and features which the entire population possess.

**Independence:** The second quality required in sampling techniques is that all possible samples which can be selected should be independent of each other. This helps to make an unbiased selection of samples where selection of one unit of population does not depend upon that of another unit.

**Adequacy:** The number of units in a sample, that is sample size, selected from the population should not be too less as this would result in failure to capture the diversity of population. At the same time, it should not be too large. It should be just sufficient to enable a researcher to derive a correct conclusion about population based on his sample.

**Homogeneity:** To make a sample scientific, the element or unit selected within the sample should be identical with another element or unit of that sample.

### **Types of Sampling**

There are basically two types of sampling methods:

a. Probability Sampling Method

The probability or chance of every unit in the population being included in the sample is known. Selection of the specific units in the sample depends entirely on chance. Types of Probability sampling are:

- *Simple random sampling*
- *Systematic sampling*
- *Stratified random sampling*
- *Cluster sampling*

b. Non-probability Sampling Method.

The probability of inclusion of any unit (of population) in a sample is not known. The selection of units within a sample involves human judgment rather than pure chance. The maximum information available “per rupee” which can be determined from a probability sample is not possible in this case and moreover, the degree of accuracy is not known. Although probability sampling is scientific and accurate, however, because of convenience and economy, the non-probability samples are preferred. Many times, samples are selected by interviewers “at

random” meaning that the actual sample selection is left to the choice of the researcher, such sample are non-probability samples and not probability samples.

- *Judgement sampling*
- *Convenience sampling*
- *Quota sampling*
- *Snowball sampling*

#### ❖ **Probability Sampling Methods**

The major sampling methods under probability sampling are:

- Simple Random Sampling
- Systematic Sampling
- Stratified Sampling
- Cluster Sampling

- **Simple Random Sampling**

Simple Random Sampling is the simplest type of sampling, in which we draw a sample of size (n) in such a way that each of the ‘N’ members of the population has the same chance of being included in the sample. A sample selected in this way is called a simple random sample.

#### **Selection of Random Samples**

One way of drawing a simple random sample is to number every individual, put the numbers on slips of paper and draw lots. But such a procedure is impracticable in most situations. A more convenient method is to use a table of random digits. Such tables contain lists of digits so chosen that each digit between 0 to 9 has an equal chance of appearing at a given spot in a single column and each two digits between 00 to 99 has the same chance of appearing at a given spot in a double column and so on. Printed tables of random numbers are used in practice. For example, if we want to select ten items from the item produced during a shift from an automatic machine, which normally produces 500 items in a shift, we need to select 10 random numbers between 001 and 500. The selected 10 numbers arranged in ascending order will give the serial number of the items to be included in the sample.

## Simple Random Sample in Practice

The use of random sampling is made by researchers for the following situations:

- When a small sample is needed from the list of sample frame (list of universe items).
- The cost per interview is practically independent of the location of the sample item.
- Other than a list of items, no other information is available.
- In simple random sampling, the sample mean provides an unbiased estimate of the universe mean. However, the use is severely limited by the following factors.

**Cost:** Sampling frame is required for a simple random sample. In most situation, it is very difficult to have a frame, if not impossible and moreover it is very time consuming and hence is uneconomical also.

**Statistical Efficiency:** One sample designing is said to be statistically more efficient than another when the frame is of a small size and a smaller standard error is obtained. Most large populations are not homogeneous but can be broken down into more homogeneous units.

In such conditions, one can have sampling design such as stratified sampling which is statistically more efficient. Similarly, the use of cluster sampling whenever we can pick up members from geographically closer areas, reduces the cost involved.

**Supervision:** The problem of selecting a sample with the help of random numbers etc., though appears to be very simple but in practice it is not so. Suppose, we want to select a sample size of 2,5000 from a list of 2 million, it's a difficult job and errors are bound to be there. Moreover, the cost of supervision is also high. As compared to this, systematic random sampling is much easier.

- **Systematic Sampling**

The systematic sampling also employs the principle of Random Sampling. However, in this method of sample, selection of a unit depends upon the selection of a preceding unit in contrast to simple Random Sampling. Where the selection of a unit is independent of each other. Systematic Random Sampling in this sense is called quasi-random-sampling.

### Advantages

One of the biggest advantage of this method is its simplicity of drawing samples. Except for population with periodic behaviour, systematic sampling variances are often somewhat smaller than those for alternative procedures.

### Disadvantages:

If sampling interval is related to a periodic coding ordering of the universe, increased variability may be introduced. Systematic sampling should be used in practice only when the researcher is sufficiently acquainted with the data. Only then he can be able to demonstrate the periodicities do not exist or sampling interval is not a multiple or submultiples of the period.

- **Stratified Random Sampling**

Another useful type of sampling procedure is called stratified random sampling. In this procedure, the members of the population are first assigned to strata or groups, on the basis of some characteristic and a simple random sample is drawn from each stratum. The individuals in all the samples taken together constitute the sample from the population as a whole, viz.

Stratum	Number in stratum	Number in sample
1	$N_1$	$n_1$
2	$N_2$	$n_2$
3	$N_3$	$n_3$
4	$N_4$	$n_4$
5	$N_5$	$n_5$
6	$N_6$	$n_6$
7	$N_7$	$n_7$
K	$N_k$	$n_k$
Total	$N$	$n$

If the component sample sizes  $n_1, n_2, \dots, n_k$  are so chose that

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = \dots = \frac{n_k}{N_k} = \frac{n}{N}$$

In this way we do proportional stratified random sampling.

Stratification does not mean absence of randomness. All it means is that the population is first divided into certain strata that are mutually exclusive and collectively exhaustive.

Populations divided into three (3) strata  $s_1, s_2, s_3$ .

A stratum, as is clear, is a sub-population which is more homogenous than the complete population.

The members of a stratum are similar to each other.

### Major Issues

There are three major issues involved in stratified sampling.

- i. **Bases of stratification:** Depends upon the variable being studied. Quite often it is desired to have more than one variable. In Marketing Research, usually demographic characteristics such as age, sex or income and geographical distribution such as rural-urban break-up, break-up by region, state or city form the basis.
- ii. **Number of strata:** Since stratification would enhance the cost of the survey, one would weigh the benefits resulting from it vs. the cost involved. As a rule of thumb, not more than 6 strata should be used for a single overall estimate.
- iii. **Sample sizes within strata:** Sample size within strata depends upon the budget and the cost per observation. Suppose in the above example, if the budget is Rs 50,000 and the cost per observation is Rs 250, then the sampling size =  $50,000/250 = 200$  and not 450. However, this should be allocated among the various strata. Researchers can use either a proportional or disproportional sample.

### Stratified Sampling in Practice

The main reasons for using stratified sampling for managerial applications are:

- i. It can obtain information about different parts of the universe, i.e., it allows to draw separate conclusion for each stratum.
- ii. It often provides universe estimates of greater precision than other methods of random sampling say simple random sampling.

### **Advantages**

It enables the researcher to make a comparison of properties of the strata as well as to estimate the popular characteristics. One of the advantages of this method is that the investigator has greater control over the selection of the sample. The superiority of stratified sampling over random sampling is that the possibility of all groups of population being selected in the sample is high here. Due to the satisfaction method, the possibility of sample representing the various sections of population is very high and this can be achieved with fewer items.

### **Disadvantages**

One of the biggest disadvantages of this method is that the sampling can be highly biased if the stratification is based on some biases or is non-scientific. Another disadvantage is that for a researcher is not easy to attain proportion becomes particularly difficult when there is a wide variance in the sizes of different strata. Apart from this if the strata are not well defined, it may not be easy to decide in which unit of stratum a particular case is to be placed.

- **Cluster Sampling**

In the probability sampling methods, we have seen that each item in the sample is chosen one at a time from the complete list of universe elements. However, it would be more expedient to select entire groups or clusters at random. Let us take, for example, a residential colony comprising 15 Blocks : A to O. Let us treat each block as a cluster and then select 3 (say) blocks (cluster) out of 15 blocks at random and then collect information from all families residing in these 3 blocks (clusters).

### **Cluster Sampling in Practice**

This method is usually convenient for collection of data as a cluster is ideally a mini-population and has all the features of the population. Clusters are Heterogeneous within themselves and not like the homogeneous strata. Hence, collection of data would be far easy as compared to other method. Moreover, apart from cost considerations, such a cluster sample would be desirable in the absence of a suitable sampling frame. Frames, for cluster sampling, are needed for the selected cluster only and hence reduce the cost of developing a frame as compared for simple random sampling or stratified random sampling.

## Use of Cluster Sampling in a National Survey

In a National survey, first we'll select a few districts in the whole country, (these districts) would act as cluster and then we would take stratified sampling/simple random sampling so that the survey can be completed with less cost and more accuracy.

- **Multistage and Multiphase Sampling**

Multistage sampling, as the name implies, means that the selection of units is done in more than one stages. The number of stages in a multistage sampling is based on convenience and the availability of suitable frames at different stages. In case of a national survey, It is easier to understand that this can involve the following four stages:

1 <sup>st</sup> stage	Districts
2 <sup>nd</sup> stage	Cities
3 <sup>rd</sup> stage	Wards or localities
4 <sup>th</sup> stage	Households

### Example

Suppose, we want to have 7500 households from all over the country. In such a case, from the first stage, District, say 30 districts out of 600 are selected from all over the country.

**II stage** – Cities: Suppose 5 cities are selected out of each 30 districts; and

**III stage** – Wards/Localities: say 10 wards/localities are selected from each city

**IV stage** – Households: 50 households are selected from each ward/locality.

In stage I, we can employ stratified sampling

In stage I, we can employ stratified sampling

In stage II, we can use cluster sampling

In stage III, we can have simple random sampling.

Thus, the use of various methods shall give individually contribute towards accuracy, cost, time, etc. this leads us to conclude that multistage sampling leads to saving of time, labour and money.

Apart from this wherever an appropriate frame is not available, the use of multistage sampling has universal appeal.

#### ❖ **Non-probability Sampling Methods**

Though the probability samples give an unbiased sample and the parameters used for the study can be tested for a given confidence internally, still non-probability sampling finds frequent use in many situations because of difficult conditions such as frame (list of all sampling units), time and cost involved. Major Non-probability sampling methods are discussed below.

- **Convenience Sampling** : As the name implies, the selection of the sample is left to the researcher who is to select the sample. The researcher normally interviews persons in groups at some retail outlet, supermarket or may stand at a prominent point and interview the persons who happen to be there. This type of sampling is also called ‘accidental sampling’ as the respondents in the sample are included merely because of their presence on the spot. The data collection and sample cost is minimum in this case. However, the method suffers greatly from the quality, i.e., accuracy point of view which can in no way be determined. However, this type of sampling is more suitable in ‘exploratory research’ where focus is on getting new ideas/insights into a given problem.
- **Judgement Sampling**: In judgement sampling, the judgement or opinion of some experts forms the basis of the sampling method. It is expected that these samples would be better as the experts are supposed to know the population. However, as the use of randomness is not there and moreover there is no way to find the accuracy of the samples, hence the method has its limitations and is used mainly for situations requiring extremely small size of samples, i.e., use of rare events, members having extreme positions, etc.
- **Snowball sampling**: The sampling technique involves the selection of additional respondents based on the referral of initial respondents is known as snowball sampling. This sampling technique is often used when the population is rarely found and not easily accessible. Hence the sampling depends upon the chain system of referrals. It means if the first respondent is found, then the investigator can ask him to provide contact details of another individuals who fall in the same population.

**Advantages:** One of the important elements of this sampling technique is that it is highly useful in case of rare population. A researcher does not have to make a lot of efforts to find the units of population of his own.

**Disadvantages:** it is at times very difficult to get the track of respondents. Moreover, since respondents are referred names, they may not be cross-sectional or representing the entire population.

- **Quota Sampling :** This is the most frequently used non-probability sampling method and is used when. Employing Stratification: i.e., age, sex, income, family etc. More often compound stratification is used in groups with sex.

#### **Quota Sample in Practice**

- It is economical as travelling costs are reduced.
- It is easy to administer.
- When field work is to be done quickly, this method provides the biggest advantage.
- No sampling frame is required.

However, since the method is not based on random selection, it is not possible to ascertain the accuracy being achieved. Moreover, the sample may not be a representative sample. Also, the quality may suffer if skilled interviewers are not employed. In many real life situations, the use of cluster sampling method and that of convenience sampling is being made mainly because the cost per observation is much less as compared to other method. Sophisticated research techniques and methods ensuring greater accuracy are not finding much use.

➤ *Write short note on sampling process.*

- **The Sampling Process**

There are seven steps involved in this process.

**Step 1 : Define the population:** It is the aggregate of all elements, usually defined prior to the selection of the sample. The population is said to be completely defined if at least the following terms are specified:

- v. Elements
- vi. Sampling Units

- vii. Extent
- viii. Time

For example, for monitoring the sales of our product, the population might be

- Element – Our Product.
- Sampling Units – Retail outlets, super markets.
- Extent – Mumbai.
- Time – December 1-31<sup>st</sup>, 2009

**Step 2 : Identify the sampling frame:** The sampling frame should be so selected which consists of almost all the sampling units. Though it is not possible to have one-to-one correspondence between frame units and sampling units, however, we should choose a sampling frame which yields unbiased estimates with a variance as low as possible. Popularly known sampling frame are :

Census reports, electoral registers, lists of member units of trade and industry associations, lists of members of professional bodies, lists of dwelling units maintained by local bodies, returns from an earlier survey and large scale maps etc.

**Step 3: Specify the sampling unit:** The sampling unit is the basic unit containing the elements of the target population.

**Step 4: Specify the sampling method:** The sampling method indicates, how the sample units are selected. The most important decision in this regard is to determine which of the two-probability or non-probability samples is to be chosen.

**Step 5: Determine the sample size (n):** The decision about the number of elements to be chosen, i.e., number of observations in each sample (n) of the target population.

**Step 6: Specify the sampling plan:** This means that one should indicate, how decisions made so far are to be implemented. All expected pertinent issues in a sampling survey must be answered by the sampling plan.

**Step 7: Select the sample:** This is the final step in the sampling process. A good deal of field work and office work is introduced in the actual selection of the sample elements. However it depends mainly upon the sampling plan and the sample size required.

## **Characteristics of a Good sample design**

To adopt an appropriate and unbiased sampling techniques, a researcher has to maintain certain qualities of sampling techniques. They are as follows:

**Representativeness:** An ideal sample is the sample which represents the characteristics of entire population. Thus the selection procedure for sampling should be such that the sample selected has all those qualities and features which the entire population possess.

**Independence:** The second quality required in sampling techniques is that all possible samples which can be selected should be independent of each other. This helps to make an unbiased selection of samples where selection of one unit of population does not depend upon that of another unit.

**Adequacy:** The number of units in a sample, that is sample size, selected from the population should not be too less as this would result in failure to capture the diversity of population. At the same time, it should not be too large. It should be just sufficient to enable a researcher to derive a correct conclusion about population based on his sample.

**Homogeneity:** To make a sample scientific, the element or unit selected within the sample should be identical with another element or unit of that sample.

**Goal Orientation:** Sample design should be oriented to the research objectives.

**Measurability:** The parameters under study should be how or the other measurable so that accuracy can be ensured.

**Usability:** The sample size should be convenient to collect and analyse.

**Cost factor:** The total cost of sample design, collection and analysis of the data should be minimum.

Thus one has to weigh the pros and cons of various sample designs before selecting the best possible one.

## **Sampling and Non-Sampling Errors**

The basic objective of a sample is to draw inferences about the population from which the sample is drawn. Therefore, it is necessary that the sampling technique be a reliable one. The randomness of the sample is especially important because of the principle of statistical regularity which states that a sample taken from a population is likely to possess almost the same characteristics as those

of the population. In the total process starting from data collection to inferring results, errors are bound to come in. these errors can be classified into two groups.

### **Sampling Errors**

Sampling errors are those which arise due to drawing of faulty inferences about the population based on the results obtained from the samples. In other words, it is the difference between the results which would have obtained if the entire population was taken for such a study and the results obtained from the samples drawn from it. The sampling error would be smallest if the sample size is large in relation to the population and vice versa.

### **Non-sampling Errors**

Non-sampling errors are introduced due to technically faulty observations or calculation during the processing of the data. The faulty methods could be used at different stages:

- Methods of data collection.
- Incomplete coverage of the population.
- Inaccurate information provided by the participants.
- Errors occurring during editing, tabulating and mathematical manipulation of data.
- These errors shall arise even if the entire population is taken under study.

Both the sampling as well as non-sampling errors must be reduced to a minimum in order to get a representative sample of the population as far as possible.

For an appropriate study of any problem, it is important to have proper sampling. It means that the sample should be of proper size. If the sample is either too small or too large, it would make the study difficult. The ideal sample size is calculated as follows:

The size of the population: The larger the size of population, the larger should be the sample size.

The degree of Accuracy: If the researcher requires a higher level of accuracy, he needs to have larger sample size.

Homogeneity or Heterogeneity of the population: Depending upon the homogeneity and heterogeneity, the sample size can be determined. Small sample size can be appropriate if the population is homogenous, but if the population is heterogeneous in nature, then the larger sample size is required to cover the diversity.

Method of sampling: If the researcher is adopting simple random sampling, then a large sample size is required, but in case of stratified or cluster sampling, the small number may serve the same purpose.

Nature of Responses: If there is a possibility of a large number of respondents not completing the questionnaire or not giving authentic reply, then a large sample size is required. All the factors should be properly weighed before arriving at the sample size. However, the selection of optimum sample size is not that simple as it might seem to be. Moreover depending upon the time and cost available, a researcher has to rationally determine the sample size. Larger sample size may require more time and money whereas if quick research has to be done, the researcher has to rationally determine the sample size. Larger sample size may require more time and money whereas if quick research has to be done, the researcher cannot afford large sample size.

