Problem Identification

Problem Identification

There are countless problems with family planning and health programs. Finding a problem is not hard, but identifying one for the purpose of research is not always easy. One of the most important tasks of research is to identify and define clearly the problem you wish to study. If you are uncertain about the research problem, if you are not certain in your own mind about what you want to study, you may be sure that others who read your proposal will also be uncertain. A well-defined problem leads naturally to the statement of research objectives, to the hypotheses, to a definition of key variables, and to a selection of a methodology for measuring the variables. A poorly defined research problem leads to confusion.

All research is set in motion by the existence of a problem. A problem is a perceived difficulty, a feeling of discomfort with the way things are, a discrepancy between what someone believes **should be** and **what is.** While problems are the initiating force behind research, not all problems require research. A potential research situation arises when three conditions exist:

1 A perceived **discrepancy** between what is and what should be.
2 A **question** about why the discrepancy exists.
3 At least **two possible and plausible answers** to the question.

The last point is important. If there is only one possible and plausible answer to the question about the discrepancy, then a research situation does not exist. Consider the example given below.

1 Example of a Nonresearch Problem

Problem Situation: A recent survey in District A found that 1,000 women were continuous users of contraceptive pills. But last month's service statistics indicate that none of these women are using contraceptive pills.

Discrepancy: All 1,000 women **should be using** contraceptive pills, but all 1,000 women **are not using** contraceptive pills.

Problem Question: What factor or factors are responsible for 1,000 women discontinuing their use of contraceptive pills?

Answer: A monsoon flood has prevented all new supplies of pills reaching District A, and all old supplies have been exhausted.

In the above example, a problem situation exists, but the reason for the problem is already known. Therefore, assuming that all the facts are correct, there is no reason to conduct research on the factors associated with pill discontinuation among the 1,000 women. On the other hand, there may very well be a need to conduct research on the question of why the supply and logistics system is incapable of providing contraceptives to women during the monsoon. Study the next example.

2 Example of a Research Problem

Problem Situation: District A is always flooded during the monsoon season. Recognizing this problem, the national family planning program established a new supply logistics system for the district. Each pill user is given a four-month supply before the monsoon begins. During the monsoon, small motorboats are available to transport new supplies to selected distribution centers accessible to village-level family planning workers. Despite these new measures, this year service statistics indicated that there are no pill supplies in District A.

Discrepancy: The new logistics system **should be able** to assure a continuous supply of pills, but this year **there are no supplies.**

Problem Question: Why has the new supply logistics system been incapable of delivering contraceptive pills to users?

Possible Answers:

1 An order for new pill supplies was not placed in time before the monsoon rains.
2 The riverboats used to transport the supplies are out of order.
3 Field-workers were not told about the new system and failed to give users a four-month supply of pills before the monsoon.

In this example, there are several possible and plausible reasons for the problem situation. One or more of these reasons might be correct. Therefore, this is a potential research situation.

In some situations, it is relatively easy to identify the problem, to define it, to hypothesize the reasons for it, and to conduct research to determine which reason is correct or more nearly correct. The reasons for the supply and logistics problem in the above example could probably be determined fairly easily and certainly would not require an expensive research study. Other problems, such as the one in the next example, are not so easy to identify or to study.

3 Example of a Research Problem

Problem Situation: A recent family planning survey revealed great differences between villages in the rate of contraceptive prevalence. Despite the fact that all villages receive the same level of health and family planning services, some villages have a prevalence rate as high as 80 percent, while others have a rate as low as 6 percent.

Discrepancy: All villages **should have** approximately the same rate of contraceptive prevalence, but in fact **there is great variation** between villages.

Problem Question: What factors are responsible for the areal variation in contraceptive prevalence rates?

Possible Answers:

1 Villages differ in their socioeconomic environments. Some villages are agricultural; some are fishing communities. Some villages are Hindu; others are predominantly Muslim or Buddhist. Some villages have access to markets in towns; others do not. Some villages have schools, health clinics, electricity, and a good water supply; others do not have these facilities. These socioeconomic differences affect the level of contraceptive practice.
2 Villages differ in institutional support for contraceptive acceptance. In some villages, local influentials strongly support the national family planning program. In other villages, they do not support it. In some villages, there are active Mothers' Clubs that support family planning. In other villages, there are

no Mothers' Clubs. These differences in institutional support for family planning affect the level of contraceptive practice.

3 Village-level health and family planning workers differ in their effectiveness. Some workers are highly motivated and very active in their assigned areas. Other workers are less motivated and less active. These differences in worker effectiveness affect the level of contraceptive practice.

While the problem situation presented above is fairly clear, the reasons for the problem are complex. Three reasons have been given, but it is likely many more could be stated. In situations such as this, the researcher must devote considerable time and attention to identifying the problem situation. The aim is to focus the research on the most important aspects of the problem.

An Introduction to Research Methodology

Meaning of Research

Research may be very broadly defined as systematic gathering of data and information and its analysis for advancement of knowledge in any subject. Research attempts to find answer intellectual and practical ques6tions through application of systematic methods. Webster's Collegiate Dictionary defines research as "studious inquiry or examination; esp: investigation or experimentation aimed at the discovery and interpretation of facts, revision of accepted theories or laws in the light of new facts, or practical application of such new or revised theories or laws". Some people consider research as a movement, a movement from the known to the unknown.

It is actually a voyage of discovery. We all possess the vital instinct of inquisitiveness for, when the unknown confronts us, we wonder and our inquisitiveness makes us probe and attain full and fuller understanding of the unknown. This inquisitiveness is the mother of all knowledge and the method, which man employs for obtaining the knowledge of whatever the unknown, can be termed as research.

Research is an academic activity and as such the term should be used in a technical sense. According to Clifford Woody research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organizing and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis. D. Steiner and M. Stephenson in the Encyclopedia of Social Sciences define research as "the manipulation of things, concepts or symbols for the purpose of generalizing to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art."

Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the pursuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research. The systematic approach concerning generalization and the formulation of a theory is also research. As such the term 'research' refers to the systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analyzing the facts and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalizations for some theoretical formulation.

Objectives of Research:

The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has not been discovered as yet. Though each research study has its own specific purpose, we may think of research objectives as falling into a number of following broad groupings:

1. To gain familiarity with a phenomenon or to achieve new insights into it (studies with this object in view are termed as exploratory or formulative research studies);
2. To portray accurately the characteristics of a particular individual, situation or a group(studies with this object in view are known as descriptive research studies);
3. To determine the frequency with which something occurs or with which it is associated with something else (studies with this object in view are known as diagnostic research studies);
4. To test a hypothesis of a causal relationship between variables (such studies are known as hypothesis-testing research studies).

Need of Research Methodology

It is necessary for a researcher to design a research methodology for the problem chosen. One should note that even if the research method considered for two problems are the same the research methodology may be different. It is important for the researcher to know not only the research methods necessary for the research undertaken but also the methodology. For example, a researcher not only needs to know how to calculate the mean, variance, and distribution function for a set of data, how to find a solution to a physical system described by a mathematical model, how to determine the roots of algebraic equations and how to apply a particular method but also need to know (i) which is a suitable method for the chosen problem?, (ii) what is the order of accuracy of the result of a method?, (iii) what is the efficiency of the method? And so on. Considerations of these aspects constitute a research methodology. More precisely, research methods help us get a solution to a problem. On the other hand, the research methodology is concerned with the explanation of the following:

1. Why is a particular research study undertaken?

2. How did one formulate a research problem?

3. What types of data were collected?

4. What particular method has been used?

5. Why was a particular technique of analysis of data used?

The study of research methods gives the training to apply them to a problem. The study of research methodology provides us with the necessary training in choosing research methods, materials, scientific tools, and training in techniques relevant to the problem chosen.

Utility of Research

It has been observed that research is of extensive use for a manager in planning, forecasting, coordinating, motivating, controlling, decision- making, etc. While managerial research helps in managerial analysis, academic research helps in academic objectives. In a nutshell, managerial research acts as an aid to managerial decision-making, while academic research is of academic utility. Utility of social research includes social prediction, social enlightenment social welfare, social growth, social cohesion, social control, improving and perfecting the tools of social research, etc. On the other hand, managerial research has substantial relevance from the functional point of view of a manager.

Utility of managerial research can be summed up as:

a) Research is an aid to decision-making.
b) Research facilitates the process of thinking, analysis, evaluation, and interpretation of the business environment; and of the various business situations.
c) Research provides a basis for innovation.
d) Research and development helps to develop new products and to modify the existing products.
e) Research identifies problem areas.
f) Research establishes the relationship not only between variables in each functional area, but also between the various functional areas.
g) Research is an aid to forecasting, which is an effective tool in the hands of managers.
h) Research helps all the managerial functions.
i) Research helps in the economic utilization of resources
j) Market and marketing analysis may be based on research.
k) Research is an aid to management information systems and
l) Research is helpful in the formulation of policy and strategy.

Research Methods

It seems appropriate at this juncture to explain the difference between research methods and research methodology. Research methods may be understood as all those methods/techniques that are used for conduction of research. Research methods or techniques, thus, refer to the methods the researchers use in performing research operations. In other words, all those methods

which are used by the researcher during the course of studying his research problem are termed as research methods. Since the object of research, particularly the applied research, it to arrive at a solution for a given problem, the available data and the unknown aspects of the problem have to be related to each other to make a solution possible. Keeping this in view, research methods can be put into the following three groups:

1. In the first group we include those methods which are concerned with the collection of data. These methods will be used where the data already available is not sufficient to arrive at the required solution;
2. The second group consists of those statistical techniques which are used for establishing relationships between the data and the unknowns;
3. The third group consists of those methods which are used to evaluate the accuracy of the results obtained.

Research methods falling in the above stated last two groups are generally taken as the analytical tools of research.