Statistical method can be employed in determining the sample size, though it is not mandatory. There are various formulas devised for this purpose depending upon the availability of information. It is given as:

$$N = (z^*(\delta/d))^2$$

N = sample size

Z = value of specified level of confidence or degree of precision

$\delta$  = standard deviation of population

d = difference between population mean and sample mean.

However, in practice, this formula is rarely used due to awareness of population mean.

## QUESTIONS

1. Explain the various concepts related with data sampling.
2. What are the salient features of sampling method? .Explain the limitations of sampling method.
3. Write short note on type of sampling error.
4. Write short note calculation of sample size.
5. What are the various types of sampling techniques used in management research?
6. Write short note on the characteristic of a good sample design .

\*\*\*\*\*

## **Chapter :5**

### **METHODS OF DATA COLLECTION**

**Primary data-questionnaire & interviews, collection of secondary data, use of computer and information technology in data collection.**

#### **Meaning of Data**

The search for answers to research questions is called collection of data. Data are facts, and other relevant materials, past and present, serving as bases for study and analyses. Some example of data are:

- The types of Loans secured by borrowers (for a credit survey)
- The items of raw materials required for a product line (Materials management)
- The quantity of each material required for a unit of output.
- The sex, age, social class, religion, income level of respondents in a consumer behaviour study.
- The opinions of eligible couples on birth control devices (Family Planning Survey)
- The capital expenditure proposals considered by a firm during a year (Financial Management) The marks obtained by students of a class in a test on a particular subject (Performances of students)
- The opinions of people on voting in a general election (Opinion Poll)
- The types of news read by newspaper readers (Readership Survey)
- The aspirations of management trainees (The emerging Managers in Indian Enterprises)
- The types and frequency of breakdowns occurred in particular brand of scooter (Post-purchase Behaviour Survey) and so on.

A researcher has to be quite alert about the nature of data. Because the nature of data suggests the use of particular statistical techniques for collection and also analysis.

## **Type of Data**

The data needed for a social science research may be broadly classified into (a) Data pertaining to human beings,

(b) Data relating to organizations, and

(c) Data pertaining to territorial areas.

### **(a) Personal data or data related to human beings consist of-**

(1) Demographic and socio-economic characteristics of individuals: Age, sex, race, social class, religion, marital status, education, occupation, income, family size, location of the household, life style, etc.

(2) Behavioural variables: Attitudes, opinions, awareness, knowledge, practice, intentions, etc.

(b) **Organisational data** consist of data relating to an organization's origin, ownership, objectives, resources, functions, performance and growth.

(c) **Territorial data** are related to geophysical characteristics, resources endowment, population, occupational pattern, infrastructure, structure, degree of development, etc. of spatial divisions like villages, cities, talukas, districts, state and the nation.

Data may also be classified as qualitative (descriptive, attribute) data and quantitative (numerical) data.

## **Importance of data**

The data serve as the bases or raw materials for analysis. Without an analysis of factual data, no specific inferences can be drawn on the questions under study. Inferences based on imagination or guess work cannot provide correct answers to research questions. The relevance, adequacy and reliability of data determine the quality of the findings of a study.

Data form the basis for testing the hypotheses formulated in a study. Data also provide the facts and figures required for constructing measurement scales and tables, which are analysed with statistical techniques. Inferences on the results of statistical analysis and tests of significance provide the answers to research questions. Thus, the scientific process of

measurements, analysis, testing and inferences depends on the availability of relevant data and their accuracy. Hence, the importance of data for any research study.

## **SOURCES OF DATA**

The sources of data may be classified into

**(a) Primary sources** and

**(b) Secondary sources.**

### **Primary Sources**

Primary sources are original sources from which the researcher directly collects data that have not been previously collected, e.g., collection of data directly by the researcher on brand awareness, brand preference, brand loyalty and other aspects of consumer behaviour from a sample of consumers by interviewing them. Primary data are first-hand information collected through various methods such as observation, interviewing, mailing etc.

### **Secondary Sources**

These are sources containing data which have been collected and compiled for another purpose. The secondary sources consists of readily available compendia and already compiled statistical statements and reports whose data may be used by researches for their studies, e.g. census reports, annual reports and financial statements of companies, Statistical statements, Reports of Government Department, Annual Reports on currency and finance published by the Reserve Bank of India, Statistical Statements relating to Cooperatives and Regional Rural Banks, published by the NABARD, Reports of the National Sample Survey Organisation, Reports of trade associations, publications of international organizations such as UNO, IMF, World Bank, H.O. WHO, etc., Trade and Financial Journals, newspapers, etc.

Secondary sources consists of not only published records and reports, but also unpublished records. The latter category includes various records and registers maintained by firms and organizations, e.g., accounting and financial records, personnel records, register of members, minutes of meetings, inventory records, etc.

**Features of Secondary Sources:** Though secondary sources are diverse and consist of all sorts of materials, they have certain common characteristics. They are readymade and readily

available, and do not require the trouble of constructing tools and administering them. They consist of data over which a researcher has no original control over collection and classification. Both the form and the content of secondary sources are shaped by others. Clearly, this is a feature which can limit the research value of secondary sources. Finally, secondary sources are not limited in time and space. That is, the researcher using them need not have been present when and where they were gathered.

### **Advantages**

Secondary sources have some advantages:

1. Secondary data, if available can be secured quickly and cheaply. Once their source of documents and reports are located, collection of data is just a matter of desk work. Even the tediousness of copying the data from the source can now be avoided, thanks to Xeroxing facilities.
2. Wider geographical area and longer reference period may be covered without much cost. Thus, the use of Secondary data extends the researcher's space and time reach.
3. The use of Secondary data broadens the data base from which scientific generalizations can be made. This is especially so when data from several environmental and cultural settings are required for the study.
4. The use of secondary data enables a researcher to verify the findings based on primary data. It readily meets the need for additional empirical support. The researcher need not await the time when additional primary data can be collected.

### **Disadvantages/Limitations**

The use of secondary data has its own limitations:

1. The most important limitation is the available data may not meet our specific needs. The definitions adopted by those who collected those data may be different; units of measure may not match; and time periods may also be different.
2. The available data may not be as accurate as desired. To assess their accuracy we need to know how the data were collected.

3. The Secondary data are not up-to-date and become obsolete when they appear in print, because of time lag in producing them. For example, population census data are published two or three years later after compilation, and no new figures will be available for another ten years.
4. Finally information about the whereabouts of sources may not be available to all social scientists. Even if the location of the source is known, the accessibility depends primarily on proximity. For example, most of the unpublished official records and compilations are located in the capital city, and they are not within the easy reach of researchers based in far off places.

### **Choice of Methods of Data Collection**

Which of the above methods of data collection should be selected for a proposed research project? This is one of the questions to be considered while designing the research plan. One or more methods has/have to be chosen. The choice of a method or methods depends upon the following factors:

1. The nature of the study of the subject-matter: If it is a study of opinions/preferences of persons, interviewing or mailing may be appropriate depending on the educational level of the respondents. On the other hand, an impact study may call for experimentation; and a study of behavioural pattern may require observation.
2. The unit of enquiry: The unit of enquiry may be an individual, household institution or community. To collect data from households, interviewing is preferable. Data from institutions may be collected by mail survey and studies on communities call for observational method.
3. The size and spread of the sample: If the sample is small and the area covered is compact interviewing may be preferable, but a large sample scattered over a wider area may require mailing.
4. Scale of the Survey: A large scale may require mailing or interviewing through trained investigators.
5. The educational level of respondents: For a simple survey among educated persons concerned with the subject-matter of study, a mail survey may be appropriate. But for a

survey of less educated/illiterate persons like industrial workers, slum dwellers, rural people interviewing is the only suitable method.

6. The type and depth of information to be collected: For collection of general, simple, factual and non-emotional data, interviewing or mailing is appropriate. For an in-depth survey of personal experiences and sensitive issues, in-depth interview is essential. For collection of data on behaviour, culture, customs, life style etc., observational method is required.
7. The availability of skilled and trained manpower: In this case, even for a large general survey entailing many complicated questions, interviewing can be adopted.
8. The rate of accuracy and representative nature of the data required: Interviewing is the most appropriate method for collecting accurate data from a representative sample of population. Interviewing can achieve a higher response rate.

A researcher can select one or more of the methods keeping in view the above factors. No method is universal. Each method's unique features should be compared with the needs and conditions of the study and thus the choice of the methods should be decided.

### **Evaluation of Data Collection Methods**

The appropriateness of a method of data collection may be evaluated on the basis of the following criteria:

1. The efficiency i.e., the speed and cost of data collection,
2. Data quality and adequacy i.e., response rate, accuracy and objectivity,
3. Naturalness of setting,
4. Anonymity,
5. Interviewer supervision,
6. Control of context and question order,
7. Ability to use visual aids,
8. Potential for controlling variables and
9. Dependence on respondent's reading and writing ability.

Primary data are directly collected by the researcher from their original sources. In this case, the researcher can collect the required data precisely according to his research needs, he can collect them when he wants them and in the form he needs them. But the collection of primary data is costly and time consuming. Yet, for several types of social science research required data are not available from secondary sources and they have to be directly gathered from the primary sources.

In such cases where the available data are inappropriate, inadequate or obsolete, primary data have to be gathered. They include: socio-economic surveys, social anthropological studies of rural communities and tribal communities, sociological studies of social problems and social institutions, marketing research, leadership studies, opinion polls, attitudinal surveys, readership, radio listening and T.V.viewing surveys, knowledge-awareness practice (KAP) studies, farm management studies, business management studies, etc.

### **Types of Data**

Once the researcher has decided the 'Research Design', the next job is of data collection. For data to be useful, our observations need to be organized so that we can get some patterns and come to logical conclusions. Statistical investigation requires systematic collection of data, so that all relevant groups are represented in the data. To determine the potential market for a new product, for example, the researcher might study 500 consumers in a certain geographical area. It must be ascertained that the group contains people representing variables such as income level, race, education and neighbourhood. The quality of data will greatly affect the conclusions and hence, utmost importance must be given to this process and every possible precaution should be taken to ensure accuracy, while gathering and collecting data. Depending upon the sources utilized, whether the data has come from actual observations or from records that are kept for normal purposes, statistical data can be classified into two categories. Primary and secondary:

#### **❖ Primary Data**

Primary data is one which is collected by the investigator himself for the purpose of a specific inquiry or study. Such data is original in character and is generated by surveys conducted by individuals or research institutions.

#### **❖ Secondary Data**

When an investigator uses the data which has already been collected by others, such data is called secondary data. This data is primary data for the agency that collects it and becomes

secondary data for someone else who uses this data for his own purposes. The secondary data can be obtained from journals, reports, government publications, publication of professional and research organizations and so on. For example, if a researcher desires to analyse the weather conditions of different regions, he can get the required information or data from the records of the meteorology department.

### **Distinction between Primary Data and Secondary Data**

<b>Description</b>	<b>Primary Data</b>	<b>Secondary Data</b>
1. Source	Original source	Secondary source
2. Methods of Data Collection	Observation Method Questionnaire Method	Published data of government agencies Trade Journal etc.
3. Statistical Process	Not done	Done
4. Originality of Data	Original first time collected by user	No data are collected by some other agency
5. Use of Data	For specific purpose data are compiled	Data are taken from other sources and used for decision making
6. Terms and Definitions of Units	Incorporated	Not included
7. Copy of the Schedule	Included	Excluded
8. Methods of Data Collection	Given	Not given
9. Description of Sample Selection	Given	Not given
10. Time	More	Less
11. Cost	Expensive	Cheaper
12. Efforts	More	Less
13. Accuracy	More accurate	Less accurate
14. Training Personnel Required	Experts/trained required	Less trained personnel

#### ➤ **Primary Data**

Primary data is one which is collected by the investigator himself for the purpose of a specific inquiry or study. Such data is original in character and is generated by surveys conducted by individuals or research institutions.

### ➤ **Secondary Data**

When an investigator uses the data which has already been collected by others, such data is called secondary data. This data is primary data for the agency that collects it and becomes secondary data for someone else who uses this data for his own purposes. The secondary data can be obtained from journals, reports, government publications, publication of professional and research organizations and so on. For example, if a researcher desires to analyse the weather conditions of different regions, he can get the required information or data from the records of the meteorology department.

### **Data Collection Procedure for Primary Data**

#### ❖ **Planning the study**

Since the quality of results gained from statistical data depends upon the quality of information collected, it is important that a sound investigative process be established to ensure that the data is highly representative and unbiased. This requires a high degree of skill and also certain precautionary measures may have to be taken.

#### ❖ **Modes of Data Collection**

There are basically three widely used methods for collection of primary data:

- Observation
- Experimentation
- Questionnaire
- Interviewing
- Case Study Method

**OBSERVATION PROCESS** : Information is collected by observing the process at work. The following are a few examples.

- i. Service Stations: Pose as a customer, go to a service station and observe
- ii. To evaluate the effectiveness of display of Dunlop pillow cushions in a departmental store, observer notes:
  - a. How many pass by?
  - b. How many stopped to look at the display?
  - c. How many decide to buy?

- iii. Super Market: What is the best location in the shelf? Hidden cameras are used.
- iv. Concealed tape recorder with the investigator helps to determine typical sales arguments and find out sales enthusiasm shown by various salesmen.

By this method, response bias is eliminated.

The method can be used to study sales techniques, customer movements, customer response, etc. However, the customer's / consumer's state of mind, their buying motives, their images are not revealed. Their income and education is also not known. It also takes time for the investigator to wait for particular sections to take place.

**EXPERIMENTATION METHOD** : Many of the important decisions facing the marketing executive cannot be settled by secondary research, observation or by surveying the opinions of customers or experts. Experimental method may be used in the following situations.

- i. What is the best method for training salesmen?
- ii. What is the best remuneration plan for salesman?
- iii. What is the best shelf arrangement for displaying a product?
- iv. What is the effectiveness of a point-of-purchase display?
- v. What package design should be used?
- vi. Which copy is the most effective?
- vii. What media are the most effective?
- viii. Which version of a product would consumers like best?

In a marketing experiment, the experimental units may be consumers, stores, sales territories, etc. Factors or marketing variables under the control of the researcher which can be studied are price, packaging, display, sales incentive plan, flavour, colour, shape, etc. To study the effect of the marketing variables in the presence of environmental factors, a sufficiently large sample should be used. Or sometimes a control group is set up. A control group is a group equivalent to the experimental group and differing only in not receiving any treatment. The result/response of a marketing experiment will be in the form of sales, attitudes or behavior.

**QUESTIONNAIRE TECHNIQUE** : The survey method is the technique of gathering data by asking questions from people who are thought to have the desired information.

## **Advantages**

One cannot know by observation, why a buyer makes particular purchases or what is his opinion about a product. Compared with either direct observation or experimentation, surveys yield a broader range of information and are effective for producing information on socio-economic characteristics, attitudes, opinions, motives, etc and to gather information for planning product features, advertising copy, advertising media, sales promotions, channels of distribution and other marketing variables. Questioning is usually faster and cheaper than observation.

## **Limitations**

- a. Unwillingness of respondents to provide information: This requires salesmanship on the part of the interviewer. The interviewer may assure that the information will be kept secret. Motivating respondents with some token gifts often yield result.
- b. Inability of the respondents to provide information: This may be due to:
  - i. Lack of knowledge.
  - ii. Lapse of memory.
  - iii. Inability to identify their motives and provide 'reasons why' for their actions.
- c. Human biases of the respondents: i.e., ego, etc.
  - i. Semantic difficulties: It is difficult, if not impossible, to state a given question in such a way that it will mean exactly the same thing to every respondent. Similarly, two different wordings of the same question will frequently generate quite different results.

**INTERVIEWING :** Interview on samples may be carried out either with a structured framework or with an undirected approach. The structured framework involves use of some predetermined questions. Such pre-determination enables the researcher to standardize the responses with some fixed alternatives. The samples here are merely directed to choose answers/responses from different pre-determined alternatives. Thus the researcher can or may quantify the responses in line with his research object. Standardizing the responses with pre-determination involves great amount of risk unless the researcher acquaints himself with the intricacies of the research matter in much greater details. However, this approach is more scientific in nature for its feasibility of quantifications with least trouble and application of scientific techniques with more rationality.

Unstructured or undirected interview approach enables the respondents or the samples to answer the researcher's queries with greater amount of flexibility. Since no predetermined responses here are advised, the researcher may proceed, keeping in tune with the research matter, with greater amount of flexibility too. However, quantification of the responses from unstructured interviews are difficult unless the researcher fixes the standard of all response with some amount of control. If unstructured approach may defeat the purpose and object of research. This approach is resorted to usually in cases where the selected samples need to be interviewed in a more intensive way.

Interviewing the subjects or the samples is more advantageous than sending questionnaires through mail. Interview method enables the researcher to personally feel the problems of samples. Moreover interviewer/researcher, being present on the spot, case study certain evaluative variables like facial expressions and gestures of the samples. For high reliability and feasibility of scoring using test devices, interview approach is more scientific than mailing questionnaire. Experimentation is a research process used to study the causal relationships between variables. It aims at studying the effect of an independent variable on a dependent variable, by keeping the other independent variable constant through some type of control.

Why Experiment?

Experimentation requires special efforts. It is often extremely difficult to design, and it is also a time consuming process. The experiment is the only method which can show the effect of an independent variable on dependent variable. In experimentation, the researcher can manipulate the independent variable and measure its effect on the dependent variable. Moreover, experiment provides "the opportunity to vary the treatment (experimental variable) in a systematic manner, thus allowing for the isolation and precise specification of important differences.

### **Planning and Conducting Experiments**

- 1 Determine the hypothesis to be tested and the independent and dependent variables involved in it.
- 2 Operationalize the variables by identifying their measurable dimensions.
- 3 Select the type of experimental plan. The types of experimental design based on types of control may be classified into: (1) one group plan, using the same group as

experimental and control group and measuring it before and after experimental treatment; (2) matched groups plan, consisting of two identical groups, one to be used as control group and another as experimental group, with (a) Post-test only measurement or (b) pretest-post-test measurements.

- 4 Choose the setting. The setting may be field or laboratory.
- 5 Make the experimental conditions as nearly the same as the expected real life conditions. This is essential in order to make the findings reliable.
- 6 Make a record of pre-experimental conditions.
- 7 Introduce appropriate methods for controlling extraneous variables that are not manipulated in the experiment.
- 8 Apply the experimental treatment and record observations and measurements using appropriate measurement devices. If feasible, repeat the tests several times in order to insure the accuracy of results.
- 9 Analyse the results, using appropriate statistical devices. Last, interpret the results, giving consideration to all possible extraneous conditions. No possible cause should be overlooked, as unforeseen conditions might influence the results.

**The advantages are:**

1. Its power to determine causal relationships between variables surpasses that of all other methods. The influence of extraneous variables can be more effectively controlled in this method.
2. The element of human errors is reduced to the minimum.
3. In this method better conditions for conducting experiments may be created, than is possible in other methods.
4. Experimentation yields generally exact measurements and can be repeated for verifying results.

**The disadvantages:**

- 1 It is difficult to establish comparable control and experimental groups.

- 2 The scope for experimentation with human beings is extremely limited.
- 3 Experiment is often difficult to design, tends to be expensive and time-consuming.
- 4 It is artificial to some extent and may lack realism.
- 5 Experimentation can be used only in studies of the present but not in studies relating to past or future.
- 6 it is of no use in determining opinions, motives and intentions of persons.

Simulation is one of the forms of observational methods. It is a process of conducting experiments on a symbolic model of representing a phenomenon. Abelson defines simulation as “the exercise of a flexible limitation of process and outcomes for the purpose of clarifying or explaining the underlying mechanisms involved.” It is a symbolic abstraction, simplification and substitution for some referent system. In other words, simulation is a theoretical model of the elements, relations and processes which symbolize some referent system, e.g., the flow of money in the economic system may be simulated in an operating model consisting of a set of pipes through which liquid moves. Simulation is thus a technique of performing sampling experiments on the model of the systems. The experiments are done on the model instead of on the real system, because the latter would be too inconvenient and expensive.

### **The Process of Simulation**

- 1 The process or system to be simulated is identified.
- 2 The purpose of this simulation is decided. It may be to ‘clarify’ or ‘explain’ the process.
- 3 On the basis of the available information on the process or system – its components and the set of conditions assumed to operate in and between the components a mathematical model is developed.
- 4 Several sets of inputs data to be used are collected. Inputs may be samples of actual data or synthetic data based on the general characteristics of real input data.
- 5 The type of simulation – computer simulation or man simulation or man-computer simulation-to be used is determined.

Lastly, the simulation is operated with the various sets of input data, and the results are analysed to determine the best solution.

## **Major Steps in Conducting a Survey**

### **❖ Deciding on the Research Objectives**

Every effort should be made to state the objectives in specific terms. Surveys, in particular, can proceed in an almost unlimited number of directions. To prevent all kinds of questions from being asked, clear informational objectives should be developed and put in writing. If possible.

### Methods of Collection of Data

Following methods are in use for collection of data for questionnaire technique:

- a. Telephone enquiries.
- b. Postal or Mail questionnaire.
- c. Personal interviewing.
- d. Panel Research.
- e. Group Interview Technique.
- f. Special Survey techniques.

Each of this method has its own advantages and disadvantages. Telephone interviewing stands out as the best method for gathering quickly needed information. It has the advantage over a mailed questionnaire as it permits the interviewer to talk to one or more persons and to clarify his questions, if they are not understood. The response rate for telephone interviewing seems to be a little better than for mailing questionnaires. The two main drawbacks of telephone interviewing are that only people with telephones can be interviewed and only short, not too personal interviews can be carried out.

The questionnaire mailing may be the best way to reach persons who would not give personal interviews or who might be biased by interviewers. It is typically the least expensive than other major methods. On the other hand, mailing questionnaires require simple and clearly worded questions. The response rate to mailed questionnaires is typically low.

Personal interviewing is the most versatile of the three methods. The personal interviewer can ask more questions and can supplement the interview with personal observations. These advantages come at a high cost, however. Personal interviewing is the most expensive method and requires much more technical and administrative planning and supervision. In a real sense,

companies turn to telephone interviewing or questionnaire mailing as a second choice out of cost consideration.

### ❖ **Construction of a Questionnaire**

When information is to be collected by asking questions to people who may have the desired data, a standardized form called questionnaire is prepared. The questionnaire has a list of questions to be asked and spaces in which the respondents record the answers. Each question is worded exactly as it is to be asked. Also, the questions are listed in an established sequence.

Questionnaire construction is discussed below in nine steps. These steps may vary in importance in individual projects, but each one must be thought through. The nine steps are:

- i. Decide what information is wanted.
- ii. Decide what type of questionnaire (personal interview, mail telephone) to use.
- iii. Decide on the content of individual questions.
- iv. Decide on the type of question (open, multiple choice, dichotomous) to use.
- v. Decide on the wording of the questions.
- vi. Decide on question sequence.
- vii. Decide on layout and method of reproduction of questionnaire.
- viii. Make a preliminary draft and pre-test it.
- ix. Revise and prepare the final questionnaire.

Basically, a questionnaire must serve two functions; translate research objectives into specific questions and motivate the respondent to cooperate with the survey and furnish the information correctly. Therefore, before a questionnaire can be formulated, a specific statement of the information which is needed must be made. The complete analysis must be anticipated

**Determine the manner questionnaire to be used:** Questionnaire can be used by personal interview, mail, or telephone. The choice among these alternatives is largely determined by the type of information to be obtained and by the type of respondents from whom it is to be obtained. It is necessary to decide on the type of questionnaire at this point since the questions asked, the way in which they are asked and the sequence in which they are asked will all be influenced by this decision.

**Determine the content of individual questions:** Once the specific information needed is known and the method of communication is decided, the researcher is ready to begin formulating his questionnaire. A first problem is to decide what to include in individual questions. The following points are in the nature of standards against which to check possible questions; obviously, they leave much to the originality of the researcher.

- Is the question necessary?
- Are several questions needed instead of one?
- Does the respondent have the information requested?
- Is the point within the respondent's experience?
- Can the respondent remember the information?
- Will the respondent have to do a lot of work to get the information?
- Will respondents give the information?

Some respondents do not want to answer particular questions. It goes without saying that such questions hurt cooperation for the rest of the interview and, therefore, should be eliminated. It is often possible, however, to change such questions so as to secure the desired information.

**Determine the type of question to use:** Once the content of individual questions is decided, the researcher is ready to begin framing the actual question. Before he can work on the wording of each question, he must decide on the type of questions to use. There are three major types from which he may choose: (1) open (2) multiple choice, and (3) dichotomous.

Multiple choice questions overcome some of the disadvantages of open questions, but incur some new ones. Open questions are subject to interviewer bias in recording of answers. This is not the problem with multiple choice questions where answers are in one or more of the stated alternatives. All that the interviewer has to do is to check the applicable reply. Thus, the multiple choice questions are faster and less subject to bias. Also, multiple choice questions simplify the tabulating process. The difficult and time-consuming editing process is reduced to a rapid check for mechanical accuracy.

Multiple choice questions give a list of alternative answers. It is important to point out that this list must include all alternatives or there will be a bias against those options which are omitted. The

alternatives mentioned in a multiple choice questions will be reported by more respondents compared to a situation when open questions are used.

Dichotomous Question or two-way question, is an extreme of the multiple choice question. The idea is to offer only two choices – yes or no, did or didn't, cash or credit, railroad or airline, etc. such questions are the most widely used of the three basic types. The following are examples of dichotomous questions:

Would the service proposed by X lines make motor freight service more useful to you?

Is any of this discount normally passed on to other?

Did you buy it or was it a gift?

Was it new or used when you got it?

**Decide on question sequence:** Once the wording of the individual question has been determined, it is necessary to set them up in some order. The order chosen can change the results obtained. There are three major sections in a questionnaire (1) The basic information sought (2) Classification information and (3) Identification information. Since questions pertaining to these sections tend to be of declining interest to the respondent, the sections are usually put in the order shown. Questions relating to the basic information are place first. To help in analyzing this information, it is usually necessary to be able to classify respondents on such bases as age, sex, income, education and nationality. Questions on these points form the classification section. The identification section identifies all parties involved. This includes the name and address of the respondent, and the names of such individuals as the interviewer, editor, and card puncher. These are used to permit checking for cheating among interviewers and to assign responsibility for the tasks done.

**Decide on layout and reproduction :** The physical layout and reproduction of a questionnaire can influence its success. Three major points should be considered in planning the layout and reproduction of the questionnaire:

- Securing acceptance of the questionnaire by respondents
- Making it easy to control the questionnaire, and
- Making it easy to handle the questionnaire.

**Pre-test :** Before a questionnaire is ready for the field it needs to be pretested under field conditions. No researcher can prepare a questionnaire perfectly in the first attempt; improvements

can hence be suggested in field tests. Pre-tests are best done by personal interview even if the survey is to be handled by mail or telephone. Interviewers can note the respondents reactions and attitudes which cannot otherwise be obtained. After any pertinent changes in the questionnaire have been made, another pre-test can be done by mail or telephone, if those methods are to be used in the survey. This latter pre-test should uncover any weakness peculiar to the method of communication.

**Revisions and final draft:** After each significant revision of the questionnaire, another pre-test should be done. When the last pre-test suggests no new revisions, the researcher is ready to print the actual questionnaire to be used in the survey.

**Selection of a sample:** An elementary unit is an element or group of elements on which observations can be made or from which the required statistical information can be ascertained according to a well defined procedure. Elementary units or groups of such units which are convenient for purposes of sampling are called sampling units.

The totality of all sampling units belonging to the population to be studied with their proper identification particulars is termed as the sampling frame. Editing would also help to eliminate inconsistencies or obvious errors due to arithmetical treatment. When the data is to be processed by computers, then it must be coded and converted into the computer language. This coding job should be done while editing the data.

### **Definition**

Interviewing is one of the prominent methods of data collection. It may be defined as a two-way systematic conversation between an investigator and an informant, initiated for obtaining information relevant to a specific study. It involves not only conversation, but also leaning from the respondent's gestures, facial expressions and pauses, and his environment. Interviewing requires face-to-face contact or contact over telephone and calls for interviewing skills. It is done by using a structured schedule or an unstructured guide.

### **Importance**

Interviewing may be used either as a main method or as a supplementary one in studies of persons. Interviewing is the only suitable method for gathering information from illiterate or less educated respondents. It is useful for collecting a wide range of data from factual demographic data to highly personal and intimate information relating to a person's opinions, attitudes, values, beliefs, past

experience and future intentions. When qualitative information is required or probing is necessary to draw out fully, then interviewing is required. Where the area covered for the survey is a compact, or when a sufficient number of qualified interviewers are available, personal interview is feasible.

Interview is often superior to other data-gathering methods. People are usually more willing to talk than to write. Once rapport is established, even confidential information may be obtained. It permits probing into the context and reasons for answers to questions.

Interview can add flesh to statistical information. It enables the investigator to grasp the behavioural context of the data furnished by the respondents. It permits the investigator to seek clarifications and brings to the forefront those question, that for one reason or another, respondents do not want to answer.

### **Evaluation of Interviewing**

**Advantages:** There are several real advantages to personal interviewing.

- 1 The greatest value of this method is the depth and detail of information that can be secured. When used with a well-conceived schedule, an interview can obtain a great deal of information. It far exceeds mail survey in amount and quality of data that can be secured.
- 2 The interviewer can do more to improve the percentage of responses and the quality of information received than other method. He can note the conditions of the interview situation, and adopt appropriate approaches to overcome such problems as the respondent's unwillingness, incorrect understanding of question, suspicion, etc.
- 3 The interviewer can gather other supplemental information like economic level, living conditions etc. Through observation of the respondent's environment.
- 4 The interviewer can use special scoring devices, visual materials and like in order to improve the quality of interviewing.
- 5 The accuracy and dependability of the answers given by the respondent can be checked by observation and probing.
- 6 Interview is flexible and adaptable to individual situations. Even more control can be exercised over the interview situation.

**Limitations:** Interviewing is not free from limitations.

- 1 Its greatest drawback is that it is costly, both in money and time.
- 2 The interview results are often adversely affected by interviewer's mode of asking questions and interactions, and incorrect recording and also be the respondent's faulty perception, faulty memory, inability to articulate etc.
- 3 Certain types of personal and financial information may be refused in face-to-face interviews. Such information might be supplied more willingly on mail questionnaires, especially if they are to be unsigned.
- 4 Interview poses the problem of recording information obtained from the respondents. No foolproof system is available. Note taking is invariably distracting to both the respondent and the interviewer and affects the thread of the conversation.
- 5 Interview calls for highly skilled interviewers. The availability of such persons is limited and the training of interviewers is often a long and costly process.

### **Characteristics**

Interviewing as a method of data collection has certain characteristics. They are:

1. The participants-the interviewer and the respondent-are strangers. Hence, the investigator has to get himself introduced to the respondent in an appropriate manner.
2. The relationship between the participants is a transitory one. It has a fixed beginning and termination points. The interview proper is a fleeting, momentary experience for them.
3. Interview is not a mere casual conversational exchange, but a conversation with a specific purpose, viz., obtaining information relevant to a study
4. Interview is a mode of obtaining verbal answers to questions put verbally.
5. The interaction between the interviewer and the respondent need not necessarily be on a face-to-face basis, because interview can be conducted over the telephone also.
6. Although interview is usually a conversation between two persons, it need not be limited to a single respondent. It can also be conducted with a group of persons, such as family members, or a group of children or a group of customers, depending on the requirements of the study.

7. Interview is an interactional process. The interaction between the interviewer and the respondent depends up-on how they perceive each other. The respondent reacts to the interviewer's appearance, behaviour, gestures, facial expression and intonation, his perception of the thrust of the questions and his own personal needs. As far as possible, the interviewer should try to be closer to the socio-economic level of the respondents. Moreover, he should realize that his respondents are under no obligation to extend response. He should, therefore, be tactful and be alert to such reactions of the respondents as lame-excuse, suspicion, reluctance or indifference, and deal with them suitably. He should not also argue or dispute. He should rather maintain an impartial and objective attitude.
8. Information furnished by the respondent in the interview is recorded by the investigator. This poses a problem of seeing that recording does not interfere with the tempo of conversation.
9. Interviewing is not a standardized process like that of a chemical technician; it is rather a flexible psychological process.

The implication of this feature is that the interviewer cannot apply unvarying standardized technique, because he is dealing with respondents with varying motives and diverse perceptions. The extent of his success as an interviewer is very largely dependent upon his insight and skill in dealing with varying socio-psychological situations.

### **Requirements**

The requirements or conditions necessary for a successful interviews are:

1. Data availability: The needed information should be available with the respondent. He should be able to conceptualize it in terms useful to the study, and be capable of communicating it.
2. Role perceptioOn: The respondent should understand his role and know what is required of him. He should know what is a relevant answer and how complete it should be. He can learn much of this from the interviewer's introduction, explanations and questioning procedure.

3. The interviewer should also know his role. He should establish a permissive atmosphere and encourage frank and free conversation. He should not affect the interview situation through subjective attitude, argumentation, etc.
4. Respondent's motivation: The respondent should be willing to respond and give accurate answer. This depends partly on the interviewer's approach and skill. The interview has interest in it for the purpose of his research, but the respondent has no personal interest in it. Therefore, the interviewer should establish a friendly relationship with the respondent, and create in him an interest in the subject-matter of the study.

The interviewer should try to reduce the effect of demotivating factors like desire to get on with other activities, embarrassment at ignorance, dislike of the interview content, suspicious about the interview, and fear of consequences. He should also try to build up the effect of motivating factors like curiosity, loneliness, politeness, sense of duty, respect of the research agency and liking for the interviewer.

The above requirement reminds that the interview is an interactional process. The investigator should keep this in mind and take care to see that his appearance and behaviour do not distort the interview situation.

### **Types of Interviews**

The interviews may be classified into: **(a) structured or directive interview, (b) unstructured or non-directive interview, (c) focused interview, and (d) clinical interview and (e) depth interview.**

- **Structured, Directive Interview**

This is an interview made with a detailed standardized schedule. The same questions are put to all the respondents and in the same order. Each question is asked in the same way in each interview, promoting measurement reliability. This type of interview is used for large-scale formalized surveys.

Advantages: This interview has certain advantages. First, data from one interview to the next one are easily comparable. Second, recording and coding data do not pose any problem, and greater precision is achieved. Lastly, attention is not diverted to extraneous, irrelevant and time-consuming conversation.

Limitation: However, this type of interview suffers from some limitations. First, it tends to lose the spontaneity of natural conversation. Second, the way in which the interview is structured may be such that the respondent's views are minimized and the investigator's own biases regarding the problem under study are inadvertently introduced. Lastly, the scope for exploration is limited.

- **Unstructured or Non-directive Interview**

This is the least structured one. The interviewer encourages the respondent to talk freely about a given topic with a minimum of prompting or guidance. In this type of interview, a detailed pre-planned schedule is not used. Only a broad interview guide is used. The interviewer avoids channelling the interview directions. Instead, he develops a very permissive atmosphere. Questions are not standardized and not ordered in a particular way.

This interviewing is more useful in case studies rather than in surveys. It is particularly useful in exploratory research where the lines of investigation are not clearly defined. It is also useful for gathering information on sensitive topics such as divorce, social discrimination, class conflict, generation gap, drug-addition etc. It provides opportunity to explore the various aspects of the problem in an unrestricted manner.