Research methodology It is a way to systematically solve the research problem. It may be understood as a science of studying how research is done scientifically. In it we study the various steps that are generally adopted by a researcher in studying his research problem along with the logic behind them. It is necessary for the researcher to know not only the research methods / techniques but also the methodology. Researchers not only need to know how to develop certain indices or tests , how to calculate the mean , the mode , the median or the standard deviation or chi - square , how to apply particular research techniques , but they also need to know which of these methods or techniques , are relevant and which are not , and what would they mean and indicate. Researchers also need to understand the assumptions underlying various techniques and they need to know the criteria by which they can decide that certain techniques and procedures will be applicable to certain problems and others will not. All this means that it is necessary for the researcher to design a methodology for his problem as the same may differ from problem to problem. For example , an architect , who designs a building , has to consciously evaluate the basis of his decisions , i.e. , he has to evaluate why and on what basis he selects particular size , number and location of doors , windows and ventilators , uses particular materials and not others and the like. Similarly, in research the scientist has to expose the research decisions to evaluation before they are implemented. He has to specify very clearly and precisely what decisions he selects and why he selects them so that they can be evaluated by others also.

From what has been stated above, we can say that research methodology has many dimensions and research methods do constitute a part of the research methodology. The scope of research methodology is wider than that of research methods. *Thus, when we talk of research*

methodology we not only talk of the research methods but also consider the logic behind the methods we use in the context of our research study and explain why we are using a particular method or technique and why we are not using others so that research results are capable of being evaluated either by the researcher himself or by others. Why a research study has been undertaken, how the research problem has been defined, in what way and why the hypothesis has been formulated what data have been collected and what particular method has been adopted, why particular technique of analyzing data has been used and a host of similar other questions are usually answered when we talk about research methodology concerning a research problem or study.

Deduction: Testing Theory

Deduction is the research approach used to test a theory. It involves the development of a theory that is subjected to a rigorous test. The stages involved in such research are:

a) deducing a hypothesis from the theory ,
b) expressing the hypothesis in operational terms ,
c) testing the operational hypothesis ,
d) examining the specific outcomes of the enquiry , and
e) Modifying the theory in the light of the findings, if necessary.

An attempt is thus made to verify the revised theory by going back to the first step and repeating the whole cycle

Induction: Building Theory

The other approach to conducting research is induction. In this approach, the researcher tries to understand the nature of the problem, gather the required quantitative and qualitative data, and analyze them to draw conclusions. Hence, developing an understanding about the problem and making proper analysis of its different dimensions are the strengths of inductive research. The result of this analysis would be the formulation of a theory. Thus, in an inductive approach, theory is built from the empirical evidences gathered through different sources. In deductive approach, data would follow a theory.

Characteristics of a Scientific Method:

The chief characteristic of a scientific method are:

1. Verifiability: The conclusion drawn through a scientific method is subjected to verification at any time. The preposition is that the phenomenon under investigation must be capable of being observed and measured. In case direct observation could not be done, other methods such as interview can be utilized for verification. For instance, 2 man's order of preference for various jobs, although incapable of being observed can still be verified by means of an interview.

2. Generality: Laws derived through scientific method are universal in their applications. They are not limited to individual objects or individual groups of objects. The individual groups or objects considered as specimen or instances, and there are relationships discovered through these individual groups should be applicable to whole group called Universe. However, complete universality is rarely achieved in social sciences because of heterogeneous nature of social phenomenon. Universality in case of the laws of social sciences thus means limited universality and they are true only under given conditions.

3. Predictability: Another characteristic of a Scientific method is that its results can be predicted with sufficient accuracy. For example, we can say with certainty that if water is heated to 100°C, it will vaporize, and if it is cooled to O'C, it will turn to ice. Also we can say that if an unbiased coin is tossed 500 times, head will turn about 250 times. Predictability is fixed on two factors i.e. fixing of relationship between the cause and the effect and the stability of causative factors.

Predictability depends on one hand upon the nature of phenomenon and on the other hand upon our knowledge of various causative factors. If the number of causative factors is large and they pull their weights in different direction, accurate prediction becomes rather difficult. This is the kind of difficulty that is found in most of the social phenomenon and it is because of this reason people do not acknowledge its claim to be called a science at all.

4. Objectivity: The results obtained through a scientific method should be free from investigator's own views, wishes or prevalent notions i. e., they must be subjected to objective observations. The main criterion of objectivity is that all persons should arrive to the same conclusion about the phenomenon. For example, when we say Coal is black, it is objective statement because coal will appear black to all people. But when we say Coal is useful mineral, the statement may not be objective, for every one may not agree to the statement.

Objectivity is essential for verification. It permits repetition of observations under practically identical conditions. This facilitates the verification of an observation by many observers. Apparently objectivity seems very easy to achieve but in actual practice, it is not often so. It is very difficult some- times to have a detached view of a phenomenon in which the investigator himself involved. This is why objectivity is more difficult to achieve in social Sciences.

5. System: In every scientific study, there is an accepted mode of investigation. The result arrived true, at by means of a haphazard method, even true, cannot be called scientific because its accuracy is purely accidental.

Types of Research

Research can be divided into two broad types relative to its purposes; applied and fundamental

1. Applied Research

Applied research is conducted in response to a specific problem, which requires a solution. The major purpose of applied research is to answer practical and useful questions about policies programs, projects, procedures, or organizations. Business executives, therefore, take interest in applied research. They often hire outside researchers and consultants to study a problem of concern to them in order to find solutions that can be implemented to rectify the problem situation.

As applied research is concerned with knowledge that has immediate applications, it is also called decisional research. The attempt to get a cure for Bird Flu is a case in point.

EXAMPLE: *The Dairy Development Corporation (DDC) has to improve its productivity in order to remain competitive in the market. There are two alternative strategies to improve its productivity. One is to pay attention to all of its existing brands and make continuous improvement; the other is to focus on new brand development. Each of these alternatives has some advantages and disadvantages. The Corporation will now have to research into each of these strategies and see which one would best be suitable to it, taking into account its capabilities, know-how, resources and so on.*

The above example indicates the need for an applied research to work out a strategy in view of the strengths and weaknesses of the DDC. Applied research is thus more concerned with knowledge that has immediate application and would be useful in making decisions and formulating policies.

The defining quality of applied research is that the researcher attempts to conduct a study whose results can be applied directly to a specific situation. To accomplish this task, the researcher must choose a research strategy that maximizes the applicability of findings. Applied research often results in recommendations on decisions or actions.

2. ***Fundamental Research***

Fundamental research is undertaken to improve our understanding of certain problems that commonly occur in organizational setting, and how to solve them. It is undertaken for the sole purpose of adding to our knowledge that is fundamental and generalizable. It is conducted without any practical end-use in mind. It is also known as pure or basic research.

The purpose of fundamental research is not to apply the findings to solve an immediate problem at hand, but rather to generate more knowledge and understanding of the phenomena and problems that occur in several organizations, and to build theories based on the research results. Such theories subsequently become the foundation for further study of the phenomena.

Fundamental research is also concerned with the development, examination, verification and refinement of research methods, procedures, techniques and tools that form the body of research methodology. Thus, fundamental research simply aims to advance knowledge and to identify and explain relationships between variables.

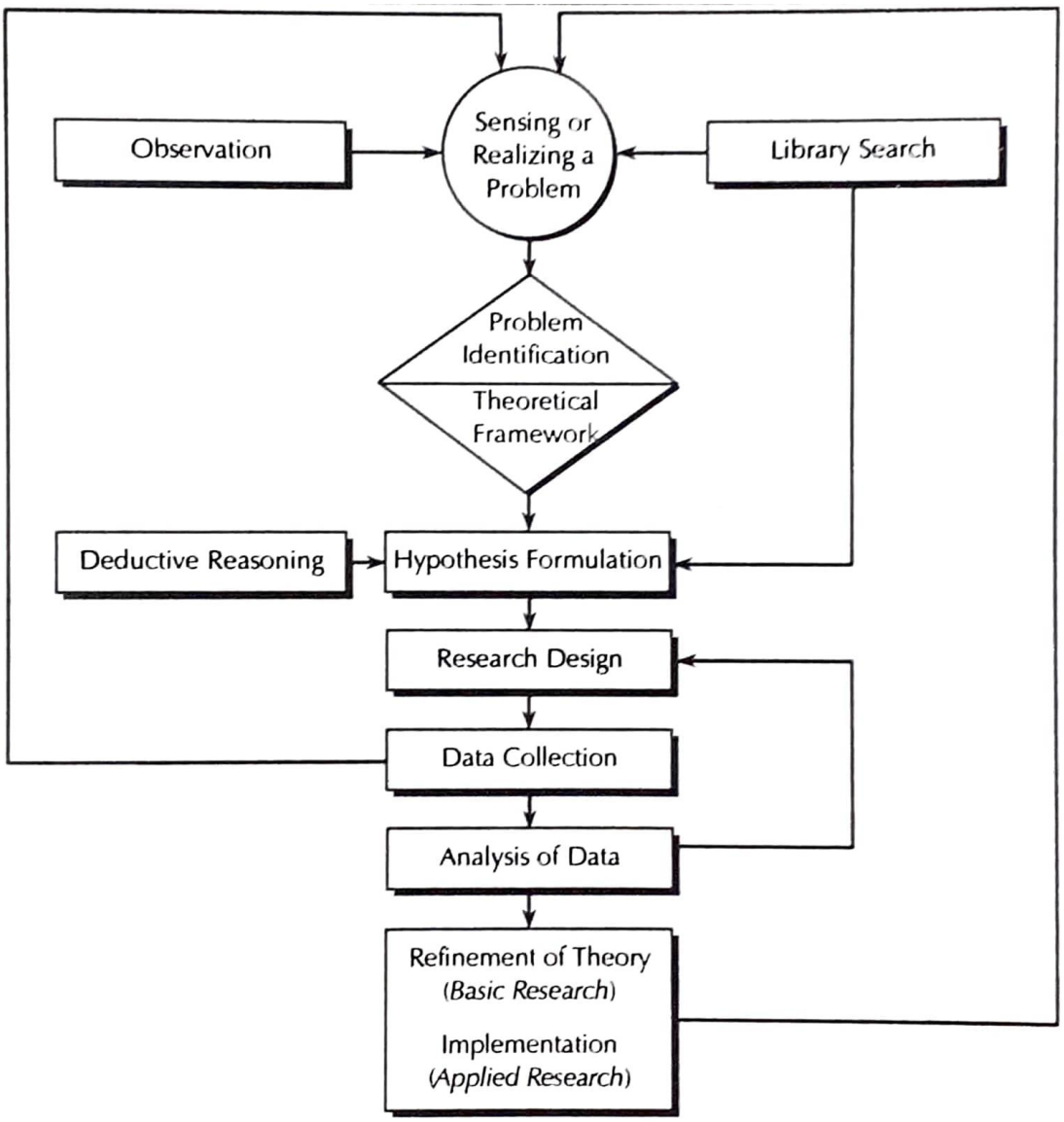
EXAMPLE: The HRD Managers' Conference recently held in Kathmandu focused on the issue of employee socialization, training and commitment. The participants in that HRD Conference commonly believed that socialization and training should have a great impact on the productivity and organizational commitment of employees However, through the years; it has been observed that the productivity of workers over forty years of age does not improve from