**Advantages :** This type of interview has certain special advantages. It can closely approximate the spontaneity of a natural conversation. It is less prone to interviewer's bias. It provides greater opportunity to explore the problem in an unrestricted manner.

**Limitations:** Though the unstructured interview is a potent research instrument, it is not free from limitations.

- 1 One of its major limitations is that the data obtained from one interview is not comparable to the data from the next. Hence, it is not suitable for surveys.
- 2 Time may be wasted in unproductive conversations. By not focusing on one or another facet of a problem, the investigator may run the risk of being led up blind ally.
- 3 As there is no particular order or sequence in this interview, the classification of responses and coding may require more time.
- 4 This type of informal interviewing calls for greater skill than the formal survey interview.

- **Focused Interview**

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 involved in a particular experience, e.g., seeing a particular film, viewing a particular programme on T.V., involved in a train/bus accident, etc. The situation is analysed prior to the interview. An interview guide specifying topics relating to the research hypothesis is used. The interview is focused on the subjective experiences of the respondent, i.e. his attitudes, and emotional responses regarding the situation under study. The focused interview permits the interviewer to obtain details of personal reactions, specific emotions.

**Merits:** This type of interview is free from the inflexibility of formal methods, yet gives the interview a set form and insures adequate coverage of all the relevant topics. The Respondent is asked for certain information, yet he has plenty of opportunity to present his views. The interviewer is also free to choose the sequence of questions and determine the extent of probing.

- **Clinical Interview**

This is similar to the focused interview but with a subtle difference. While the focused interview is concerned with the effect of a specific experience, clinical interview is concerned with broad underlying feelings or motivations or with the course of the individual's life experiences.

The 'Personal history' interview used in social case work, prison administration, psychiatric clinics and in individual life history research is the most common type of clinical interview. The specific aspects of the individual's life history to be covered by the interview are determined with reference to the purpose of the study and the respondent is encouraged to talk freely about them.

- **Depth Interview**

This is an intensive and searching interview aiming at studying the respondent's opinion emotions or convictions on the basis of an interview guide. This requires much more training interpersonal skills than structured interviewing. This deliberately aims to elicit unconscious as well as extremely personal feelings and emotions. This is generally a lengthy procedure designed to encourage free expression of affectively charged information. it requires probing. The interviewer should totally avoid advising or showing disagreement. Of course, he should use encouraging expressions like "uh-huh" or "I see" to motivate the respondent to continue narration.

Sometimes the depth interviewer has to face the problem of affection, i.e., the respondent may hide expressing affective feelings. The interviewer should handle such situation with great care.

### **Interviewing process**

The interviews process consists of the following stages:

- Preparation
  - Introduction
  - Developing rapport
  - Carrying the interview forward
  - Recording the interview, and
  - Closing the interview.
- **Preparation** : The interviewing requires some preplanning and preparation. The interviewer should keep the copies of interview schedule/guide (as the case may be) ready for use. He should also have the list of names and addresses of respondents, he should regroup them into contiguous groups in terms of location in order to save time and cost in travelling. The interviewer should find out the general daily routine of the respondents in order to determine the suitable timings for interview. Above all, he should mentally prepare himself for the interview. he should think about how he should approach a respondent, what mode of introduction he could adopt, what situations he may have to face and how he could deal with them.

The interviewer may come across such situations as respondents' avoidance, reluctance, suspicion, diffidence, inadequate responses, distortion, etc. the investigator should plan the strategies for dealing with them. If such preplanning is not done, he will be caught unaware and fail to deal appropriately when he actually faces any such situation. it is possible to plan in advance and keep the plan and mind flexible and expectant of new development.

- **Introduction:** The investigator is a stranger to the respondents. Therefore he should be properly introduced to each of the respondents. what is the proper mode of introduction? There is no one appropriate universal mode of introduction. Mode varies according to the type of respondents.

When making a study of an organization or institution, the head of the organization should be approached first and his cooperation secured before contacting the sample inmates/employees. When studying a community or a cultural group, it is essential to approach the leader first and to enlist his cooperation.

For a survey of urban households, the research organization's letter of introduction and the interviewer's identify card can be shown. In these days of fear of opening the door for a stranger, residents' cooperation can be easily secured, if the interviewer attempts to get him introduced through a person known to them, say a popular person in the area e.g., a social worker.

For interviewing rural respondents, the interviewer should never attempt to approach them along with someone from the revenue department, for they would immediately hide themselves, presuming that they are being contacted for collection of land revenue or subscription to some government bond. He should not also approach them through a local political leader, because persons who do not belong to his party will not cooperate with the interviewer. It is rather desirable to approach the rural respondents through the local teacher or social worker.

After getting himself introduced to the respondent in the most appropriate manner, the interviewer can follow a sequence of procedures as under, in order to motivate the respondent to permit the interview:

- *With a smile greet the respondent in accordance with his cultural pattern*
- *Identify the respondent by name*
- *Describe the method by which the respondent was selected*
- *Mention the name of the organization conducting the research*
- *Assure the anonymity or confidential nature of the interview*
- *Explain their usefulness of the study*
- *Emphasize, the value of respondent's cooperation, making such statements as. "You are among the few in a position to supply the information." "your response is invaluable." I have come to learn from your experience and knowledge."*

- **Developing rapport:** Before starting the research interview, the interviewer should establish a friendly relationship with the respondent. This is described as “rapport.” It means establishing a relationship of confidence and understanding between the interviewer and the respondent. It is a skill which depends primarily on the interviewer’s commonsense, experience, sensitivity, and keen observation. Start the conversation with a general topic of interest such as weather, current news, sports event, or the like perceiving the probable interest of the respondent from his context.

Such initial conversation may create a friendly atmosphere and a warm interpersonal relationship and mutual understanding of the other. However the interviewer should “guard against the over-rapport” as cautioned by Herbert Hyman. Too much identification and too much courtesy result in tailoring replies to the image of a “nice interviewer.” The interviewer should use his discretion in striking a happy medium.

- **Carrying the interview forward:** After establishing rapport, the technical task of asking questions from the interview schedule starts. This task requires care, self-restraint, alertness and ability to listen with understanding, respect and curiosity. In carrying on this task of gathering information from the respondent by putting questions to him, the following guidelines may be followed:

1. Start the interview. Carry it on in an informal, natural conversational style.
2. Ask all the applicable questions in the same order as they appear on the schedule without any elucidation and change in the wording. Ask all the applicable questions listed in the schedule. Do not take answers for granted.
3. If interview guide is used, the interviewer may tailor his questions to each respondent, covering, of course, the areas to be investigated.
4. Know the objectives of each questions so as to make sure that the answers adequately satisfy the question objectives.
5. If a question is not understood, repeat it slowly with proper emphasis and appropriate explanation, when necessary.
6. Take all answers naturally, never showing disapproval or surprise. When the respondent does not meet with interruptions, denial, contradiction and other

- harassment, he may feel free and may not try to withhold information. He will be motivated to communicate when the atmosphere is permissive and the listener's attitude is non-judgement and is genuinely absorbed in the revelations.
7. Listen quietly with patience and humility. Give not only undivided attention, but also personal warmth. At the same time, be alert and analytic to incomplete, nonspecific and inconsistent answers, but avoid interrupting the flow of information. If necessary, jot down unobtrusively the points which need elaboration or verification for later and timelier probing.
  8. Neither argue nor dispute.
  9. Show genuine concern and interest in the ideas expressed by the respondent, at the same time, maintain an impartial and objective attitude.
  10. Should not reveal your own opinion or reaction. Even when you are asked for your views, laugh off the request, saying "Well, your opinions are more important than mine."
  11. At times the interview 'runs dry' and needs re-stimulation. Then use such expressions as "Uh-huh" or "That's interesting" or "I see," "can you tell me more about that?" and the like.
  12. When the interviewee fails to supply his reactions to related past experiences, represent the stimulus situation, introducing appropriate questions which will aid in revealing the past "Under what circumstances did such and such a phenomenon occur?" or "How did you feel about it and the like.
  13. At times, the conversation may go off the track. Be alert to discover drifting, steer the conversation back to the track by some such remark as, "You know, I was very much interested in what you said a moment ago. Could you tell me more about it?"
  14. When the conversation turns to some intimate subjects, and particularly when it deals with crises in the life of the individual, emotional blockage may occur. then drop the subject for the time being and pursue another line of conversation for a while a while so that a less direct approach to the subject can be made later.

15. When there is a pause in the flow of information, do not hurry the interview. Take it as a matter of course with an interested look or a sympathetic half-smile. If the silence is too prolonged, introduce a stimulus saying “ You mentioned that...What happened then?”

- **Recording the interview:** It is essential to record responses as they take place. If the note-taking is done after the interview, a good deal of relevant information may be lost. Nothing should be made in the schedule under respective question. It should be complete and verbatim. The responses should not be summarized or paraphrased. How can complete recording be made without interrupting the free flow of conversation? Electronic transcription through devices like tape recorder can achieve this. It has obvious advantages over note-taking during the interview. But it also has certain disadvantages. Some respondents may object to or fear “going on record.” Consequently the risk of lower response rate will arise especially for sensitive topics.

If the interviewer knows short-hand, he can use it with advantage. Otherwise, he can write rapidly by abbreviating words and using only key words and the like. However, even the fast writer may fail to record all that is said at conversational speed. At such times, it is useful to interrupt by some such comment as “that seems to be a very important point, would you mind repeating it, so that I can get your words exactly.” The respondent is usually flattered by this attention and the rapport is not disturbed. The interviewer should also record all his probes and other comments on the schedule in brackets to set them off from responses. With the pre-coded structured questions, the interviewer's task is easy. He has to simply ring the appropriate code or tick the appropriate box, as the case may be. He should not make mistakes by carelessly ringing or ticking a wrong item.

- **Closing the interview:** After the interview is over, take leave of the respondent, thanking him with a friendly smile. In the case of a qualitative interview of longer duration, select the occasion for departure more carefully. Assembling the papers for putting them in the folder at the time of asking the final question sets the stage for a final handshake, a thank you and a good-bye. If the respondent desires to know the result of the survey, note down his name and address so that a summary of the result could be posted to him when ready.

- **Editing** : At the close of the interview, the interviewer must edit the schedule to check that he has asked all the questions and recorded all the answers and that there is no inconsistency between answers. Abbreviations in recording must be replaced by full words. He must ensure that everything is legible. It is desirable to record a brief sketch of his impressions of the interview and observational notes on the respondent's living environment, his attitude to the survey, difficulties, if any, faced in securing his cooperation and the interviewer's assessment of the validity of the respondent's answers.

### **Interview Problems**

In personal interviewing, the researcher must deal with two major problem, inadequate response, non-response and interviewer's bias.

**Inadequate response:** Kahn and Cannel distinguish five principal symptoms of inadequate response. They are partial response, in which the respondent gives a relevant but incomplete answer, non-response, when the respondent remains silent or refuses to answer the question, irrelevant response, in which the respondent's answer is not relevant to the question asked; inaccurate response, when the reply is biased or distorted; and the verbalized response problem, which arises on account of respondent's failure to understand a question or lack of information necessary for answering it. The problem of inaccurate response is common in economic surveys. the respondents have difficulty in furnishing accurate information on 'sensitive' matters like asset holdings, income, expenditure, saving and investments. It is difficult to deal with this problem. perhaps one possible approach is to use indirect questions instead of direct questions for securing information on above matters, and to cross-check with information furnished on other related questions. For example, data on income can be verified with the details on occupation.

**Interviewer's bias** “: The interviewer is an important cause of response bias. he may resort to cheating by 'cooking up' data without actually interviewing. The interviewers can influence the responses by inappropriate suggestions, word emphasis, tone of voice and question rephrasing. His own attitudes and expectations about what a particular category of respondents may say or think may bias the data. The respondent's perception of the interviewer's characteristics (education, apparent social status, etc) may also bias his answers.

Another source of response bias arises from interviewer's perception of the situation. If the regards the assignment as impossible or sees the results of the survey as a possible threat to

personal interests or beliefs he is likely to introduce bias. As interviewers are human beings, such biasing factors can never be overcome completely, but their effects can be reduced by careful selection and training of interviewers, proper motivation and supervision, standardization of interview procedures (use of standard wording in survey questions, standard instructions on probing procedure and so on) and standardization of interviewer behaviour. There is need for more research on ways to minimize bias in the interview.

### **Non-response**

Non-response refers to failure to obtain responses from some sample respondents. There are many sources of non-response; non-availability, refusal, incapacity, inaccessibility.

: Some respondents may not be available at home at the time of call. This depends upon the nature of the respondent and the time of calls. For example employed persons may not be available during working hours. Farmers may not be available at home during cultivation season. Selection of appropriate timing for calls could solve this problem. Evenings and weekends may be favourable interviewing hours for such respondents. If someone is available, then the respondent's hours of availability can be ascertained and the next visit can be planned accordingly.

<b>QUESTIONS</b>
------------------

1. Efficient Data collection eases the task of a researcher. Describe?
2. Describe the factors on which data collection methods depend.
3. Explain the types of data collection.
4. What are various means of collecting data ?
5. Write shortnote on Experimentation.
6. Write shortnote on Simulation as a means of data collection.
7. Describe survey method as a means of data collection.
8. Explain the interview method of data collection in detail.



## **Chapter : 6**

### **DATA PROCESSING**

Collection & processing data-filed work, survey errors, data coding, editing & tabulation.

The data, after collection, has to be processed and analysed in accordance with the outline laid down for the purpose at the time of developing the research plan. This is essential for a scientific study and for ensuring that we have all relevant data for making contemplated comparisons and analysis. Technically speaking, processing implies editing, coding, classification and tabulation of collected data so that they are amenable to analysis. The term analysis refers to the computation of certain measures along with searching for patterns of relationship that exist among data-groups. Thus, “in the process of analysis, relationships or differences supporting or conflicting with original or new hypotheses should be subjected to statistical tests of significance to determine with what validity data can be said to indicate any conclusions”. But there are persons (Selltitz, Jahoda and others) who do not like to make difference between processing and analysis. They opine that analysis of data in a general way involves a number of closely related operations which are performed with the purpose of summarizing the collected data and organizing these in such a manner that they answer the research question (s).

Data in the real world often comes with a large quantum and in a variety of formats that any meaningful interpretation of data cannot be achieved straightaway. Data processing is an intermediary stage of work between data collection and data interpretation. The data gathered in the form of questionnaires/interview schedules/field notes/data sheets is mostly in the form of a large volume of research variables. The research variables recognized is a result of the preliminary research plan, which also sets out the data processing methods beforehand. Processing of data requires advance planning and this planning may cover such aspects as identification of variables, hypothetical relationship among the variables and the tentative research hypotheses.

The various steps in processing of data may be stated as:

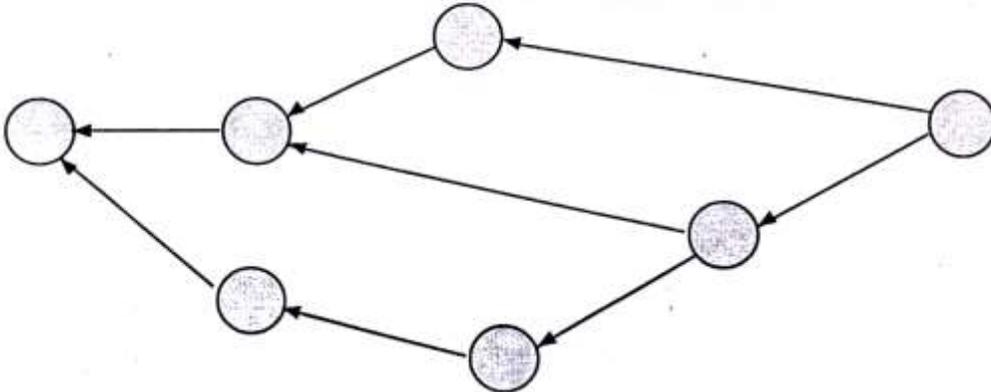
- (a) Identifying the data structures
- (b) Editing the data
- (c) Coding and classifying the data

- (d) Transcriptions of data
- (e) Tabulation of data.

**IDENTIFYING THE DATA STRUCTURE** : In the data preparation step, the data are prepared in a data format, which allows the analyst to use modern analysis software such as SAS or SPSS. The major criterion in this is to define the data structure. A data structure is a dynamic collection of related variables and can be conveniently represented as a graph whose nodes are labeled by variables. The data structure also defines and states the preliminary relationship between variables/groups of variables that have been preplanned by the researcher. Most data structures can be graphically presented to give clarity as to the framed research hypotheses. A simple structure could be a linear structure, in which one variable leads to the other and finally to the resultant end variable. Graphically, it may be shown as below :



Alternatively, a more serial data structure could be drawn as follows:



Alternatively, a more serial data structure could be drawn as follows:

The identification of the nodal points and the relationships among the nodes, could sometimes be a complex. When the task is complex, which involves several types of instruments being collected for the same research question, the procedure for drawing the data structure would involve a series of steps. In several intermediate steps, the heterogeneous data structures of the individual data sets can be harmonized to a common standard and the separate data sets are then integrated into a single data set. However, the clear definition of such data structures would help in the further processing of data.

**EDITING** : The next step in the processing of data is editing of the data. Editing is a process of checking to detect and correct errors and omissions. Data editing happens at two stages, one at the time of recording the data and second at the time of analysis of data.

### **Data Editing at the Time of Recording of Data**

Documented editing and testing of the data at the time of data recording is done considering the following questions in mind.

- Do the filters agree or are the data inconsistent?
- Have “missing values” been set to standardized values, which are the same for all research questions?
- Have variable descriptions been specified?
- Have labels for variable names and value labels been defined and written?

All editing and cleaning steps are documented, so that the redefinition of variables or later analytical modification requirements could be easily incorporated into the data sets.

### **Data Editing at the Time of Analysis of Data**

Data editing is also a requisite before the analysis of data is carried out. This ensures that the data is complete in all respect for subjecting them to further analysis. Some of the usual check list questions that can be had by a researcher for editing data sets before analysis would be:

- Is the coding frame complete?
- Is the documentary material sufficient for the methodological description of the study?
- Is the storage medium readable and reliable?
- Has the correct data set been framed?
- Is the number of cases correct?
- Are there differences between questionnaire, coding frame and data?
- Are there undefined and so-called “wild codes” or duplicate cases?
- Comparison of the first counting of the data with the original documents of the researcher.

**CODING AND CLASSIFICATION** : The edited data are then subject to codification and classification. Coding process assigns numerals or other symbols to the several responses of the data set. It is therefore a pre-requisite to prepare a coding scheme for the data set. The recording

of the data is done on the basis of this coding scheme. The responses collected in a data sheet varies, sometimes the responses could be the choice among a multiple response, sometimes the response could be in terms of values and sometimes the response could be alphanumeric. At the recording stage itself, if some codification were done to the responses collected, it would be useful in the data analysis. When codification is done, it is imperative to keep a log of the codes allotted to the observations. This code sheet will help in the identification of variables/observations and the basis for such codification.

The variables or observations in the primary instrument would also need codification, especially when they are categorized. The categorization could be on a scale i.e., most preferable to not preferable, or it could be very specific such as Gender classified as Male and Female. Certain classifications can lead to open-ended classification such as education classification, Illiterate, Graduate, Professional, Others, please specify. In such instances, the codification needs to be carefully done to include all possible responses under “Others, please specify”. If the preparation of the exhaustive list is not feasible, then it will be better to create a separate variable for the “Others, please specify” category and record all responses as such.

**Numeric coding** : Coding need not necessarily be numeric. It can also be alphabetic. Coding has to be compulsorily numeric, when the variable is to be subject to further parametric analysis.

**Alphabetic coding** : A mere tabulation or frequency count or graphical representation of the variable may be given an alphabetic coding.

**Zero coding** : A coding of zero has to be assigned carefully to a variable. In many instances, when manual analysis is done, a code of 0 would imply a “no response” from the respondents. Hence, if a value of 0 is to be given to a specific response in the data sheet, it should not lead to the same interpretation of “no response”. The coding sheet needs to be prepared carefully if the data recording is not done by the researcher, but is outsourced to a data entry firm or individual. In order to enter the data in the same perspective, as the researcher would like to view it, the data coding sheet is to be prepared first and a copy of the data coding sheet should be given to the outsourcer to help in the data entry procedure. Sometimes, the researcher might not be able to code the data from the primary instrument itself. He may need to classify the responses and then code them. For this purpose, classification of data is also necessary at the data entry stage.

## **Classification**

1. Most research studies result in a large volume of raw data which must be reduced into homogenous groups if we are to get meaningful relationships. This fact necessitates classification of data which happens to be the process of arranging data in groups or classes on the basis of common characteristics. Data having a common characteristic are placed in one class and in this way the entire data get divided into a number of groups or classes. Classification can be one of the following two types, depending upon the nature of the phenomenon involved :

(a) **Classification according to attributes** : As stated above, data are classified on the basis of common characteristics which can either be descriptive (such as literacy, sex, honesty. Etc.) or numerical (such as weight, height, income etc.) Descriptive characteristics refer to qualitative phenomenon which cannot be measured quantitatively; only their presence or absence in an individual item can be noticed. Data obtained this way on the basis of certain attributes are known as statistics of attributes and their classification is said to be classification according to attributes. Such classification can be simple classification or manifold classification. In simple classification we consider only one attribute and divide the universe into two classes one class consisting of items possessing the given attribute and the other class consisting of items which do not possess the given attribute. But in manifold classification we consider two or more attributes simultaneously, and divide that data into a number of classes (total number of classes of final order is given by  $2^n$ , where  $n$  = number of attributes considered). Whenever data are classified according to attributes, the researcher must see that the attributes are defined in such a manner that there is least possibility of any doubt/ambiguity concerning the said attributes.

(b) **Classification according to class-intervals** : Unlike descriptive characteristics, the numerical characteristics refer to quantitative phenomenon which can be measured through some statistical units. Data relating to income, production, age weight, etc. come under this category. Such data are known as statistics of variables and are classified on the basis of class intervals. For instance, persons whose incomes, say are within Rs. 201 to Rs.400 can form one group, those whose incomes are within Rs.401 to Rs. 600 can form another group and so no. in this way the entire data may be divided into a number of groups of classes or what are usually called, 'class-intervals. Each

group of class-interval, thus has an upper limit as well as a lower limit which are known as class limits. The difference between the two class limits is known as class magnitude. We may have classes with equal class magnitudes or with unequal class magnitudes. The number of items which fall in a given class is known as the frequency of the given class. All the classes or groups, with their respective frequencies taken together and put in the form of a table, are described as group frequency distribution or simply frequency distribution.

Classification is the process of arranging or categorizing data according to common characteristics or features possessed by items of data. The purpose of classification is to categories/group heterogeneous items of data into homogeneous classes/groups. Usually the bases used for classification are :

- i) Spatial or geography or area
- ii) Temporal or chronology or time
- iii) Quality or attribute
- iv) Quantity or magnitude

**TRANSCRIPTIONS OF DATA** : When the observations collected by the researcher are not very large, the simple inferences that can be drawn from the observations can be transferred to a data sheet, which is a summary of all responses on all observations from a research instrument. The main aim of transcription is to minimize the shuffling process between several responses and several observations. Suppose a research instrument contains 120 responses and the observations has been collected from 200 respondents, to prepare a simple summary of one response from all 200 observations would require a shuffling of 200 pages. The process is quite tedious if several summary tables are to be prepared from the instrument. The transcription process helps in the presentation of all responses and observations on data sheets, which can help the researcher to arrive at preliminary conclusions as to the nature of the sample collected etc. transcription is hence an intermediary process between data coding and data tabulation.

The researcher may adopt a manual or computerized transcription. Long work sheets, sorting cards or sorting strips could be used by the researcher to manually transcript the responses. The computerized transcription could be done using a data base package such as spreadsheets, text files or other databases.

## PRELIMINARIES FOR COMPUTERIZED DATA PROCESSING

When the sample size is large and/or when the variables studied is vast and interrelated, data can be transcribed to the computer for further easy processing. Many data base software packages are available using which data can be easily transcribed. The data base software shows a worksheet similar to the manual worksheet prepared above. Some important points that can be kept in mind while preparing the coding for computerized data entry are:

- a) **Use a ‘natural’ coding scheme** : Logical ordering of responses would be preferable. For instance, when the response is “very high” “more than once”, “only once” “nil”, the coding could be in the integer format of “4”, “3”, “2”, “1”. The alphabetical coding could also be used such as “H”, “F”, “O”, “N”. The use of numbers would help in further metric evaluations of that variable, while the alphanumeric codes would be amenable only for categorical analysis. If they are to be used as a metric measure, recoding of these variables into numeric format should be carried out using the software packages.
- b) **Avoid the use of blank space as a coding category**: Many computer software programs may not distinguish between blank space and zero unless otherwise specifically marked. Sometimes blank space may also be inferred as “no response”/ “no information” by the software packages. When using alphanumeric coding also the sue of blank space would differentiate it as an altogether separate response. Thus, in an alphanumeric coding the entry “KBALA” will be different from entry “KBALA”. The blank space between A to B makes it a unique response though the researcher might have erroneously used it.
- c) **Do not use the “-”, “+” symbols**: These would be taken as a numeric expression indicating negative and positive integers. Hence, avoidance of these symbols as a code is recommended.

## TABULATION

The transcription of data can be used to summarize and arrange the data in a compact form for further analysis. The process is called tabulation. Thus, tabulation is a process of summarizing raw data and displaying them on compact statistical table for further analysis. It involves counting the number of cases falling into each of the categories identified by the researcher.

In other words, arranging the classified data into vertical columns and horizontal rows is called tabulation.

Tabulation can be done manually or through the computer. The choice depends upon the size and type of study. Cost considerations, time pressures and the availability of software packages. Manual tabulation is suitable for small and simple studies.

### **Manual Tabulation**

When data are transcribed in a classified form as per the planned scheme of classification, category-wise totals can be extracted from the respective columns of the work sheets. A simple frequency table counting the number of “yes” and “No” responses can be made by easily counting the “y” response column and “N” response column in the manual worksheet table prepared earlier. This is one way frequency table and they are readily inferred from the totals of each column in the worksheet. Sometimes, the researcher has to cross tabulate two variables for instance the age group of vehicle owners. This requires a two-way classification and cannot be inferred straight from the worksheet. For this purpose, tally sheets are used. This process of tabulation is simple and does not require any technical knowledge or skill. If one wants to prepare a table showing the distributing of respondents by age, a tally sheet showing the age groups horizontally is prepared. Tally marks are then made for the respective group i.e., ‘vehicle owners’ from each line of response in the worksheet. After every four tallies, the fifth tally is cut across the previous four tallies. This represents a group of five items.

Computerized tabulation is easy with the help of software packages. The input requirement will be the column and row variables. The software package then computes the number of records in each cell of the row/ column categories. The most popular package is the statistical package for social sciences (SPSS). It is an integrated set of programs suitable for analysis of social science data. This package contains programs for a wide range of operation and analysis such as handling missing data, recoding, variable information, simple descriptive analysis, cross tabulation multivariate analysis and non-parametric analysis.

### **CONSTRUCTION OF FREQUENCY TABLE**

Frequency table provide a “shorthand” summary of data. The importance of presenting statistical data in tabular form needs no emphasis. Tables facilitate comprehending masses of data

at a glance, they conserve space and reduce explanations and descriptions to a minimum. They give a visual picture of relationships between variables and categories. They facilitate summation of items and the detection of errors and omissions and they provide a basis for computations.

It is important to make a distinction between the general purpose tables and specific tables. The general purpose tables are primary or reference tables designed to include large amounts of source data in convenient and accessible form. The special purpose tables are analytical or derivative ones that demonstrate significant relationships in the data or the results of statistical analysis. Tables in reports of government on population, vital statistics, agriculture, industries etc., are of general purpose type. They represent extensive repositories of statistical information. Special purpose tables are found in monographs, research reports and articles and are used as instruments of analysis. In research, we are primarily concerned with special purpose tables.

**Components of a table :** The major components of a table are :

**A. Heading**

- i. Table Number
- ii. Title of the Table
- iii. Designation of units

**B. Body**

- i. Stub-head : Heading of all rows or blocks of stub items
- ii. Body head : Headings of all columns or main captions and their sub- captions
- iii. Field/body : The cells in rows and columns

**C. Notations**

- i. Footnotes, wherever applicable
- ii. Source, wherever applicable

There are certain generally accepted principles or rules relating to construction of tables.

They are:

- (a) Every table should have a title. The title should represent a succinct description of the contents of the table. It should be clear and concise. It should be placed above the body of the table.

- (b) A number to facilitate easy reference should identify every table. The number can be centered above the title. The table numbers should run in a consecutive serial order. Alternatively tables in Chapter 1 be numbered as 1.1, 1.2, 1.,....., in Chapter 2 as 2.1, 2.2, 2.3 ..... and so on.
- (c) The captions (or column headings) should be clear and brief.
- (d) The units of measurement under each heading must always be indicated.
- (e) Any explanatory footnotes concerning the table itself are placed directly beneath the table and in order to obviate any possible confusion with the textual footnotes such reference symbols as the asterisk (\*) dagger (+) and the like may be used.
- (f) If the data in series of tables have been obtained from different sources, it is ordinarily advisable to indicate the specific sources in a place just below the table.
- (g) Usually lines separate columns from one another . lines are always drawn at the top and bottom of the table and below the captions.
- (h) The columns may be numbered to facilitate reference
- (i) All column figures should be properly aligned. Decimal points and ‘plus’ or ‘minus’ signs should be in perfect alignment.
- (j) Columns and rows that are to be compared with one another should be brought close together.
- (k) Totals of rows should be placed at the extreme right column and totals of columns at the bottom.
- (l) In order to emphasize the relative significance of certain categories, different kinds of type, spacing and identifications can be used.
- (m)The arrangement of the categories in a table may be chronological, geographical, alphabetical or according to magnitude. Numerical categories are usually arranged in descending order of magnitude.
- (n) Miscellaneous and exceptional items are generally placed in the last row of the table.
- (o) Usually the larger number of items is listed vertically. This means that a table’s length is more than its width.
- (p) Abbreviations should be avoided whenever possible and ditto marks should not be used in a table.
- (q) The table should be made as logical, clear accurate and simple as possible.

Text references should identify tables by number, rather than by such expressions as “the table above” or “the following table”. Tables should not exceed the page size by photo stating. Tables that are too wide for the page may be turned sidewise, with the top facing the left margin or binding of the script. Where should tables be placed in a research report or thesis? Some writers place both special purpose and general-purpose tables in an appendix and refer to them in the text by numbers. This practice has the disadvantage of inconveniencing the reader who wants to study the tabulated data as the text is read. A more appropriate procedure is to place special purpose tables in the text and primary tables, if needed at all, in an appendix.

### **ELEMENTS/TYPES OF ANALYSIS**

As stated earlier, by analysis we mean the computation of certain indices or measures along with searching for patterns of relationship that exist among the data groups. Analysis, particularly in case of survey or experimental data, involves estimating the values of unknown parameters of the population and testing of hypotheses for drawing inferences. Analysis may, therefore, be categorized as descriptive analysis and inferential analysis (inferential analysis is often known as statistical analysis). “Descriptive analysis is largely the study of distributions of one variable. This study provides us with profiles of companies, work groups, persons and other subjects on any of a multiple of characteristics such as size. Composition, efficiency, preferences, etc.” this sort of analysis may be in respect of one variable (described as unidimensional analysis), or in respect of two variables (described as bivariate analysis) or in respect of more than two variables (described as multivariate analysis). In this context we work out various measures that show the size and shape of a distribution (s) along with the study of measuring relationships between two or more variables.

Way may as well talk of correlation analysis and causal analysis. Correlation analysis studies the joint variation of two or more variables for determining the amount of correlation between two or more variables. Causal analysis is concerned with the study of how one or more variables affect changes in another variable. It is thus a study of functional relationships existing between two or more variables. This analysis can be termed as regression analysis. Causal analysis is considered relatively more important in

experimental researches, whereas in most social and business researches our interest lies in understanding and controlling relationships between variables then with determining causes per se and as such we consider correlation analysis as relatively more important.

In modern times, with the availability of computer facilities, there has been a rapid development of multivariate analysis which may be defined as “all statistical methods which simultaneously analyse more than two variables on a sample of observations. Usually the following analyses are involved when we make a reference of multivariate analysis :

- (a) **Multiple regression analysis** : This analysis is adopted when the researcher has one dependent variable which is presumed to be a function of two or more independent variables. The objective of this analysis is to make a prediction about the dependent variable based on its covariance with all the concerned independent variables.
- (b) **Multiple discriminant analysis** : This analysis is appropriate when the researcher has a single dependent variable that cannot be measured, but can be classified into two or more groups on the basis of some attribute. The object of this analysis happens to be to predict an entity’s possibility of belonging to a particular group based on several predictor variables.
- (c) **Multivariate analysis of variance (or multi-ANOVA)** : This analysis is an extension of two-way ANOVA, wherein the ratio of among group variance to within group variance is worked out on a set of variables.
- (d) **Canonical analysis** : This analysis can be used in case of both measurable and non-measurable variables for the purpose of simultaneously predicting a set of dependent variables from their joint covariance with a set of independent variables.

Inferential analysis is concerned with the various tests of significance for testing hypotheses in order to determine with what validity data can be said to indicate some conclusion or conclusions. It is also concerned with the estimation of population values. It is mainly on the basis of inferential analysis that the task of interpretation (i.e., the task of drawing inferences and conclusions is performed).

## **STATISTICS IN RESEARCH**

The role of statistics in research is to function as a tool in designing research, analyzing its data and drawing conclusions therefrom. Most research studies result in a large volume of raw data which must be suitably reduced so that the same can be read easily and can be used for further analysis. Clearly the science of statistics cannot be ignored by any research worker, even though he may not have occasion to use statistical methods in all their details and ramifications. Classification and tabulation, as stated earlier, achieve this objective to some extent, but we have to go a step further and develop certain indices or measures to summarise the collected/classified data. Only after this we can adopt the process of generalization from small groups (i.e., samples) to population. In fact, there are two major areas of statistics viz., descriptive statistics and inferential statistics. Descriptive statistics concern the development of certain indices from the raw data, whereas inferential statistics concern with the process of generalization. Inferential statistics are also known as sampling statistics and are mainly concerned with two major type of problems : (i) the estimation of population parameters, and (ii) the testing of statistical hypotheses.

The important statistical measures that are used to summarise the survey/research data are:

- 1) Measures of central tendency or statistical averages; 2) measures of dispersion; 3) measures of asymmetry (skewness); 4) measures of relationship; and 5) other measures.

Amongst the measures of central tendency, the three most important ones are the arithmetic average or mean, median and mode. Geometric mean and harmonic mean are also sometimes used.

From among the measures of dispersion, variance, and its square root the standard deviation are the most often used measures. Other measures such as mean deviation, range, etc. are also used. For comparison purpose, we use mostly the coefficient of standard deviation or the coefficient of variation.

In respect of the measures of skewness and kurtosis, we mostly use the first measure of skewness based on mean and mode or on mean and median. Other measures of skewness, based on quartiles or on the methods of moments, are also used sometimes, Kurtosis is also used to measure the peachiness of the curve of the frequency distribution.