socialization and training. On the other hand, organizational commitment of employees over 40 years is relatively higher. Why is this so? The participating HRD managers were looking for answers to these phenomena. What factors might be responsible for these phenomena? What type of socialization and training should be given to employees to improve their productivity and organizational commitment?

To answer the questions raised by the HRD managers, one has to undertake a basic research. Our existing knowledge is not enough to answer these questions. The purpose of the research is thus to simply increase the amount of knowledge on the issue of employee commitment and training, not to actually come up with a practical solution to a problem. The researcher therefore, has to design an investigation and conduct a study observing socialization and training programs in different work settings and recording the effects of such programs on the productivity and commitment of younger and older employees.

Several of such experiments conducted in different work situations would give the research some idea about the relationship between socialization, training, commitment and productivity of employees belonging to different age groups. The main purpose of conducting fundamental research is thus to advance the level of scientific knowledge.

Scientific Research Process



Scientific research is systematic and follows the steps of the scientific method. From the inception of a research idea to the final report of results, the research process has several crucial steps. However, these steps do not provide a rigid pattern into which you must force your thinking. Thinking simply cannot be scheduled. An investigator does not tackle one step at a time, complete that process and then move on to the next step. Some steps can go simultaneously. Others need proper sequencing and logical arrangement.

Generally, research is understood to follow a certain structural process. There are eight steps in scientific method. These eight steps cover the total spectrum of a research endeavor, starting from problem formulation through to refinement of theory or practice. However, these steps may vary depending on the subject matter and the researcher, and also are interdependent with considerable back and forth interaction. These steps are diagrammatically shown in Figure. Brief description of each of these steps is given in the following paragraphs:

Sensing or Realizing Problem

The first step in any scientific inquiry is to identify an issue you want to study. There are many sources of research idea (observing the situation or sensing the problem). New problems keep on emerging in the environment. You somehow sense these developments occurring in the environment. At this stage, you may not know exactly what is happening, but you can definitely sense that things are not going on as smoothly as they should be.

Problem Identification

Once you increase your level of awareness of what is happening in the environment, you would then focus on the problem and the associated factors through further search of information. In this step, you try to identify what exactly are the problems or issues in the situation. There is a saying in research that "a problem well defined is a problem half solved". If the research problem is unclear and poorly defined, the result could be a lot of time and resources wasted on gathering OP potentially useless information and data.

Theoretical Framework

In the third step of scientific research, you make an attempt to integrate the information logically so that the reason for the problem can be conceptualized. The critical variables are examined and the association among them is identified. Putting all the variables and their association together, a theoretical framework is developed.

Hypothesis Formulation

In the fourth step of scientific research, hypotheses are formulated. Hypotheses are logically conjectured relationship between two or more variables expressed in the form of testable statements. Hypotheses for the study arc drawn from the theoretical framework as developed in Step 111. Research hypotheses are even more focused. They provide the specific answers to questions that the research will examine often in an empirical way. Hypotheses are particularly useful in quantitative research, where there is statistical analysis.

Research Design

The fifth step is devising the plan for the research. Once you have narrowed your research hypothesis, you must next decide on a design or plan of attack for your research. The research design is thus a strategy for conducting research. It describes the general framework for analyzing and evaluating data after identifying:

a) What you want to know, and
b) What has to be dealt with in order to obtain required information.

The decision of which design to use can totally depend on the nature of the research project. Once you have decided on which earth design to undertake for your project, you can then develop the collection methods and data

Collection of Data

Data collection, the sixth step in scientific research, is also known as fieldwork. At this stage, you have to administer the research instruments (questionnaire, interview schedules, observation schedules, etc.) to gather data as expected in Step V. However, the procedures used to obtain the data vary depending on the research design chosen and the source of the data. It is important to note that this step is the key part of the scientific research process and is crucial to the success of the research project. All of the research and planning effort so far is of little use if the data is incorrectly collected.

Data Analysis

After you have collected data, you must summarize and analyze them. Data analysis is in fact the statistical analysis of data that have been edited, coded and tabulated. It especially important in cases, where you have amassed large amounts of information from many respondents. You can analyze data in several ways, and some types of data are better analyzed with one method than another. In most cases, you will probably calculate some descriptive statistics that provide a "nutshell" description of your data and inferential statistics that assess the reliability of our data. With the use of these different statistical techniques, the hypotheses are tested.

Refinement of Theory or Practice

The final step involves interpretation and generalization of the findings into the larger body of knowledge about the phenomenon. In the case of applied research, specific implementation strategy is proposed to solve the problem identified by the study. Through research existing theories or practices are refined and modified.

By carefully following the major steps outlined above, you can reduce the possibility of making major errors and increases the possibility that meaningful research results will be obtained. Scientific research helps you to state your findings more accurately and with confidence.

Consider the following case as an example, which shows the nature and process of scientific research:

EXAMPLE *A dealer of a car producing company was concerned with the complaints received from the car users that the cars it produces have some problems with rating sound at the dash board and the rear passenger seat after a few thousand kilometers of driving.*

- *He obtained information from the company workers to identify the various factors influencing the problem.*
- *He then formulated the problem and generated guesses (hypotheses).*
- *He constructed a checklist and obtained requisite information from a representative sample of cars.*
- *He analyzed the data thus collected, interpreted the results in the light of his hypotheses and reached conclusions.*

In this example, the dealer went through a sequence of steps which were in order and thus systematic. Secondly, the dealer did not just jump at the conclusions. He used a step-wise scientific method of inquiry in reaching at conclusions. This research study, thus, met the important characteristics of research: first, it was a systematic process, and secondly it followed a step-wise scientific method of enquiry to reach at the conclusion.

The most characteristic feature of the scientific research process is its cyclical nature. Research usually starts with a problem and ends in a tentative empirical generalization. The generalization at the end of one cycle is the beginning of the next cycle. The cyclical process continues indefinitely, reflecting the process of a scientific discipline and the ongoing accumulation of scientific knowledge. The research process is also self-correcting. Tentative generalizations to research problems are tested logically and empirically. If these generalizations are rejected, new ones are formulated and tested.

Research Language

- Theory
- Conceptualization
- Operationalization
- Variables
- Hypotheses
- Assumptions
- Population
- Sample
- Validity
- Reliability
- Data
- The research process
- Summary
- Methodological queries
- References

Concepts And Constructs

The terms "concept" and "construct" have similar meanings. Yet there is an important distinction. A concept expresses an abstraction formed by generalization from particulars. "Weight" is a concept: it expresses numerous observations of things that are more or less "heavy" or "light." "Mass," "energy," and "force" are concepts used by physical scientists. They are, of course, much more abstract than concepts such as "weight," "height," and "length."

A concept of more interest is "achievement." It is an abstraction formed from the observation of certain behaviors of children. These behaviors are associated with the mastery or "learning" of school tasks—reading words, doing arithmetic problems, drawing pictures, and so on. The various observed behaviors are put together and expressed in a word—"achievement." "Intelligence", "aggressiveness", "conformity," and "honesty" are all concepts used to express varieties of human behavior.

A construct is a concept. It has the added meaning, however, of having been deliberately and consciously invented or adopted for a special scientific purpose. "Intelligence" is a concept, an abstraction from the observation of presumably intelligent and non-intelligent behaviors. But as a scientific construct, "intelligence" means both more and less than it may mean as a concept. It means that scientists consciously and systematically use it in two ways. One, it enters into theoretical schemes and is related in various ways to other constructs. We may

say, for example, that school achievement is in part a function of intelligence and motivation. Two, "intelligence" is so defined and specified that it can be observed and measured. We can make observations of the intelligence of children by administering X intelligence test to them, or we can ask teachers to tell us the relative degrees of intelligence of their pupils. Constructs can be conceptually defined in that they have meaning in theoretical terms. They can be abstract and do not necessarily need to be directly observable. Examples of constructs include intelligence or life satisfaction.

Unit 2: Research Problem Identification and Formulation

Identification of Research Problem

Formulation of problem is the first and foremost step in a research work. The research problem can be formulated and selected rationally and then the whole research work can be conducted only if the identification of the problem is done precisely. To identify the problem a researcher should have some basic knowledge, which is then developed through discussion with experts or through the literature or by continuation of activities in the related field.

The research problem undertaken for the study must be carefully selected. This task is a difficult one, although it may not appear to be so. A problem must coin from the researcher's mind like a plant springing from its own seed. Help may be taken from a supervisor in this connection. A research supervisor can at the most only help a researcher to choose a subject. Identifying the exact nature and dimensions of a problem is of major importance in research work. It is very essential that an investigator should learn how to recognize and define a problem. He should proceed step by step in locating the research problem.

The following steps are to be followed in identifying a research problem:

1. Determining the field of research in which a researcher is keen to do the research work.
2. Develop the mastery on the area or in field of specialization.
3. Review the recent research conducted in the selected area.
4. On the basis of review, select the priority field of the study.
5. Draw an analogy and insight in identifying a problem or employ the personal experience of the field in locating the problem. In this process researcher can take help of supervisor or expertise of the field.
6. Pin-point specific aspect of the problem which is to be investigated.

Ways of Understanding Research Problem

The selection of a suitable problem is not an easy task. It is a serious responsibility to commit oneself to a problem that will inevitably require much time and energy and which is so academically significant. Specifically, the concept for separating out the research problem from the diversified field can be made by the

i. Discussion among the colleagues
ii. Discussion with the research guide
iii. Discussion with some experts and
iv. Intensive reading all the available literature.

The following are the general ways for understanding problem to which one may proceed for a suitable research:

1. Personal experience of the investigator in any field is the main means for understanding a suitable problem. Many of the problems confronted in our daily life lend ourselves to investigation and they are perhaps more appropriate for the beginning researcher than are problems more remote from our experiences.
2. The other ways of understanding of problem, most frequently used by the investigator as suggested by the supervisors, is the extensive study of available literature-research abstracts, journals, hand-books of research international abstracts etc. He can draw an analogy for selecting a research problem or can think parallel problem in the field studied.
3. In the choice of a suitable problem, the researcher has to decide his field of investigation. He should study the field intensively in the specific area; this may enable him to identify a problem from the specific field.
4. The new innovations, technological changes and curricular developments are constantly bringing new problems and new-opportunities for social research.
5. The most practical ways of understanding problem is to consult supervisor, experts of the field and most experienced person of the field. They may suggest most significant problem of the area. He can discuss certain issues of the area to emerge a problem.
6. It is a general practice that researchers suggest some problems in their research reports. The researcher can understand a suitable problem for his own study.

Steps in Research Problem Formulation

If one does not know to single out the research problem, he cannot enter into research work. Therefore, he should know how the problems are understood and formulated. In a scientific inquiry, the formulation of general topic into a specific problem is the formulation of research problem. The researcher must try to understood the whole problem thoroughly and then rephrasing the same into meaningful terms from an analytical point of view. Therefore, the following points are suggested as the steps for the formulation of research problem.

• Statement of problem in general way

• Understanding the nature of the problem

• Surveying the available literature

• Developing the ideas through discussion

• Rephrasing the research problem.

Criteria of a Good Problem

Fred. N. Kerlinger has defined a good problem as an inquisitive sentence statement that asks what relation exists between two or more variables. The research questions, research objectives and the hypotheses of the research lie on the problem statement of the research work. So for selecting research questions, setting research objectives and hypothesis of the research the selected problem should be an ideal one. An ideal research problem must have the following three main criteria:

1. *The problem should express a relation between two variables.* The topic of the research work reflects the research problem and it would be inquisitive and interesting if stated in terms of relation between two variables.

2. *The problem should be stated rigidly and unambiguously.* If a research problem has ambiguity and haziness in its interpretation consequently it affects research design and the whole research process and hence the result of the study. If a formulation of research topic has some dilemma then these should be identified and eliminated carefully to make a sensible research question

3. *The problem statement should be such as to imply possibilities of empirical testing.* In case, when problem statement expresses the relationship between two variables and is stated rigidly but if the testing the relationship in the circumstances of adopted research design and within the available facility is impossible then the selected problem cannot be good problem. Thus, a good research problem should always be linked with the research design and it must consider the available facility and capability of the study.

Problem Statement

Developing a problem statement includes some combination of the interrelated tasks like generation of an issue, exploration of an issue, and from that exploration determining worthwhile research questions. The problem statement has thus the following three characteristics:

- It should raise a question about a relationship between variables.
- The relationship between the variables should be stated and explained clearly.
- The problem statement should suggest a method of researching the question.

A statement of problem could be stated either in declarative or in interrogative form.

EXAMPLES

Declarative: Factors contributing to the excessive absenteeism among Nepalese workers.

Interrogative: Why is absenteeism so high among workers in Nepalese organizations?

The following are some examples of well-defined research problems. These are stated in interrogative form:

EXAMPLES

- To what extent do age, education, length of service, level of earning, and place of residence of employees predict occupational aspirations?
- Do long work hours, lack of development opportunities, and discrimination account for the lack of inward mobility of women in civil service?
- Can cultural differences account for the differences in the nature of hierarchical relationships between supervisors and subordinates?

A problem statement that appeared in a research report is given here as an example:

EXAMPLES A review of research studies on leadership and age (Kabacoff & Stoffey, 2001) reveals the controversy and recent interest in examining the relationship between the variables. These studies explain some reasons for the importance of the association between age and leadership. In the first place, they explain that older workers remain in employment for longer years and they work side by side with younger members in various work and leadership roles. It is no secret that today's workers, as a result of several reasons; live much longer than their counterparts in the last generation.

As older and younger employees abound in organizations there is the need to manage both of them effectively in order to realize organizational goals more fully. Both age groups have something to offer: technology has driven the rapid promotion of younger workers while experience has made the older worker very relevant. In addition, today's flatter organizations give greater interaction between younger and older workers and the practice of leadership is no longer an exclusive domain of the older people, as it used to be. In several organizations, therefore, top management team will comprise of multigenerational members. It is argued that the ability to understand, learn, and effectively leverage multigenerational diversity will be necessary for organizations now and in the future to build and maintain high performance systems (Kabacoff & Stoffey, 2001, p.2).

Another reason that underscores the importance of the study of the relationship between leadership and age is the finding by kakabedse. et al. (1998) that age, along with other time-related dimensions, has a powerful effect in shaping the attitudes and hence the behaviors of senior leaders within organizations. In a study undertaken in Australia, three leader profiles emerged - the radicals, the Bureaucrats and the team players. The radicals were the youngest (between 26- 35years), the team players were the oldest (56 years and over), while the bureaucrats were in between (45-55 years). Older workers were mature, saw challenges and had longer-term perspectives in managing people and systems. On the other hand, younger employees were competitive, result oriented, energetic and adopted an open style management.

There is a need for further study of the leadership styles and behavior of younger versus older organizational leaders. The results of such a study will hopefully identify similarities and differences of younger versus older workers for use by managers in directing the affairs of their organizations towards attaining their goals efficiently and effectively. For example, are there some tasks, such as computer-related activities, which may be better performed by younger, rather than old workers? On the other hand, are there some personnel projects that are better managed by older, rather than younger workers? Answers to such questions may be helpful in formulating and revising some organizational policies.

Source: Oshagbemi, T. (2002). Age influence on the leadership styles, *Employee Relations,* 26 (1), 37-52.

Research Questions

Research is intended to help us learn something new. The research process encloses the research questions, the most important element of any research, for the effective execution of research activity. We often define our research goals in terms of questions and hence research questions describe the ideas contained in the research objectives. Research questions are the interrogative form of research objectives. Research questions are such questions that can help the researcher learn something new- fruitfully Formulation of research questions is the real starting point in preparation of a research process. The data required to be collected for the study are determined

by the help of the research questions. The research design is necessarily based on the research questions; the research method to be adopted for the study of specific problem is also set on these questions. The data analysis tools and methods, result interpretation procedures and writing phases of the reports are also determined by the research questions. The questions have to be related to three aspects: What, Why and How? What question seeks descriptions, why question seek explanation and understanding and how question seek interventions to bring about change. If the researcher does not have clear formulated research questions at the starting point of the research in his practical field, then He (she) may face different challenges to precede the direction of research. In such case he (she) may have some loosely connected ideas about what should be researched.

Types of Research Questions

Generally there are three types of research questions, they are: What questions (concerned with description), Why questions (concerned with explaining causes) and how questions (concerned with bringing about change). 'What questions' pertain to describing the characteristic of trend and pattern in the given situation. For example: What are the types of community involved in transformation of skills?, What are the socio economic characteristics of community?, what are the needs of the community ? etc. Why question relate to the cause or reasons for the characteristics of the particular phenomenon and the behavior of the individual involved. They also explain the relationship between events and activities. For example: Why do drug abusers commit thefts? , Why does stressful living result in heart attacks?, Why do some people use a product while other do not? How questions are concerned with bringing about change and the outcomes of change. For example: How has caste system changed in Nepal in last century?, How does technology create unemployment?, How do MC health service affect infant mortality? Besides these three types of question different authors have proposed different types of research questions. Lin (1993) has proposed four more types of research questions they are Who, Where, How many and How much. Similaly Hedrick et al (1993) had identified four types of research questions as: descriptive, normative, correlative and impact. Marshall and Rossrnan (1995) have classified research questions as theoretical, site-specific and population-specific.

Identification of Research Question

The main purpose of formulating research question is to define the scope of the research. It is used to determine what is to be studied and the extent to which it will be studied. Neuman _ (1997) offered some techniques of developing research questions, which are:

- Record all questions that occur in mind after reading literature or after discussions with other or after thinking on various aspects of study.
- Review all these questions whether each question is necessary and delete those which are beyond the scope of the study. This will also remove overlapping between questions.
- Classify questions on the basis of their nature, i.e. What, Why and How questions

- Examine the scope of the questions for availability of time and money; chose these which can be answered within manageable time and resources.
- Separate major or key questions (which for the core of the research) from subsidiary questions.