Amongst the measures of relationship, Karl Pearson's coefficient of correlation is the frequently used measure in case of statistics of variables, whereas Yule's coefficient of association is used in case of statistics of attributes. Multiple correlation coefficient, partial correlation coefficient, regression analysis, etc., are other important measures often used by a researcher.

Index numbers, analysis of time series, coefficient of contingency, etc., are other measures that may as well be used by a researcher, depending upon the nature of the problem under study.

We give below a brief outline of some important measures (our of the above listed measures) often used in the context of research studies.

### **MEASURES OF CENTRAL TENDENCY**

Measures of central tendency (or statistical averages) tell us the point about which items have a tendency to cluster. Such a measure is considered as the most representative figure for the entire mass of data. Measure of central tendency is also known as statistical average. Mean, median and mode are the most popular averages. Mean, also known as arithmetic average, is the most common measure of central tendency and may be defined as the value which we get by dividing the total of the values of various given items in a series by the total number of items. We can work it out as under :

Mean is the simplest measurement of central tendency and is a widely used measure. Its chief use consists in summarizing the essential features of a series and in enabling data to be compared. It is amenable to algebraic treatment and is used in further statistical calculations. It is a relatively stable measure of central tendency. But it suffers from some limitations viz., it is unduly affected by extreme items; it may not coincide with the actual value of an item in a series, and it may lead to wrong impressions, particularly when the item values are not given with the average. However, mean is better than other averages, specially in economic and social studies where direct quantitative measurements are possible.

Median is the value of the middle item of series when it is arranged in ascending or descending order of magnitude. It divides the series into two halves; in one half all items

are less than median, whereas in the other half all items are less than median, whereas in the other half all items have values higher than median. If the values of the items arranged in the ascending order are : 60,74,80,90,95,100, then the value of the 4<sup>th</sup> item viz. 88 is the value of median. We can also write thus :

If we use assumed average A, then mean would be worked out as under:

Median is a positional average and is used only in the context of qualitative phenomena, for example, in estimation intelligence, etc., which are often encountered in sociological fields. Median is not useful where items need to be assigned relative importance and weights. It is not frequently used in sampling statistics.

Mode is the most commonly or frequently occurring value in a series. The mode in a distribution is that item around which there is maximum concentration. In general, mode is the size of the item which has the maximum frequency, but at items such an item may not be mode on account of the effect of the frequencies of the neighboring items. Like median, mode is a positional average and is not affected by the values of extreme items. It is, therefore, useful in all situations where we want to eliminate the effect of extreme variations. Mode is particularly useful in the study of popular sizes. For example, a manufacturer of shoes is usually interested in finding out the size most in demand so that he may manufacture a larger quantity of that size. In other words, he wants a modal size to be determined for median or mean size would not serve his purpose. But there are certain limitations of mode as well. For example, it is not amenable to algebraic treatment and sometimes remains indeterminate when we have two or more modal values in a series. It is considered unsuitable in cases where we want to give relative importance to items under consideration.

Geometric mean is also useful under certain conditions. It is defined as the  $n$ th root of the product of the values of  $n$  times in a given series. Symbolically, we can put it thus:

The most frequently used application of this average is in the determination of average per cent of change i.e., it is often used in the preparation of index numbers or when we deal in ratios.

Harmonic mean is defined as the reciprocal of the average of reciprocals of the values of items of a series. Symbolically, we can express it as under :

Harmonic mean is of limited application, particularly in cases where time and rate are involved. The harmonic mean gives largest weight to the smallest item and smallest weight to the largest item. As such it is used in cases like time and motion study where time is variable and distance constant.

From what has been stated above, we can say that there are several types of statistical averages. Researcher has to make a choice for some average. There are no hard and fast rules for the selection of a particular average in statistical analysis for the selection of an average mostly depends on the nature, type of objectives of the research study. One particular type of average cannot be taken as appropriate for all types of studies. The chief characteristics and the limitations of the various averages must be kept in view; discriminate use of average is very essential for sound statistical analysis.

### **MEASURES OF DISPERSION**

An averages can represent a series only as best as a single figure can, but it certainly cannot reveal the entire story of any phenomenon under study. Specially it fails to give any idea about the scatter of the values of items of a variable in the series around the true value of average. In order to measure this scatter, statistical devices called measures of dispersion are calculated. Important measures of dispersion are (a) range, (b) mean deviation, and (c) standard deviation.

(a) Range is the simplest possible measures of dispersion and is defined as the difference between the values of the extreme items of a series. Thus,

Range = (Highest value of an item in a series) – (Lowest value of an item in a series)

The utility of range is that it gives an idea of the variability very quickly, but the drawback is that range is affected very greatly by fluctuations of sampling. Its value is never stable, being based on only two values of the variable. As such, range is mostly used as a rough measure of variability and is not considered as an appropriate measure in serious research studies.

(b) Mean deviation is the average of difference of the values of items from some average of the series. Such a difference is technically described as deviation. In calculating mean deviation we ignore the minus sign of deviations while taking their total for obtaining the mean deviation. Mean deviation is, thus, obtained as under:

When mean deviation is divided by the average used in finding out the mean deviation itself, the resulting quantity is described as the coefficient of mean deviation. Coefficient of mean deviation is a relative measure of dispersion and is comparable to similar measure of other series. Mean deviation and its coefficient are used in statistical studies for judging the variability, and thereby render the study of central tendency of a series more precise by throwing light on the typicalness of an average. It is a better measure of variability than range as it takes into consideration the values of all items of a series. Even then it is not a frequently used measure as it is not amenable to algebraic process.

(c) Standard deviation is most widely used measure of dispersion of a series and is commonly denoted by the symbol ' $\sigma$ ' (pronounced as sigma). Standard deviation is defined as the square-root of the average of squares of deviations, when such deviations for the values of individual items in a series are obtained from the arithmetic average. It is worked out as under :

When we divide the standard deviation by the arithmetic average of the series, the resulting quantity is known as coefficient of standard deviation which happens to be a relative measure and is often used for comparing with similar measure of other series. When this coefficient of standard deviation is multiplied by 100, the resulting figure is known as coefficient of variation. Sometimes, we work out the square of standard deviation, known as variance, which is frequently used in the context of analysis of variation.

The standard deviation (along with several related measures like variance, coefficient of variation, etc.) is used mostly in research studies and is regarded as a very satisfactory measure of dispersion in a series. It is amenable to mathematical manipulation because the algebraic signs are not ignored in its calculation (as we ignore in case of mean deviation). It is less affected by fluctuations of sampling. These advantages make standard deviation and its coefficient a very popular measure of the scatteredness of a series. It is popularly used in the context of estimation and testing of hypotheses.

## **MEASURES OF ASYMMETRY (SKEWNESS)**

When the distribution of item in a series happens to be perfectly symmetrical, we then have the following type of curve for the distribution:

Such a curve is technically described as a normal curve and the relating distribution as normal distribution. Such a curve is perfectly bell shaped curve in which case the value of  $\bar{X}$  or  $M$  or  $Z$  is just the same and skewness is altogether absent. But if the curve is distorted (whether on the right side or on the left side), we have asymmetrical distribution which indicates that there is skewness. If the curve is distorted on the right side, we have positive skewness but when the curve is distorted towards left, we have negative skewness as shown here under:

Skwness is, thus, a measure of asymmetry and shows the manner in which the items are clustered around the average. In a symmetrical distribution, the items show a perfect balance on either side of the mode, but in a skew distribution the balance is thrown to one series. The difference between the mean, median or the mode provides an easy way of expressing skewness in a series. In case of positive skewness, we have  $Z < M > X$  and in case of negative skewness we have  $X < M < Z$  Usually we measure skewness in this way:

The significance of skewness lies in the fact that through it one can study the formation of series and can have the idea about the shape of the curve, whether normal or otherwise, when the items of a given series are plotted on a graph.

Kurtosis is the measure of flat-toppedness of a curve. A bell shaped curve or the normal curve is Mesokurtic because it is kurtic in the centre; but if the curve is relatively more peaked than the normal curve, it is called Leptokurtic whereas a curve is more flat than the normal curve, it is called Platykurtic. In brief, Kurtosis is the humpedness of the curve and points to the nature of distribution of items in the middle of a series.

It may be pointed out here that knowing the shape of the distribution curve is crucial to the use of statistical methods in research analysis since most methods make specific assumptions about the nature of the distribution curve.

## **MEASURES OF RELATIONSHIP**

So far we have dealt with those statistical measures that we use in context of univariate population i.e., the population consisting of measurement of only one variable. But if we have the data on two variables, we are said to have a bivariate population and if the data happen to be on more than two variables, the population is known as multivariate population. If for every measurement of a variable, X, we have corresponding value of a second variable, Y, the resulting pairs of values are called a bivariate population. In addition, we may also have a corresponding value of the third variable, Z, or the fourth variable, W, and so on, the resulting pairs of values are called a multivariate population. In case of bivariate or multivariate populations, we often wish to know the relation of the two and/or more variables in the data to one another. We may like to know, for example, whether the number of hours students devote for studies is somehow related to their family income, to age, to sex or to similar other factor. There are several methods of determining the relationship between variables, but no method can tell us for certain that a correlation is indicative of causal relationship. Thus we have to answer two types of questions in bivariate or multivariate populations viz.,

- i) Does there exist association or correlation between the two (or more) variables? If yes, of what degree?
- ii) Is there any cause and effect relationship between the two variables in case of the bivariate population or between one variable on one side and two or more variables on the other side in case of multivariate population? If yes, of what degree and in which direction?

The first question is answered by the use of correlation technique and the second question by the technique of regression. There are several methods of applying the two techniques, but the important ones are as under:

In case of bivariate population : Correlation can be studied through (a) cross tabulation; (b) Charles Spearman's coefficient of correlation; (c) Karl Pearson's coefficient of correlation; whereas cause and effect relationship can be studied through simple regression equations.

In case of multivariate population : Correlation can be studied through (a) coefficient of multiple correlation; (b) coefficient of partial correlation; whereas cause and effect relationship can be studied through multiple regression equations.

We can now briefly take up the above methods one by one.

Cross tabulation approach is specially useful when the data are in nominal form. Under it we classify each variable into two or more categories and then cross classify the variables in these sub-categories. Then we look for interactions between them which may be symmetrical, reciprocal or asymmetrical. A symmetrical relationship is one in which the two variables vary together, but we assume that neither variable is due to the other. A reciprocal relationship exists when the two variables mutually influence or reinforce each other. Asymmetrical relationship is said to exist if one variable (the independent variable) is responsible for another variable (the dependent variable). The cross classification procedure begins with a two-way table which indicates whether there is or there is not an interrelationship between the variables. This sort of analysis can be further elaborated in which case a third factor is introduced into the association through cross-classifying the three variables. By doing so we find conditional relationship in which factor X appears to affect factor Y only when factor Z is held constant. The correlation, if any, found through this approach is not considered a very powerful form of statistical correlation and accordingly we use some other methods when data happen to be either ordinal or interval or ratio data.

Charles Spearman's coefficient of correlation (or rank correlation) is the technique of determining the degree of correlation between two variables in case of ordinal data where ranks are given to the different values of the variables. The main objective of this coefficient is to determine the extent to which the two sets of ranking are similar or dissimilar. This coefficient is determined as under:

As rank correlation is a non-parametric technique for measuring relationship between paired observations of two variables when data are in the ranked form, we have dealt with this technique in greater details later on in the book in chapter entitled 'Hypotheses Testing II (Non-parametric tests)'.

Karl Pearson's coefficient of correlation (or simple correlation) is the most widely used method of measuring the degree of relationship between two variables. This coefficient assumes the following:

- i) That there is linear relationship between the two variables;

- ii) That the two variables are casually related which means that one of the variables is independent and the other one is dependent; and
- iii) A large number of independent causes are operating in both variables so as to produce a normal distribution.

Karl Pearson's coefficient of correlation (or  $r$ )

This is the short cut approach for finding ' $r$ ' in case of ungrouped data. If the data happen to be grouped data (i.e., the case of bivariate frequency distribution), we shall have to write Karl Person's coefficient of correlation as under:

Where  $f_{ij}$  is the frequency of a particular cell in the correlation table and all other values are defined as earlier.

Karl Pearson's coefficient of correlation is also known as the product moment correlation coefficient. The value of ' $r$ ' lies between  $\pm 1$ . Positive values of  $r$  indicate positive correlation between the two variables (i.e., changes in both variables take place in the same direction), whereas negative values of ' $r$ '. indicate negative correlation i.e., changes in the two variables taking place in the opposite directions. A zero value of ' $r$ ' indicates that there is no association between the two variables. When  $r = (+)1$ , it indicates perfect positive correlation and when it is  $(-)1$ , it indicates perfect negative correlation, meaning thereby that variations in independent variable (X) explain 100 % of the variations in the dependent variable (Y). We can also say that for a unit change in independent variable, if there happens to be a constant change in the dependent variable in the same direction, then correlation will be termed as perfect positive. But if such change occurs in the opposite direction, the correlation will be termed as perfect negative. The value of ' $r$ ' nearer to +1 or -1 indicates high degree of correlation between the two variables.

### **SIMPLE REGRESSION ANALYSIS**

Regression is the determination of a statistical relationship between two or more variables. In simple regression, we have only two variables, one variable (defined as independent) is the cause of the behavior of another one (defined as dependent variable). Regression can only interpret what exists physically i.e., there must be a physical way in which

independent variable X can affect dependent variable Y. the basic relationship between X and Y is given by

Where the symbol  $\hat{Y}$  denotes the estimated value of Y for a given value of X. this equation is known as the regression equation of Y on X (also represents the regression line of Y on X when drawn on a graph) which means that each unit change in X produces a change of b in Y, which is positive for direct and negative for inverse relationships.

Then generally used method to find the 'best' fit that a straight line of this kind can give is the least-square method. To use it efficiently, we first determine

Then

These measures define a and b which will give the best possible fit through the original X and Y points and the value of r can then be worked out as under :

Thus, the regression analysis is a statistical method to deal with the formulation of mathematical model depicting relationship amongst variables which can be used for the purpose of prediction of the values of dependent variable, given the values of the independent variable.

(Alternatively, for fitting a regression equation of the type to the given values of X and Y variables, we can find the values of the two constants viz., a and b by using the following two normal equations:

And then solving these equations for finding a and b values. Once these values are obtained and have been put in the equation we say that we have fitted the regression equation of Y on X to the given data. In a similar fashion, we can develop the regression equation of X and Y viz., presuming Y as an independent variable and X as dependent variable).

## **MULTIPLE CORRELATION AND REGRESSION**

When there are two or more than two independent variables, the analysis concerning relationship is known as multiple correlation and the equation describing such relationship as the multiple regression equation. We here explain multiple correlation and regression taking only two independent variables and one dependent variable (Convenient computer programs exist for dealing with a great number of variables). In this situation the results are interpreted as shown below:

Multiple regression equation assumes the form

Where  $X_1$  and  $X_2$  are two independent variables and  $Y$  being the dependent variable, and the constants  $a$ ,  $b_1$  and  $b_2$  can be solved by solving the following three normal equations:

(It may be noted that the number of normal equations would depend upon the number of independent variables. If there are 2 independent variables, then 3 equations, if there are 3 independent variables then 4 equations and so on, are used.)

In multiple regression analysis, the regression coefficients (viz.,  $b_1$   $b_2$ ) becomes reliable as the degree of correlation between the independent variables (viz.,  $X_1$ ,  $X_2$ ) increases. If there is a high degree of correlation between independent variables, we have a problem of what is commonly described as the problem of multicollinearity. In such a situation we should use only one set of the independent variable to make our estimate. In fact, adding a second variable, say  $X_2$ , that is correlated with the first variable, say  $X_1$ , distorts the values of the regression coefficients. Nevertheless, the prediction for the dependent variable can be made even when multicollinearity is present, but in such a situation enough care should be taken in selecting the independent variables to estimate a dependent variable so as to ensure that multi-collinearity is reduced to the minimum.

With more than one independent variable, we may make a difference between the collective effect of the two independent variables and the individual effect of each of them taken separately. The collective effect is given by the coefficient of multiple correlation,

### **PARTIAL CORRELATION**

Partial correlation measures separately the relationship between two variables in such a way that the effects of other related variables are eliminated. In other words, in partial correlation analysis, we aim at measuring the relation between a dependent variable and a particular independent variable by holding all other variables constant. Thus, each partial coefficient of correlation measures the effect of its independent variable on the dependent variable. To obtain it, it is first necessary to compute the simple coefficients of correlation between each set of pairs of variables as stated earlier. In the case of two independent variables, we shall have two partial correlation coefficients denote.

## QUESTIONS

1. Explain Data Processing



## **Chapter :7**

### **TESTING OF HYPOTHESIS**

Procedure for hypothesis testing, use of statistical techniques for testing of hypothesis.

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 generalization 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.

Hypothesis, 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 counseling will show a greater increase in creativity than students not receiving counseling” 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.

**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 untreatable 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.

## BASIC CONCEPTS CONCERNING TESTING OF HYPOTHESES

Basic concepts in the context of testing of hypotheses need to be explained.

- a) **Null hypothesis and alternative hypothesis:** In the context of statistical analysis, we often talk about null hypothesis and alternative hypothesis. If we are to compare method A with method B about its superiority and if we proceed on the assumption that both methods are equally good, then this assumption is termed as the null hypothesis. As against this, we may think that the method A is superior or the method B is inferior, we are then stating what is termed as alternative hypothesis. The null hypothesis is generally symbolized as  $H_0$  and the alternative hypothesis as  $H_a$ . Suppose we want to test the hypothesis that the population mean ( $\mu$ ) is equal to the hypothesized mean ( $\mu_{H_0} = 100$ ).