Hypothesis

Definition

A hypothesis proposes relationship between two or more variables. In common usage, a hypothesis refers to a provisional idea whose merit requires evaluation. For example: political participation increases with education. This simple assertion can be seen as a hypothesis. It has a subject (the variable, political participation), a connective verb (a relationship, increases), and an object (the variable, education). This hypothesis takes two basic ideas "political participation" and "education" and suggests that they are connected to the extent that as one increases then the other increases as well. This can be stated in more mathematical terms as one variable being directly proportionate to the other.

Fred N. Kerlinger and H. B. Lee (2000): "A hypothesis is a conjectured statement that implies or states a relationship between two or more variables".

John W. Creswell (2014): "A hypothesis is a formal statement that presents the expected relationship between independent and dependent variables". A hypothesis is thus a statement about the relationship between two or more variables which needs to be investigated for its truth. It is basically a working assumption. If the relationship between two variables is found as the hypothesis predicts, then the hypothesis is supported and a new theory has been suggested. A good hypothesis states as clearly as possible the expected relationship (or difference) between two variables and defines these variables in operational and measurable terms**.**

Functions of Hypothesis

Specifically, a hypothesis serves the following functions (Kumar, 2011):

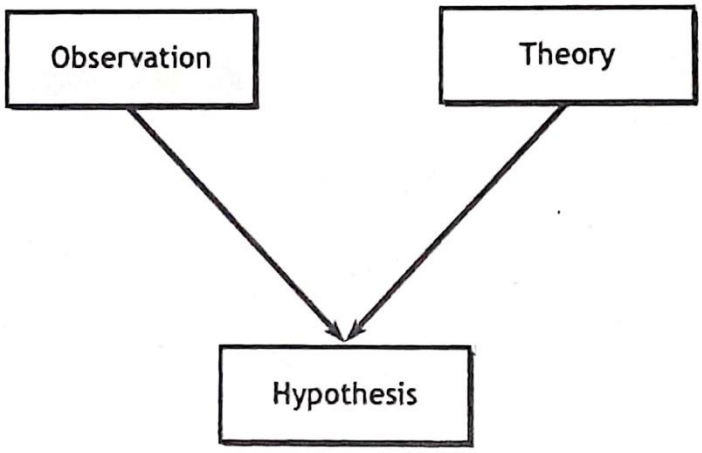
- The formulation of a hypothesis provides a study the focus. It tells you what specific aspects of a research problem to investigate.
- A hypothesis tells you what data to collect and what not to collect, thereby providing focus to the study.
- As it provides a focus, the construction of a hypothesis enhances objectivity in a study.
- A hypothesis may enable you to add to the formulation of theory. It enables you to specifically conclude what is true or what is false**.**

Hypothesis Formulation

Hypothesis can be derived in a variety of ways i.e. general culture, past research/scientific theory, personal experience, discussion and conversations and intuition. A researcher observes a social situation and come to a conclusion about some of the variables which are operating within it. You could then develop some hypotheses which connect two or more of these variables. Generally there are two grounds on which a hypothesis may be justified: logical and empirical.

Logical justification is developed from arguments based on concepts and theories relating directly to the research problem. Empirical justification is based on reference to other research found in the literature. Hence, in order to formulate a useful hypothesis, you need to have good knowledge of the background to the subject and the nature of the problem or issue which is being addressed. A hypothesis statement is derived directly from the statement of the problem. Hypothesis can be stated rather easily once the research problem is known. The hypothesis is thus more operational than the problem statement.

A diagrammatic presentation of the process of hypothesis formulation is given in figure.



Hypothesis generation and testing require an understanding of the deductive and inductive reasoning. A simple definition of deductive reasoning is "taking a known idea or theory and applying it to a situation with the intention of testing whether it is true". Deduction is thus the process of arriving at conclusion by interpreting the meaning of the results of the data analysis. In this form of reasoning, one goes from general knowledge to specific knowledge. For example, consider the following two arguments**:**

Argument 1

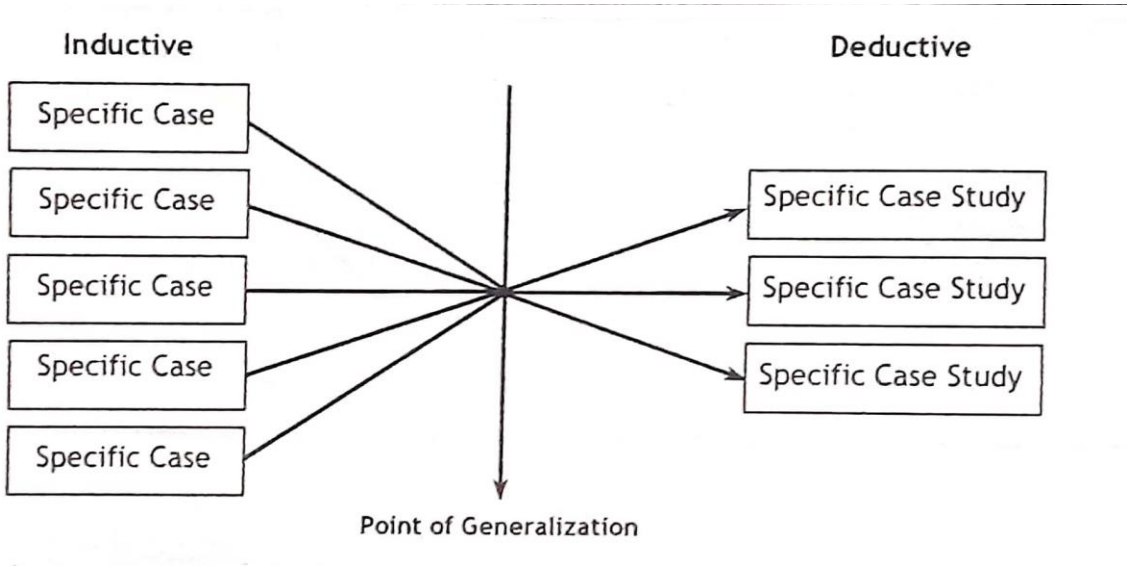
- All books have pages (theory or generalization).
- This is a book (empirical observation - fact).
- Therefore, this book has pages (logical conclusion).

Argument 2

- Lung cancer is caused by smoking (theory or generalization).
- John has lung-cancer (empirical observation-fact).
- John was a cigarette smoker (logical conclusion).

Deduction starts from a generalization and goes to a specific case. Inductive reasoning is the opposite. A simple definition of inductive reasoning is "using observation to formulate an idea or

theory". Induction is thus the process of examining many cases and then generalizing from them. In this form of reasoning, one goes from a specific knowledge to the general knowledge. Looking at the cigarette-cancer example, the researcher would investigate whether the people who have long-cancer had previously smoked before they received cancer. Both induction and deduction are used by the researchers to organize facts, describe results, develop new relationship, and suggest new research.



Deductive logic, as stated above, is useful for going from a general problem to a specific hypothesis. This statement can be understood by an example:

■ **Problem:** Nepal's population is increasing so rapidly that if it continues at its present rate, in 30 years, it will not be possible to feed all of its citizens.

■ **Theory:** Population growth can be controlled through family planning clinics.

■ **Hypothesis:** Family planning can reduce the growth of population in Nepal.

The above example is stated in very general terms and later will be made specific so that it can be tested, but it is a good example of the deductive logic used in a thesis. The problem statement has been created from the facts of Nepal's present population and from the prediction of its growth in the future. The prediction is made from examining past facts of population growth in Nepal. The theory is a very general theory, which has been true in other parts of the world; it has almost become a principle for population control. The hypothesis is the result of deductive logic from the first two statements.

It is now possible to examine the stating of problems and hypotheses. The problem asks about the relation between several facts or observations. Accordingly, the hypothesis suggests that the relationship exists. It is important to realize that the hypothesis has to be stated in a very specific terms so that the means of investigating the hypothesis are included in the statement. A problem

is formulated in the form of a question; it serves as the basis or the origin from which a hypothesis is derived. A hypothesis is a suggested solution to a problem. A problem (question) cannot be directly tested, whereas a hypothesis can be tested and verified. Hence, a problem cannot be scientifically solved unless it is reduced to hypothesis form.

EXAMPLE

Research Problem: What is the relationship between population growth in Kathmandu before the introduction of family planning and after the introduction of family planning?

Research Hypothesis: There is a significant difference in the population growth in Kathmandu between when family planning was first introduced and five years later.

Hypothesis statements should be clear if the definition of a variable is understood as some characteristic, which changes. The above hypothesis is simply stating that two groups exist in relation to some characteristic. If there is a significant difference between the two groups then the hypothesis is supported. The theory that family planning can reduce the growth of population in Nepal is supported and becomes tentative. Hence, there appears to be a solution to the population problem in Nepal.

One-sided Vs Two-sided Hypothesis

During the planning of our research, we need to specify whether we plan to use a one-sided or two-sided hypothesis. A one-sided hypothesis states a specific direction (e.g. increase or decrease). If a change in the unexpected direction is equivalent in practice to no change, then we should use a one-sided hypothesis. A two-sided hypothesis states that there is a difference between the dependent and independent variable, but does not specify the direction. If we expect that a change in either direction is possible and that changes in either direction are interested, then we should use a two-sided hypothesis.

Formats of Stating Hypothesis

The different formats of hypothesis construction based on association and correlation between variables are as follows:

Correlation There is a significant relationship between Variable A and Variable B for Group 1,

Difference between Means

There is a significant difference between mean levels of Variable A for Group 1 and Group 2.

Difference between Frequencies

There is a significant relationship between Group 1 and Group 2 for Variable A. There is a significant difference between Group 1, 2 and 3 for the following variables:

Variable A Variable B etc.

Types of Hypothesis

1. Descriptive and Relational Hypothesis

Research hypothesis can be classified as: descriptive and relational. Descriptive hypotheses are in the form of propositions that only state the existence, size, form, or distribution of some variable (Cooper & Schindler, 2011).

EXAMPLES

- Tribhuvan University (case) is experiencing budget difficulties (variable).
- The Hetauda-Narayangadh sector of the East-West Highway (case) has a higher-than-average accident rates (variable).
- The average stockholders of Nepal Development Bank (case) favor returns in the form of bonus dividends (variable).

These descriptive statements contain only one variable. Hence, the relationship between variable cannot be studied and explored. These statements do not fulfill the criteria of research hypotheses. It is, therefore, advisable to use research questions rather than descriptive hypotheses. The research questions for the above three statements could be stated as follows:

EXAMPLES

- What is the extent of budget difficulties in Tribhuvan University?
- Why is the accident rate higher in Hetauda-Narayangadh sector of the East-West Highway?
- Why do the stockholders of commercial banks favor returns in the form of bond dividends?

A relational hypothesis, on the other hand, describes the relationship between two or more variables with respect to some case. Relational hypotheses are of two types: correlational hypothesis and explanatory (causal) hypothesis. When a statement describes the relationship between two variables, it is called a correlational hypothesis. A Correlational hypothesis states that the variables occur together in some specified manner without stating that one causes the other.

EXAMPLES

- Families with higher incomes spend more for recreation.
- With education people's political participation will increase.

2. Explanatory Hypotheses

In an explanatory hypothesis, the implications of one variable on the other are stated. How one variable would cause or lead to a change in the other variable? Such causal relations can be unidirectional, in which variable A influences variable B, but not vice versa. They can also be bidirectional, in which each variable influences the other. The following are the examples of unidirectional and bidirectional relations:

EXAMPLES

- The increase in age would lead to decrease in organizational commitment.
- The productivity of skilled workers will increase if the workers are given added pay for production in access of the standard.

3. Directional and Non-directional Hypotheses

The directional hypothesis indicates the particular direction of the expected relationship between two variables. These relationships could be stated in positive or negative form. In stating the relationship between the two variables, the terms such as "positive", "negative", "more than' "less than" and the like are used. The directional hypothesis requires a one-tailed test. The following are the examples of directional hypotheses.

EXAMPLES

- Younger workers are less motivated than older workers.
- The greater the workload, the lower the job satisfaction of the workers..

The non-directional hypotheses are formulated when there are no clues available about the positive or negative relationship between two variables. Hence, these hypotheses do not indicate any direction of the relationship or difference and require a two-tailed test. Non-directional hypotheses are formulated in cases where previous studies do not exist or indicate conflicting findings (Sekaran & Bougie, 2013). The following are some examples of non-directional hypothesis:

EXAMPLES

- There is a difference between work attitudes of industrial and agricultural workers.

- There is no relationship between educated and uneducated employees in their occupational commitments.

4. Null and Alternative Hypotheses

There are the two methods of stating the hypothesis: null and alternative: A null hypothesis is a statistical hypothesis that is tested for possible rejection under the assumption that it is true. The hypothesis contrary to null hypothesis is known as alternative hypothesis. In other words, a null hypothesis is a hypothesis set up to be nullified or refuted in order to support an alternative hypothesis.

The null hypothesis is called null because it usually reflects the "no-difference" or "no-effect' situation. This hypothesis is thus the one actually tested statistically. It is an arbitrary convention hypothesizing that any relation or difference in the findings is due to chance or sampling error and puts this supposition to a probability test. Theoretically, it is a hypothesis set up for possible rejection.

The following example would clarify the concepts of null and alternative hypothesis. Suppose you are interested in a study to determine whether production would increase if the skilled workers are given a bonus or incentive pay for production in excess of a standard. For this investigation, you can formulate a research hypothesis in the following way:

EXAMPLE

The productivity of skilled workers will increase if they are given added pay for production in access of the standard.

This is a positive statement whose validity you would attempt to test through your research However, many researchers would object to the use of a positive hypothesis like this. A positive hypothesis like this may indicate a built-in prejudice on the part of the researcher toward a result favoring the hypothesis. To them, a null hypothesis is more desirable. The null hypothesis takes the form of a statement indicating no prejudice toward an answer. How can then this hypothesis be stated in a null form? The following is an example:

EXAMPLE No significant difference will exit between productivity of skilled workers on an incentives plan and productivity of skilled workers on a regular wage plan.

This null hypothesis thus indicates a definitive, exact relationship between two variables. That is, it states that the population correlation between two variables is equal to zero, or that the difference in the means of two groups in the population is equal to zero. In statistics, the only way of supporting your hypothesis is to refute your null hypothesis. Rather than trying to prove your idea (the alternative hypothesis) right, you must show that the null hypothesis is likely to be wrong. You have to refute or nullify the null hypothesis. You have to assume that your alternative hypothesis is wrong until you find evidence to the contrary. The following is another example of null hypothesis:

H_0: There is no difference between male and female statistically in their productivity.

Statistically expressed: H_o: $\mu_l = \mu_2$

Where, H_0= the null hypothesis

μ_1 = the productivity of male workers

μ_2 = the productivity of female workers

The alternate form of the above null hypothesis can be formulated as follows:

H_A : Male workers will have more productivity than female workers, or female workers will have less productivity than male workers.

Statistically expressed: H_A : $\mu_1 > \mu_2$

where, H_A = the alternate hypothesis

μ_1 = the productivity of male workers

μ_2 = the productivity of female workers

From the above example, it is clear that an alternative hypothesis, which is the opposite of the null, is a statement expressing a relationship between two variables or indicating differences between groups. The following are some other examples of null and alternative hypotheses:

• H_0: There is no relationship between working conditions and job satisfaction of employees.

H_A: If the working conditions are improved, then the job satisfaction of employees will improve.

H_0: There is no difference between male and female workers in their organizational commitment.

H_A: Male workers will have greater organizational commitment than female workers.

■ H_0: There is no relationship between pay and productivity.

H_A: Pay and productivity are positively related.

■ H_0: Working condition, pay and fringe benefits have no influence on job satisfaction of workers.

H_A: Working conditions, pay and fringe benefits all have positive influence on job satisfaction of workers.

Stating the Null Hypothesis

■ There is no difference between the means of the two populations from which the two samples were drawn at random.

■ The two means in the two populations from which the samples were respectively drawn at random are equal.

Criteria of Good Hypothesis Statement

The main requirement of hypothesis formulation is that it should fulfill certain basic criteria. Many different criteria can be found in the literature over what are the desirable qualities of a "good" hypothesis. Mason and Bramble (1997) outline the important features (criteria) of good hypothesis statement as follows: •

- Hypothesis should be stated in declarative form.
- Hypothesis should state the expected (articulated) describe a relationship between two or more variables.
- Hypothesis should be testable empirically.
- Hypothesis should be limited in scope.
- Hypothesis should be clearly and precisely stated. There should be no ambiguity in the variables or the relationships proposed.
- Hypothesis should state the conditions and circumstances under which it is supposed to apply.
- Hypothesis should reflect a guess at a solution or outcome to a problem based upon some knowledge, previous research, or identified needs .It should be consistent with most known facts.

Linkage between Research Hypothesis and Statistical Hypothesis

A research hypothesis is a specific and focused hypothesis that guides a research project or study. It is derived from a broader research question or problem, and it reflects the objectives and expectations of the researcher. For example, a research question could be: "How does gender affect the academic performance of students?". A research hypothesis could be: "Female students have higher grades than male students".

A statistical hypothesis is a hypothesis that can be formally tested with statistical methods and techniques. It is usually expressed in terms of parameters or distributions of a population or a sample. For example, a statistical hypothesis could be: "The mean height of men is equal to 175 cm".

In statistical hypothesis there are two types, null and alternative. Null is default hypothesis as statistical tests are designed to test homogeneity only. Almost Statistical tests can't test heterogeneity. This restriction is important as research question may go along with Null or

Alternative hypothesis. It is also useful to avoid further confusion as whatever research question default hypothesis will be null.

A research hypothesis is the proposed answer of the research question. The research hypothesis usually includes an explanation (‘x affects y because …’). A statistical hypothesis, on the other hand, is a mathematical statement about a population parameter. Statistical hypotheses always come in pairs: the null and alternative hypotheses.

Literature Review

Introduction

In terms of a literature review, "the literature" means the works you consulted in order to understand and investigate your research problem. Re-view (or look again) is a process of systematic, meticulous, and critical summary of the published literature in your field of research. How others have dealt with topics in your research subject and of what knowledge they have acquired? Literature review also indicates that you should summarize the broad contents of the research articles or studies and indicate clearly any linkages with other studies in the field.

The following are some definitions which explain the meaning, purposes and functions of literature review:

F. Cardesco and E.M. Gatner (1986): "A literature review is a self-contained unit in a study which analyzes critically a segment of a published body of knowledge through summary, classification, and comparison of prior research studies and theoretical articles."

P. Haywood and E.C. Wragg (1996): "A literature review is the process of locating, obtaining, reading and evaluating the research literature in the area of your interest."

N. Walliman (2010): "A literature review (or overview) is a summary and analysis of current knowledge about a particular topic or area of enquiry."

Review of literature is, thus, an essential part of all research studies. It is a way to discover what other research in the area of your problem has uncovered. A critical review of the literature helps you to develop a thorough understanding and insight into previous research works that relates to your study. It is also a way to avoid investigating problems that have already been definitely answered.

Example: A social scientist is interested to study the impact of social mobilization program on poverty alleviation. From his or her knowledge and experience in the field, he or she knows that a body of knowledge exists about the methods and policies of social mobilization, but he or she also knows that the boundaries of this body of knowledge are constantly expanding. Other social scientists have worked in the same area and have, no doubt, contributed new information to the field. Therefore, he or she seeks to identify these new contributions and add them to the established body of knowledge before he or she conducts his or her own investigation.