Then we would say that the null hypothesis is that the population mean is equal to the hypothesized mean 100 and symbolically we can express as:

$$H_0 : \mu = \mu_{H_0} = 100$$

If our sample results do not support this null hypothesis, we should conclude that something else is true. What we conclude rejecting the null hypothesis is known as alternative hypothesis. In other words, the set of alternatives to the null hypothesis is referred to as the alternative hypothesis. If we accept  $H_0$ , then we are rejecting  $H_a$  and if we reject  $H_0$ , then we are accepting  $H_a$ . for  $H_0 : \mu = \mu_{H_0} = 100$ , we may consider three possible alternative hypotheses as follows:

Alternative hypothesis	To be read as follows
$H_a : \mu \neq \mu_{H_0}$	(The alternative hypothesis is that the population mean is not equal to 100 i.e., it may be more or less than 100)
$H_a : \mu > \mu_{H_0}$	(The alternative hypothesis is that the population mean is greater than 100)
$H_a : \mu < \mu_{H_0}$	(The alternative hypothesis is that the population mean is less than 100)

The null hypothesis and the alternative hypothesis are chosen before the sample is drawn (the researcher must avoid the error of deriving hypotheses from the data that he collects and then testing the hypotheses from the same data). In the choice of null hypothesis, the following considerations are usually kept in view:

- a) Alternative hypothesis is usually the one which one wishes to prove and the null hypothesis is the one which one wishes to disprove. Thus, a null hypothesis represents the hypothesis we are trying to reject, and alternative hypothesis represents all other possibilities.
- b) If the rejection of a certain hypothesis when it is actually true involves great risk, it is taken as null hypothesis because then the probability of rejecting it when it is true is  $\alpha$  (the level of significance) which is chosen very small.
- c) Null hypothesis should always be specific hypothesis i.e., it should not state about or approximately a certain value.

Generally, in hypothesis testing we proceed on the basis of null hypothesis, keeping the alternative hypothesis in view. Why so? The answer is that on the assumption that null hypothesis is true, one can assign the probabilities to different possible sample results, but this cannot be done if we proceed with the alternative hypothesis. Hence the use of null hypothesis (at times also known as statistical hypothesis) is quite frequent.

**(b) 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.

**(c) Decision rule or test of hypothesis:** Given a hypothesis  $H_0$  and an alternative hypothesis  $H_a$ , we make a rule which is known as decision rule according to which we accept  $H_0$  (i.e., accept  $H_a$ ). for instance, if ( $H_0$  is that a certain lot is good (there are very few defective items in it) against  $H_a$ )

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  $H_0$  otherwise we will reject  $H_0$  (or accept  $H_a$ ). this sort of basis is known as decision rule.

**(d) 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  $\alpha$  (alpha) known as  $\alpha$  error, also called the level of significance of test; and Type II error is denoted by  $\beta$  (beta) known as  $\beta$  error. In a tabular form the said two errors can be presented as follows:

	Decision	
	Accept $H_0$	Reject $H_0$
$H_0$ (true)	Correct decision	Type I error ( $\alpha$ error)
$H_0$ (false)	Type II error ( $\beta$ error)	Correct decision

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 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.

**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  $\alpha$  (alpha) known as  $\alpha$  error, also called the level of significance of test; and Type II error is denoted by  $\beta$  (beta) known as  $\beta$  error. In a tabular form the said two errors can be presented as follows:

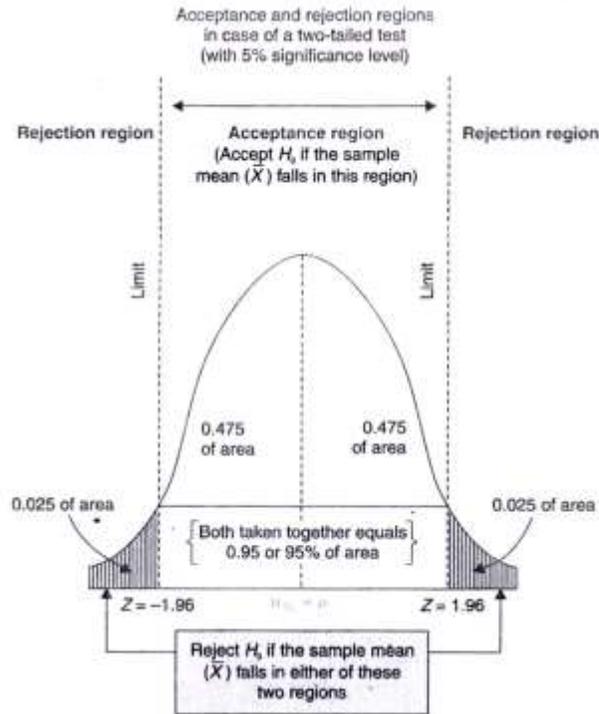
	Decision	
	Accept $H_0$	Reject $H_0$
$H_0$ (true)	Correct decision	Type I error ( $\alpha$ error)
$H_0$ (false)	Type II error ( $\beta$ error)	Correct decision

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 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 hypothesized 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. Symbolically, the two-tailed test is appropriate when we have  $H_0: \mu = \mu_0$  and  $H_a: \mu \neq \mu_0$  which may mean  $\mu > \mu_0$  or  $\mu < \mu_0$ . thus, in a two-tailed test, there are two rejection regions, one on each tail of the curve which can be illustrated as under:



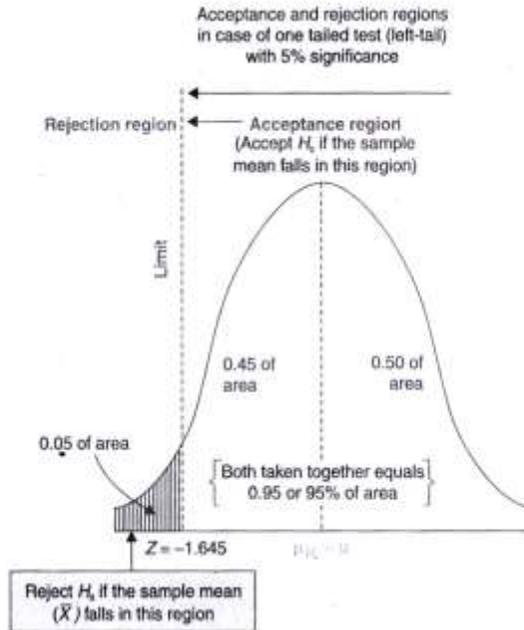
Mathematically we can state:

Acceptance Region A :  $|Z| \leq 1.96$

Rejection Region R :  $|Z| \geq 1.96$

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  $\mu$ , 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 hypothesized value. For instance, if our  $H_0 : \mu = \mu_0$  and  $H_a : \mu > \mu_0$ , then we are interested in what is known as right tailed test (wherein there is one rejection region only on the right tail) which can be illustrated as below:



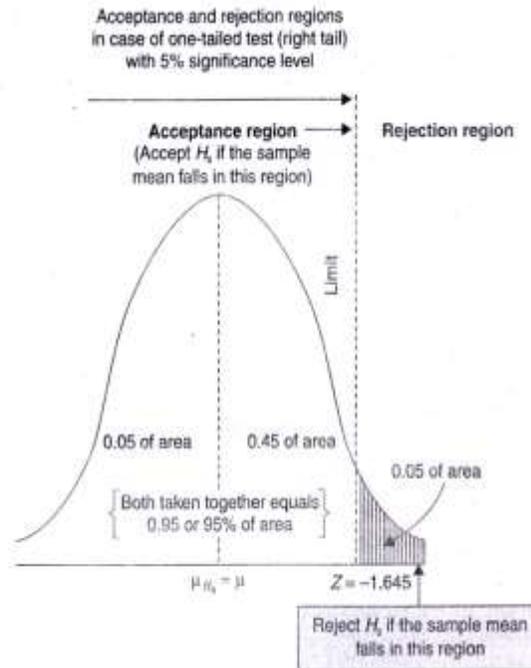
Mathematically we can state:

Acceptance Region A :  $Z > - 1.645$

Rejection Region R :  $Z \leq - 1.645$

If our  $\mu = 100$  and if our sample mean deviates significantly from 100 in the lower direction, we shall reject  $H_0$ , otherwise we shall accept  $H_0$  at a certain level of significance. If the significance level in the given case is kept at 5%, then the rejection region will be equal to 0.05 of area in the left tail as has been shown in the above curve.

In case our  $H_0 : \mu = \mu H_0$  and  $H_a : \mu \neq \mu H_0$ , we are then interested in what is known as one-tailed test (right tail) and the rejection region will be on the right tail of the curve as shown below :



Mathematically we can state:

Acceptance Region A :  $|Z| \leq 1.645$

Rejection Region R :  $|Z| > 1.645$

If our  $\mu = 100$  and if our sample mean deviates significantly from 100 in the upward direction, we shall reject  $H_0$ , otherwise we shall accept the same. If in the given case the significance level is kept at 5%, then the rejection region will be equal to 0.05 of area in the right-tail as has been shown in the above curve.

It should always be remembered that accepting  $H_0$  on the basis of sample information does not constitute the proof that  $H_0$  is true. We only mean that there is no statistical evidence to reject it, but we are certainly not saying that  $H_0$  is true (although we behave as if  $H_0$  is true).

## 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:

**i) Making a formal statement:** The step consists in making a formal statement of the null hypothesis ( $H_0$ ) and also of the alternative hypothesis ( $H_a$ ). 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  $H_0 : \mu = 10$  tons

Alternative Hypothesis  $H_a : \mu > 10$  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  $H_0 : \mu = 80$

Alternative Hypothesis  $H_a : \mu \neq 80$

The formulation of hypotheses is an important step which must 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  $H_a$  is of the type greater than ( or of the type lesser than), we use a one-tailed test, but when  $H_a$  is of the type “ whether greater or smaller” them we use a two-tailed test.

**(ii) 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 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.

**(iii) 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.

**(iv) 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.

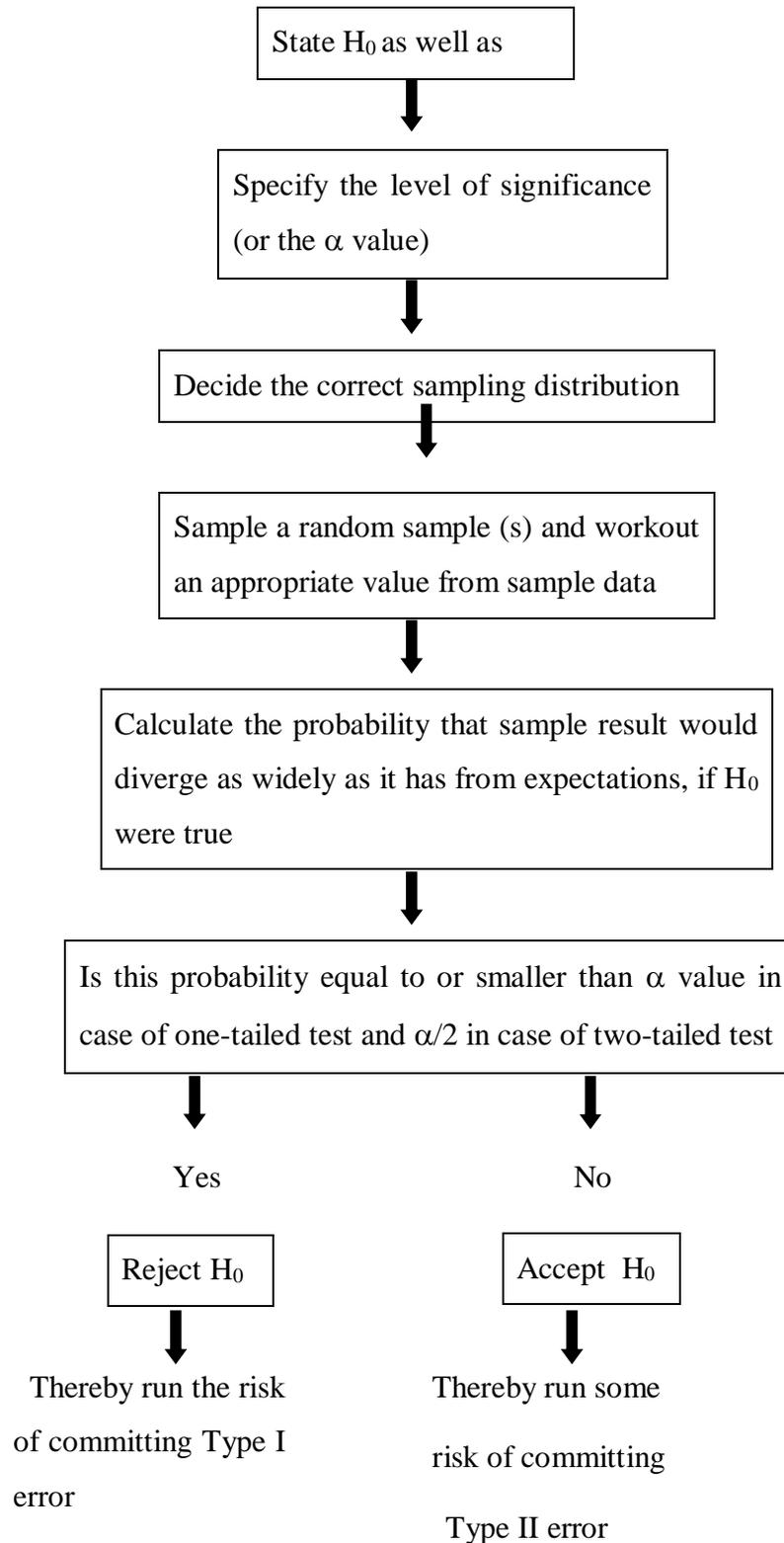
**(v) Calculation of the probability:** One has then to calculate the probability that the sample result would diverge as widely as it has from expectations, if the null hypothesis were in fact true.

**(vi) Comparing the probability:** Yet another step consists in comparing the probability thus calculated with the specified value for  $\alpha$  , the significance level. If the calculated probability is equal to or smaller than the  $\alpha$  value in case of one-tailed test (and  $\alpha/2$  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  $H_0$  , we run a risk of (at most the level of significance) committing an error of Type I, but if we accept  $H_0$  , then we run some risk (the size of which cannot be specified as long as the  $H_0$  happens to be vague rather than specific) of committing an error of Type II.

### **FLOW DIAGRAM FOR HYPOTHESIS TESTING**

The above stated general procedure for hypothesis testing can also be depicted in the form of a flow-chart for better understanding as shown below:

## FLOW DIAGRAM FOR HYPOTHESIS TESTING



## MEASURING THE POWER OF A HYPOTHESIS TEST

As stated above we may commit Type I and Type II errors while a hypothesis. The probability of Type I error is denoted as  $\alpha$  (the significance level of the test) and the probability of Type II error is referred to as  $\beta$ . Usually the significance level of a test is assigned in advance and once we decide it, there is nothing else we can do about  $\alpha$ . But what can we say about  $\beta$ ? We all know that a 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  $\beta$  (the probability of accepting  $H_0$  when  $H_0$  is not true) to be as small as possible. Alternatively, we would like that  $1 - \beta$  (the probability of rejecting  $H_0$  when  $H_0$  is not true) to be as large as possible. If  $1 - \beta$  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  $H_0$  when it is not true and if  $1 - \beta$  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  $1 - \beta$  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  $1 - \beta$  for each possible value of the population parameter (say  $\mu$ , the true population mean) for which the  $H_0$  is not true (alternatively the  $H_a$  is true), the resulting curve is known as the power curve associated with the given test. Thus power curve of a hypothesis 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 hypothesized value of the population parameter, the power of the test (i.e.,  $1 - \beta$ ) must get closer and closer to the probability of rejecting  $H_0$  when the population parameter is exactly equal to hypothesized value of the parameter. We know that this probability is simply the significance level of the test, and such the power curve of a test terminates at a point that lies at a height of  $\alpha$  (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.

## **TEST OF HYPOTHESES**

Has been stated above that hypothesis testing determines the validity of the assumption (technically described as null hypothesis) with a view to choose between two conflicting hypotheses about the of a population parameter. Hypothesis testing helps to decide on the basis of a sample data, other a hypothesis about the population is likely to be true or false. Statisticians have developed tests of hypotheses (also known as the tests of significance) for the purpose of testing of which can be classified as : (a) Parametric tests or standard tests of hypotheses; and Non-parametric tests or distribution-free test of hypotheses.

Parametric tests usually assume certain properties of the parent population from which we draw 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 tests can be used. But there are situations when the researcher cannot or does not want such assumptions. In such situations we use statistical methods for testing hypotheses which called non-parametric tests because such tests do not depend on any assumption about the of the parent population. Besides, most non-parametric tests, assume only nominal or data, whereas parametric tests require measurement equivalent to at least an interval scale. Result, non-parametric tests need more observations than parametric tests to achieve the same of Type I and Type II errors.

## **IMPORTANT PARAMETRIC TESTS**

Important parametric tests are :

- (1) z-test;
- (2) t-test;
- (3)  $X^2$  -test, and
- (4) F-test.

All these tests 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 hypothesized 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 samples ( $s$ ) 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.

**$\chi^2$ -test** is based on chi-square distribution and as a parametric test is used for comparing a sample variance to a theoretical population variance.

**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.

### QUESTIONS

1. What is a hypothesis ? What is meant by testing?
2. What is hypothesis testing ? What is the logic of hypothesis testing?
3. Explain type I and type II errors
4. Describe two-tailed and one tailed tests.
5. Explain the procedure for hypothesis testing
6. Describe the various parametric test in hypothesis testing

## **Chapter : 8**

### **INTERPRETATION OF DATA**

Techniques of interpretation, Report writing, Layout of a project report, preparing research reports.