Purpose of Literature Review

Scientific research must be based on past knowledge. The previous studies cannot be ignored because they provide the foundation to the present study and provide you with a handy guide to a particular topic. In other words, there has to be continuity in research. This continuity in research

is ensured by linking your study with the past research studies. The primary purposes of literature review are: to learn how others have defined and measured key concepts; to identify data sources that other researchers have used; to identify potential relationships between concepts; and to identify researchable hypotheses. The literature review enables you to know:

- What research has been done in the subject?
- What others have written about the topic?
- What theories have been advanced?
- What approaches were taken by other researchers?
- What were the areas of agreement or disagreement?
- Whether there are gaps that can fill through the proposed research?

The purpose of literature review is, thus, to find out what research studies have been conducted in your chosen field of study, and what remains to be done. Hence, a literature survey helps you to avoid needless duplication of effort. No matter what topic you choose, chances are that someone has already done research on it. If so, then conducting your research as originally planned would be a waste of time and resource. Specifically stating, the purposes of literature survey are as follows:

- To give continuity in research.
- To place the research in a historical context to show familiarity with state-of-the-art developments.
- To synthesize and gain new perspective, get more insight into the topic and know about the knowledge status of the proposed subject.
- To draw a theoretical framework and define the research parameters.
- To discover important variables relevant to the topic.
- To generate hypotheses.
- To identify the methodology and techniques of research.

The literature survey provides the foundation for developing a comprehensive theoretical framework from which hypothesis can be developed for testing. The literature survey also minimizes the risk of pursuing dead-ends in research. Another advantage of reviewing the literature applies to the design phase of your project. Designing a study involves several decisions as to what variables to include and how to measure them, what techniques to use, what procedures to follow, and so on. A literature review provides you with a rich source for addressing these important design questions. Yet another purpose is that a review of the literature keeps you up to date on current empirical or theoretical controversies in a particular research area. As science progresses, new ideas develop concerning age-old managerial issues. Such controversies not only provide a rich source of research ideas but also give direction to specific research hypotheses and designs (Bordens & Abbott, 2011).

Need of literature review:

- It demonstrates that you know the field. This means more than reporting what you have read and understood. Instead, you need to read it critically and to write in such a way that shows you have a feel for the area; you know what the most important issues are and their relevance to your work; you know the controversies and things that are neglected.
- It justifies the reason for research. This is closely connected with demonstrating that you know the field. It is the knowledge of your field which allows you to identify the gap which your research could fill.
- It allows you to establish your theoretical framework and methodological focus. Even if you are proposing a new theory or a new method, you are doing so in relation to what has been done.

Types of Literature Review

There are different types of literature review that can be undertaken, depending on the purpose of the research. The main kinds of literature review are as follows:

- **Historical review:** This type of literature review traces the issues, concepts or events over time.
- **Methodological review**: This kind of review assesses and evaluates methodological techniques used and the strengths of different studies.
- **Theoretical review:** This type of review focuses on the theories or concepts related to the research issue under study.
- **Integrative review:** This type of review summarizes and integrates the current state of knowledge on the topic under study.

  These kinds of literature reviews are not mutually exclusive. These are often mixed together. In the case of academic research, all these types of literature review need to be undertaken.

Functions of Literature Review

The review of literature accomplishes the following functions:

- Ensures that you are not "reinventing the wheel".
- Gives credit to those who have laid the groundwork for your research.
- Demonstrates your knowledge of the research problem.
- Demonstrates your understanding of the theoretical and research issues related to your research.
- Shows your ability to critically evaluate relevant literature information.
- Indicates your ability to integrate and synthesize the existing literature.

- Convinces your readers that your proposed research will make a significant and substantial contribution to the literature.

Encyclopedias

■ ***Encyclopedia Britannica***: It is an excellent introduction to almost any field. It includes features and relatively long articles, which may be relied upon for authenticity and scholarly quality. The original British version of the encyclopedia is sometimes reflected in the fuller treatment given certain English topics. If an American orientation is desired, the Encyclopedia Americana should be consulted. Both these encyclopedias offer extensive bibliographies. Britannica also includes a yearbook — annual, which summarizes the events and cultural trends of the preceding year.

■ ***Britannica Online:*** This is an online version of Encyclopedia Britannica. Now with the online version, the task of locating materials, events and bibliographies has become much quicker and simpler.

■ ***Encyclopedia of the Social Sciences:*** It is the first comprehensive encyclopedia covering all fields of social sciences. Though it is international in scope, its emphasis is on English-speaking and Western European nations. It includes signed articles by specialists with adequate bibliographies. This encyclopedia is a good source for biographical articles.

■ ***International Encyclopedia of the Social Sciences***: International Encyclopedia of the Social Sciences is not meant to replace the earlier Encyclopedia of the Social Sciences. Instead, they should be used together with the International Encyclopedia emphasizing recent developments and an analytic comparative approach to a subject (e.g. "Comparative Politics"). Once again, articles are written by specialists and contain up-to-date bibliographies.

■ ***Encyclopedia of Education:*** Authoritative articles are included covering the history and theory of education, structure of education, structure of educational systems in various countries, research in education, important people and educational institutions, etc. This encyclopedia should be used in conjunction with the detailed index (Vol. 10). Its orientation is based primarily on education in the US; however, a number of articles treat international or comparative topics.

• ***McGraw-Hill Encyclopedia of Science and Technology:*** Clearly written articles, intelligible to the non-specialist, treat the basic subject matter of natural sciences, including their major technological applications in engineering, agriculture, forestry, etc. Articles are profusely illustrated and have short bibliographies.

■*Business Encyclopedia and Legal Adviser.* This encyclopedia includes articles written by professionals in accounting, banking, journalism, commerce and industry and explains the concepts significant to business, including the legalities involved.

Difference between Reference and Bibliography

While writing an assignment, article or book, the writer often looks for the sources to generate an idea or data. In this context, students usually misinterpret bibliography for reference, but they are different, in the sense that you give **reference** to the sources, that you have quoted in-text, in the research report or assignment. But on the other hand, in the **bibliography**, you create a list of all the sources you have gone through to conceive the idea.

Reference and Bibliography is an important part of any project under study because it helps in acknowledging other's work and also help the readers in finding the original sources of information. It not only prevents plagiarism but also indicates that the writer has done good research on the subject by using a variety of sources to gain information.

Reference Vs Bibliography Comparison Chart

BASIS FOR COMPARISON	REFERENCE	BIBLIOGRAPHY
Meaning	Reference implies the list of sources, that has been referred in the research work.	Bibliography is about listing out all the materials which has been consulted during the research work.
Based on	Primary Sources	Both Primary and Secondary Sources
Arrangement	Alphabetically and numerically	Numerically
Includes	Only in-text citations,that have been used in the assignment or project.	Both in-text citations and other sources, that are used to generate the idea.

BASIS FOR COMPARISON	REFERENCE	BIBLIOGRAPHY
Supporting argument	A reference can be used to support an argument.	A bibliography cannot be used to support an argument.
Used for	Thesis and Dissertation	Journal Papers and Research work

Definition of Reference

Reference can be understood as the act of giving credit to or mentioning the name of, someone or something. In research methodology, it denotes the items which you have reviewed and referred to, in the text, in your research work. It is nothing but a way to acknowledge or indirectly showing gratitude, towards the sources from where the information is gathered.

While using references, one thing is to be noted that you go for reliable sources only, because it increases credence and also supports your arguments. It may include, books, research papers, or articles from magazines, journals, newspapers, etc., interview transcripts, internet sources such as websites, blogs, videos watched, and so forth.

These are used to inform the reader about the sources of direct quotations, tables, statistics, photos etc. that are included in the research work.

Definition of Bibliography

At the end of the research report, bibliography is added, which contains a list of books, magazines, journals, websites or other publications which are in some way relevant to the topic under study,that has been consulted by the researcher during the research. In finer terms, it comprises of all the references cited in the form of footnotes and other important works that the author has studied.

The bibliography is helpful to the reader in gaining information regarding the literature available on the topic and what influenced the author. For better presentation and convenient reading, the bibliography can be grouped into two parts, wherein the first part lists out the names of books and pamphlets consulted, and the other contains the names of magazines and newspapers considered.

Types of Bibliography

- **Bibliography of works cited**: It contains the name of those books whose content has been cited in the text of the research report.
- **Selected Bibliography**: As it is evident from the name itself, selected bibliography covers only those works which the author assumes that are of major interest to the reader.
- **Annotated Bibliography**: In this type of bibliography, a small description of the items covered is given by the author to ensure readability and also improve the usefulness of the book.

Key Differences Between Reference and Bibliography

The difference between reference and bibliography can be drawn clearly on the following grounds:

1. Reference implies referring to someone or something, that means it provides the list of sources, whose text is used in the assignment or research work. Conversely, bibliography represents the list of all the sources, from which the research has gained some information about the topic, irrespective of the work cited or not.
2. References are based on primary sources, whereas bibliography is created on the basis of primary and secondary sources.
3. References used in the assignment can be arranged alphabetically or numerically. On the contrary, list of sources used in the bibliography is arranged numerically.
4. The bibliography is used to list out everything you go through to obtain the information relating to the assignment, no matter if you specifically cite it in your assignment or not. Now coming to references, it only takes into account those sources which have been cited in the assignment.
5. The main objective of adding a reference at the end of the document is to improve credence or support an idea or argument. As against, the bibliography is not used for supporting an argument.
6. While reference is used in thesis and dissertation. On the other hand, bibliography is used in case of journal paper and research work.

Conclusion

To sum up, references and bibliography are almost same, but there are only subtle differences between the two, which lies in the items which are included in them. The primary use of references is to get recognition and authentication of the research work, whereas bibliography is appended with the aim of giving the reader the information on the sources relating to the topic.