After collecting and analyzing the data, the researcher has to accomplish the task of drawing inferences followed by report writing. This has to be done very carefully, otherwise misleading conclusions may be drawn and the whole purpose of doing research may get vitiated. It is only through interpretation that the researcher can expose relations and processes that underlie his findings. In case of hypotheses testing studies, if hypotheses are tested and upheld several times, the researcher may arrive at generalizations. But in case the researcher had no hypothesis to start with, he would try to explain his findings on the basis of some theory. This may at times result in new questions, leading to further researches. All this analytical information and consequential inference(s) may well be communicated, preferably through research report, to the consumers of research results who may be either an individual or a group of individuals or some public/private organization.

#### **MEANING OF INTERPRETATION**

Interpretation refers to the task of drawing inferences from the collected facts after an analytical and/or experimental study. In fact, it is a search for broader meaning of research findings. The task of interpretation has two major aspects viz., (i) the effort to establish continuity in research through linking the results of a given study with those of another, and (ii) the establishment of some explanatory concepts. "In one sense, interpretation is concerned with relationships within the collected data, partially overlapping analysis. Interpretation also extends beyond the data of the study to include the results of other research, theory and hypotheses." Thus, interpretation is the device through which the factors that seem to explain what has been observed by researcher in the course of the study can be better understood and it also provides a theoretical conception which can serve as a guide for further researches.

Interpretation is essential for the simple reason that the usefulness and utility of research findings lie in proper interpretation. It is being considered a basic component of research process because of the following reasons:

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

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

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

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

The task of interpretation is not an easy job, rather it requires a great skill and dexterity on the part of researcher. Interpretation is an art that one learns through practice and experience. The researcher may, at times, seek the guidance from experts for accomplishing the task of interpretation. The technique of interpretation often involves the following steps:

(i) Researcher must give reasonable explanations of the relations which he has found and he must interpret the lines of relationship in terms of the underlying processes and must try to find out the thread of uniformity that lies under the surface layer of his diversified research findings. In fact, this is the technique of how generalization should be done and concepts be formulated.

(ii) Extraneous information, if collected during the study, must be considered while interpreting the final results of research study, for it may prove to be a key factor in understanding the problem under consideration.

(iii) It is advisable, before embarking upon final interpretation, to consult someone having insight into the study and who is frank and honest and will not hesitate to point out omissions and errors in logical argumentation. Such a consultation will result in correct interpretation and, thus, will enhance the utility of research results.

(iv) Researcher must accomplish the task of interpretation only after considering all relevant factors affecting the problem to avoid false generalization. He must be in no hurry while interpreting results, for quite often the conclusions, which appear to be all right at the beginning, may not at all be accurate.

One should always remember that even if the data are properly collected and analyzed, wrong interpretation would lead to inaccurate conclusions. It is, therefore, absolutely essential that the task of interpretation be accomplished with patience in an impartial manner and also in correct perspective. Researcher must pay attention to the following points for correct interpretation:

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

The researcher must remain cautious about the errors that can possibly arise in the process of interpreting results. Errors can arise due to false generalization and/or due to wrong interpretation of statistical measures, such as the application of findings beyond the range of observations, identification of correlation with causation and the like. Another major pitfall is the tendency to affirm that definite relationships exist on the basis of confirmation of particular hypotheses. In fact, the positive test results accepting the hypothesis must be interpreted as "being in accord" with the hypothesis, rather than as "confirming the validity of the hypothesis". The researcher must remain vigilant about all such things so that false generalization may not take place. He should be well equipped with and must know the correct use of statistical measures for drawing inferences concerning his study.

He must always keep in view that the task of interpretation is very much intertwined with analysis and cannot be distinctly separated. As such he must take the task of interpretation as a special aspect of analysis and accordingly must take all those precautions that one usually observes while going through the process of analysis viz., precautions concerning the reliability of data, computational checks, validation and comparison of results.

He must never lose sight of the fact that his task is not only to make sensitive observations of relevant occurrences, but also to identify and disengage the factors that are initially hidden to the eye. This will enable him to do his job of interpretation on proper lines. Broad generalization should be avoided as most research is not amenable to it because the coverage may be restricted to a particular time, a particular area and particular conditions. Such restrictions, if any, must invariably be specified and the results must be framed within their limits.

The researcher must remember that "ideally in the course of a research study, there should be constant interaction between initial hypothesis, empirical observation and theoretical conceptions. It is exactly in this area of interaction between theoretical orientation and empirical observation that

opportunities for originality and creativity lie."<sup>2</sup> He must pay special attention to this aspect while engaged in the task of interpretation.

## **RESEARCH REPORT**

A research report is a formal statement of research process and its results. It narrates the problem studied, methods used for studying it and the findings and conclusions of the study. The purpose of research report is to communicate to interested person the methodology and the results of the study in such a manner as to enable them to understand the research process and to determine the validity of the conclusions. The aim of the report is not to convince the reader of the value of the result, but to convey to him what was done, why it was done, and what was its outcome. It is so written that the reader himself can reach his own conclusions as to the adequacy of the study and the validity of the reported results and conclusions.

### **Characteristics of a Report**

A research report is a narrative but authoritative document on the outcome of research effort. It presents a highly specific information for a clearly designated audience. It is no persuasive as a form of communication. Extra caution is shown in advocating a course of action even if the finding point to it. Presentation is subordinated to the matter being presented. It is simple, readable and accurate form of communication.

### **Functions of a Research Report**

A well written research report performs several functions.

It serves as a means for presenting the problem studied, methods and techniques used for collecting and analysing data, the findings, conclusions and recommendations in an organized manner.

It serves as a basic reference material for future use in developing research proposals in the same or related area.

A report serves as a means for judging the quality of the completed research project.

It is a means for evaluating the researcher's ability and competence to do research.

It provides factual base for formulating policies and strategies relating to the subject matter studied.

It provides systematic knowledge on problems and issues analysed

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 generalizations 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. Research results must invariably enter the general store of knowledge. All this explains the significance of writing research report. 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.

### **DIFFERENT STEPS IN WRITING REPORT**

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;
- (e) preparation of the final bibliography; and
- (f) writing the final draft.

Though all these steps are self explanatory, yet 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 organization 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 generalizations 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 moldering cement and loose brick."<sup>4</sup> 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.

***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.

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.

Let us deal with them separately.

### **Preliminary Pages**

Preliminary pages 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 of illustrations* so that the decision-maker or anybody interested in reading the report can easily locate the required information in the report.

### **Main Text**

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.

(i) *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? 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 generalizability 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 limitation, under which the research project was completed, must also be narrated.

(ii) *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 understood by all concerned. If the findings happen to be extensive, at this point they should be put in the summarized form.

(iii) *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. This 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 o the problem, the conclusions at which he arrived, and the bases for his conclusions."

(iv) *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 apply in similar circumstances.

(b) The conditions of the present study which may limit the extent of legitimate generalization 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 summarizes a 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 the probable future of the subject and an indication of the kind of research which needs to be done that particular field is useful and desirable.

(v) *Summary*: It has become customary to conclude the research report with a very brief summa resting in brief the research problem, the methodology, the major findings and the major conclusion drawn from the research results.

### **(C) End Matter**

At the end of the report, appendices should be enlisted in respect of all technical data such 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 v the numbers of the pages in a book or report on which they are mentioned or discussed) she 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.

### **TYPES OF RESEARCH REPORTS**

Research reports vary greatly in length and type. In each individual case, both the length and the form are largely dictated by the problems at hand. For instance, business firms prefer reports in the letter form, just one or two pages in length. Banks, insurance organizations and financial institutions : generally fond of the short balance-sheet type of tabulation for their annual reports to their customers and shareholders. Mathematicians prefer to write the results of their investigations in the form of algebraic notations. Chemists report their results in symbols and formulae. Students of literature usually write long reports presenting the critical analysis of some writer or period or the like with a liberal use of quotations from the works of the author under

discussion. In the field of education and psychology, the favourite form is the report on the results of experimentation accompanied by the detailed statistical tabulations. Clinical psychologists and social pathologists frequently find it necessary to make use of the case-history form. News items in the daily papers are also forms of report writing. They represent firsthand on-the-scene accounts of the events described or compilations of interviews with persons who were on the scene. In such reports the first paragraph usually contains the important information in detail and the succeeding paragraphs contain material which is progressively less and less important.

Book-reviews which analyze the content of the book and report on the author's intentions, his success or failure in achieving his aims, his language, his style, scholarship, bias or his point of view. Such reviews also happen to be a kind of short report. The reports prepared by governmental bureaus, special commissions, and similar other organizations are generally very comprehensive reports on issues involved. Such reports are usually considered as important research products.

### **TECHNICAL REPORT**

In the technical report the main emphasis is on (i) the methods employed, (ii) assumptions made in course of the study, (iii) the detailed presentation of the findings including their limitations and supporting data.

A general outline of a technical report can be as follows:

*Summary of results:* A brief review of the main findings just in two or three pages.

*Nature of the study:* Description of the general objectives of study, formulation of the problem in operational terms, the working hypothesis, the type of analysis and data required, etc.

*Methods employed:* Specific methods used in the study and their limitations. For instance, in sampling studies we should give details of sample design viz., sample size, sample selection, etc.

### **L PRESENTATION**

At times oral presentation of the results of the study is considered effective, particularly in cases where policy recommendations are indicated by project results. The merit of this approach lies in the fact that it provides an opportunity for give-and-take decisions which generally lead to a better understanding of the findings and their implications. But the main demerit of this sort of presentation is the lack of any permanent record concerning the research details and it may be just possible that the findings may fade away from people's memory even before an action is taken. In order to overcome this difficulty, a written report may be circulated before the oral presentation and referred frequently during the discussion. Oral presentation is

effective when supplemented by various visual devices. Use of slides, wall charts and blackboards is quite helpful in contributing to clarity and in reducing the boredom, if any. Distributing a board outline, with a few important tables and charts concerning the research results, makes the listeners attentive who have a ready outline on which to focus their thinking. This very often happens in academic institutions where the researcher discusses his research findings and policy implications with others either in a seminar or in a group discussion. Thus, research results can be reported in more than one way, but the usual practice adopted, in academic institutions particularly, is that of writing the Technical Report and then preparing several research papers to be discussed at various forums in one form or the other. But in practical field and with problems having policy implications, the technique followed is that of writing a popular report. Researches done on governmental account or on behalf of some major public or private organizations are usually presented in the form of technical reports.

There are very definite and set rules which should be followed in the actual preparation of

The research report or paper. Once the techniques are finally decided, they should be scrupulously adhered to and no deviation permitted. The criteria of format should be decided as soon as the materials for the research paper have been assembled. The following points deserve mention so far as the mechanics of writing a report are concerned:

*Size and physical design:* The manuscript should be written on unruled paper. If it is to be written by hand, then black or blue-black ink should be used. A margin of at least one-half inches should be allowed at the left hand and of at least half an inch at the right hand of the paper. There should also be one-inch margins, top and bottom. The paper should be neat and legible. If the manuscript is to be typed, then all typing should be double-spaced on one side of the page only except for the insertion of the long quotations.

*Procedure :* Various steps in writing the report should be strictly adhered.

*Layout:* Keeping in view the objective and nature of the problem, the layout of the report should be thought of and decided and accordingly adopted.

*Treatment of quotations:* Quotations should be placed in quotation marks and double spaced, forming an immediate part of the text.

*The footnotes:* Regarding footnotes one should keep in view the followings:

(a) The footnotes serve two purposes viz., the identification of materials used in quotations in the report and the notice of materials not immediately necessary to the body of the research text but still of supplemental value. In other words, footnotes are meant for cross reference citation of authorities and sources, acknowledgement and elucidation or explanation point of view. It should always be kept in view that footnote is not an end nor a means of the display of scholarship. The modern tendency is to make the minimum use of footnotes for scholarship does not need to be displayed.

(b) Footnotes are placed at the bottom of the page on which the reference or quotation which they identify or supplement ends. Footnotes are customarily separated from the material by a space of half an inch and a line about one and a half inches long.

(c) Footnotes should be numbered consecutively, usually beginning with 1 in each chapter separately. The number should be put slightly above the line, say at the end of a quotation. At the foot of the page, again, the footnote number should be indented and typed above the line. Thus, consecutive numbers must be used to correlate the reference in the text with its corresponding note at the bottom of the page, except in case of statistical tables and other numerical material, where symbols such as the asterisk (\*) or the like one may be used to prevent confusion.

(d) Footnotes are always typed in single space though they are divided from one another by double space.

*Documentation style:* Regarding documentation, the first footnote reference to any given should be complete in its documentation, giving all the essential facts about the edition used documentary footnotes follow a general sequence.

*Punctuation and abbreviations in footnotes:* The first item after the number in the footnote is the author's name, given in the normal signature order. This is followed by a comma. After the comma, the title of the book is given: the article (such as "A", "An", "The" etc.) is omitted and only the first word and proper nouns and adjectives are capitalized. The title is followed by a comma. Information concerning the edition is given next. This entry is followed by a comma. The place of publication is then stated; it may be mentioned in an abbreviated form, if the place happens to be a famous one such as Lond. for London, N.Y. for New York, N.D. for New Delhi and so on. This entry is followed by a comma. Then the name of the publisher is mentioned and this entry is closed by a comma. It is followed by the date of publication if the date is given on the title page. If the date appears in the copyright notice on the reverse side of the title page or elsewhere in the volume, the comma should be omitted and the date enclosed in square brackets [c 1978], [1978]. The entry is followed by a comma. Then follow the volume and page references and are separated by a

comma if both are given. A period closes the complete documentary reference. But one should remember that the documentation regarding acknowledgements from magazine articles and periodical literature follow a different form as stated earlier while explaining the entries in the bibliography.

*Use of statistics, charts and graphs:* A judicious use of statistics in research reports is often considered a virtue for it contributes a great deal towards the clarification and simplification of the material and research results. One may well remember that a good picture is often worth more than a thousand words. Statistics are usually presented in the form of tables, charts, bars and line-graphs and pictograms. Such presentation should be self explanatory and complete in itself. It should be suitable and appropriate looking to the problem at hand. Finally, statistical presentation should be neat and attractive.

*The final draft:* Revising and rewriting the rough draft of the report should be done with great care before writing the final draft. For the purpose, the researcher should put to himself questions like: Are the sentences written in the report clear? Are they grammatically correct? Do they say what is meant? Do the various points incorporated in the report fit together logically? "Having at least one colleague read the report just before the final revision is extremely helpful. Sentences that seem crystal-clear to the writer may prove quite confusing to other people; a connection that had seemed self evident may strike others as a *non-sequitur*. A friendly critic, by pointing out passages that seem unclear or illogical, and perhaps suggesting ways of remedying the difficulties, can be an invaluable aid in achieving the goal of adequate communication."<sup>6</sup>

10. *Bibliography:* Bibliography should be prepared and appended to the research report as discussed earlier.

11. *Preparation of the index:* At the end of the report, an index should invariably be given, the value of which lies in the fact that it acts as a good guide, to the reader. Index may be prepared both as subject index and as author index. The former gives the names of the subject-topics or concepts along with the number of pages on which they have appeared or discussed in the report, whereas the latter gives the similar information regarding the names of authors. The index should always be arranged alphabetically. Some people prefer to prepare only one index common for names of authors, subject-topics, concepts and the like ones.

Research report is a channel of communicating the research findings to the readers of the report. A good research report is one which does this task efficiently and effectively. As such it must be prepared keeping the following precautions in view:

While determining the length of the report (since research reports vary greatly in length), one should keep in view the fact that it should be long enough to cover the subject but short enough to maintain interest. In fact, report-writing should not be a means to learning more and more about less and less.

A research report should not, if this can be avoided, be dull; it should be such as to sustain reader's interest.

Abstract terminology and technical jargon should be avoided in a 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 expressions such as "it seems," "there may be" and the like.

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

The layout of the report should be well thought out and must be appropriate and in accordance with the objective of the research problem.

The reports should be free from grammatical mistakes and must be prepared strictly in accordance with the techniques of composition of report-writing such as the use of quotations, footnotes, documentation, proper punctuation and use of abbreviations in footnotes and the like.

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.

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.

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

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

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

Report must be attractive in appearance, neat and clean, whether typed or printed.

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

Objective of the study, the nature of the problem, the methods employed and the analysis techniques adopted must all be clearly stated in the beginning of the report in the form of introduction.

## QUESTIONS

Explain The Concept Of Interpretation Of Data In Research.

Explain The Need Of Interpretation In Research.

Explain The Technique Of Interpretation

What Are The Precautions That To Be Considered In Interpretation

Explain The Significance Of Report Writing

Define A Research Report And Explain Its Purpose.

What Are The Characteristics Of Research Report? What Functions Does This Report Perform ?

Describe The Layout Of The Research Report

Explain The Types Of Research Reports

Describe The Mechanics Of Writing A Research Report

Explain The Various Precautions That To Be Considered In Writing Research Reports



### **DR. SHINEY CHIB**

Dr. Shiney Chib is working as Director & Research Head with Datta Meghe Institute of Management Studies, Nagpur. She is appointed as adjunct faculty with SEGI-University, Malaysia. She has been conferred with 'Distinguished Educator Award' by 'Discovery Education Media', 'Rajiv Gandhi Excellence Award in the field of Education' and 'Mahila Ratan Gold Medal for Academic Excellence' and 'Eminent Educationist Award'. She has also received 'Academic Excellence Award' from NIPM Nagpur chapter. Eight students have been awarded PhD under her guidance. She has industrial experience of 11 years and 15 years of teaching experience. She has many publications in the reputed journals and presented research paper and also chaired the sessions in the various conferences held in India and Abroad.



# **EURO ASIA**

**RESEARCH AND DEVELOPMENT ASSOCIATION**

**IC/14, RAMESH NAGAR**

**DELHI - 110 015**

[www.euroasiapub.org](http://www.euroasiapub.org) | email: [editorijrim@gmail.com](mailto:editorijrim@gmail.com)