Research Design

Introduction

When a particular research area has been identified, research problem defined, and the related literature in the area have been reviewed; the next step is to construct the research design. A research design is the plan of attack: What approach to the problem will be taken? What methods will be used? What strategies will be most effective?

Definition

Fred N. Kerlinger (1986): "Research design is the plan, structure, and strategy of investigation conceived so as to obtain answers to research question. The plan is the overall scheme or program of the research. It includes an outline of what the investigator will do from writing the hypotheses and their operational implications to the final analysis of data".

John W. Creswell (2011): "Research designs are plans and the procedures for research that span the decision from broad assumptions to detailed methods of data collection and analysis."

William Zikmund (2013): "Research design is a master plan specifying the methods and procedures for collecting and analyzing the needed information".

Essential elements of a research design

- A research design is an *overall plan* for the activities to be undertaken during the course of a research study.
- The research design serves as a *framework for the study*, guiding the collection and analysis of the data, the research instruments to be utilized, and the sampling plan to be followed.
- It is an *organized and integrated system* that guides the researcher in formulating, implementing, and controlling the study.
- The research design is a *blueprint* specifying the method to be adopted for gathering and analyzing data.

- The research design is a *strategy* of obtaining information for the purpose of conducting a study and making generalizations about the population.

In planning a research investigation, choices have to be made about research strategy (experimental vs non-experimental), research setting (laboratory vs natural setting), measures (questionnaires, observations, interviews), the data analysis strategies (descriptive vs inferential

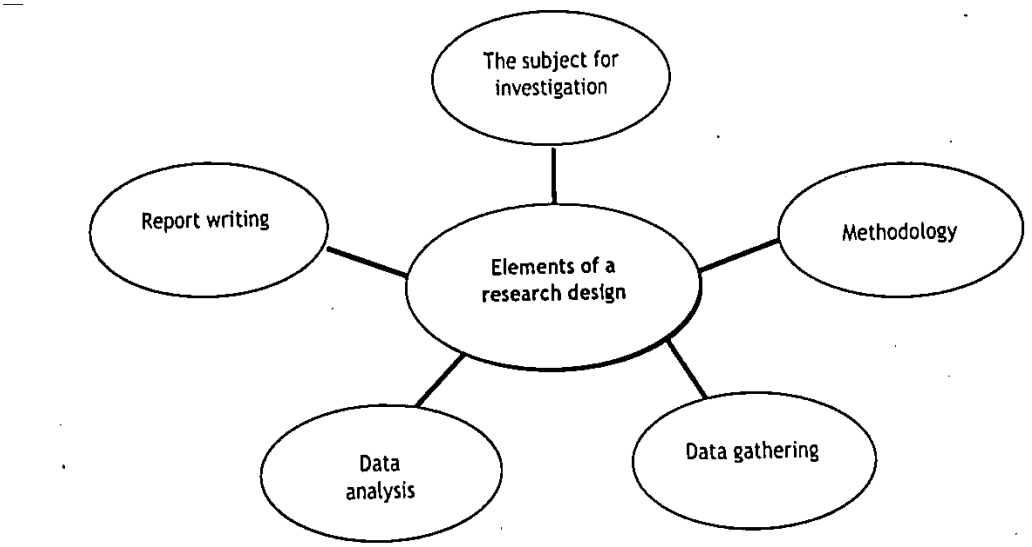
statistics), and a host of other factors. A research design thus includes all these essential factors of an investigation.

Good research is not accidental. It requires careful planning as well as careful execution. Before launching a research project, you have to prepare a research design, in which you have to set forth your plans for the research you intend to undertake.

Research design is like a philosophy of life; no one is without one, but some people are more aware of theirs and thus able to make more informed and consistent decisions. Similarly, every type of empirical research has an implicit, if not explicit, research design. Because a design always exists, it is important to make it explicit, to get it out in the open where its strengths, limitations, and implications can be clearly understood (Maxwell, 2013).

Elements of a Research Design

The basic elements of a research design are (a) the problem, (b) the methodology, (c) data gathering, (d) data analysis, and (e) report writing. These elements of research design have been shown in figure. A good research design considers all these elements. The first element of a research design is to answer the research question or test research hypothesis.



Every research work usually requires an explanation of the methodology and the sample description. What methods were used to choose the sample? Why these methods were chosen and how they were applied? Next, there should be an explanation of what the variables are in the hypothesis and how they were measured. Furthermore, the details of the data collection must he explained and a discussion on the reliability and validity of the measurements included. Finally, it is necessary to explain how the data were analyzed.

Preparation of the Research Design

A research design is a clearly planned procedure for carrying out the research. Many things need to be planned in advance. The design generally incorporates answers to the following kinds of questions (Oliver, 2011):

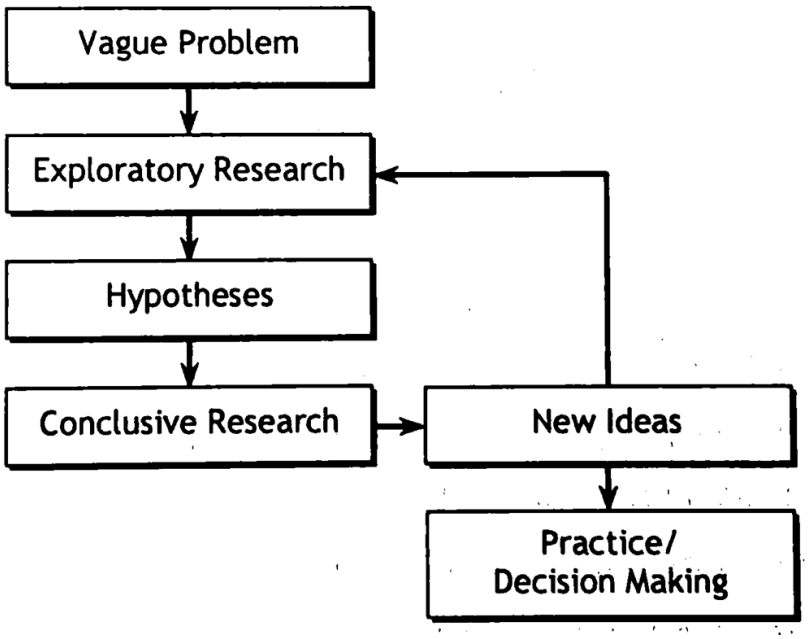
- What sort of data do I need to collect in order to test the hypothesis and/or achieve research aims?
- Where will I collect the data?
- How will I collect the data?
- What type of data-collection instruments and procedures will I use?
- Who will provide me the data?
- Do I need to ask permission before trying to collect data?
- When will I collect the data?
- How will the data be analyzed?
- Will I use a particular theoretical frame in order to interpret the data?

Exploratory Research Design

The relationship and sequence of research activities are shown in figure. When searching for hypotheses, exploratory designs are appropriate. When hypotheses have been established and are to be tested, conclusive research designs are needed. Figure given below, highlights the sequence of research activities, from vague problem to new idea generation. First, let us define exploratory research. An exploratory research is defined as "a study undertaken in areas where very little prior knowledge or information is available on the subject under investigation". It is thus the initial research conducted to study and define the nature of a problem. An exploratory study is undertaken when we do not know much about the situation at hand. In such cases, extensive preliminary work needs to be done to gain familiarity with the phenomenon in the situation.

The purpose of exploratory research is to achieve new insights into a phenomenon. The major emphasis in those studies is the discovery of new insights or ideas. There are three purposes for exploratory research:

- Diagnosing a situation
- Screening alternatives
- Discovering new ideas



An exploratory study is undertaken to orient the researcher and the study. It is, therefore' a important method of finding out what is happening, to see new insights, to ask questions, and assess phenomena in a new light. It is particularly useful if you wish to clarify your understanding of a problem. In such study, the focus is initially broad and becomes progressively narrower in the research progresses.

Exploration is thus the first stage of any research project, which is new and unexplored. It, important to note that doing a study for the first time in a particular issue does not make the research exploratory in nature. Only when knowledge is scant and a deeper understanding is needed, the study becomes exploratory.

Characteristics

■ There is no set method of conducting exploratory research. The key requirements for this research are: imagination and flexibility. It is less structured and more flexible.

■ Exploratory research studies are not characterized by formal research design. Hence, they are not very scientific in nature.

■ The researcher may utilize any number of informal approaches in attempting to define the problem and gather the data.

■ Exploratory research provides low-risk form of research that may result in good outcomes. A clear picture of the situation can emerge leading to hypothesis formulation.

■ Exploratory research provides direction for a more formal research effort.

Descriptive Research Designs

Descriptive research describes phenomena as they exist. Such studies involve the systematic collection and presentation of data to give a clear picture of a particular situation. These studies attempt to obtain a complete and accurate description of a situation. These studies can be classified in the following five categories:

(a) Historical

(b) Descriptive

(c) Developmental

(d) Survey

(e) Case Study.

These five types of descriptive research designs are not mutually exclusive. A combination of all these could also be used in some research projects. We will now look at each of them in some greater detail.

a. Historical Research

History is a meaningful and an organized record of past events. It is not merely a list of events arranged chronologically, but a valid integrated account of social, cultural, economic, and political forces that had operated to produce a historical event.

Historical research is concerned with past phenomena. It can be defined as "the systematic and objective location, evaluation, and synthesis of evidence in order to establish facts and draw conclusions about past events." Historical research is thus a process of collecting, evaluating, verifying, and synthesizing past evidence systematically and objectively to reach a conclusion. Historical research may also attempt to discern trends in the past and reconstruct the origin and development of those events. The main purpose of conducting historical research is to show the relevance of past events to the present. In other words, the purpose is to arrive at an accurate account of the past so as to gain a clearer perspective of the present.

Accuracy of gathered information is the main ingredient of success in historical research. There are two main sources from where past evidences can be found. One is the primary source, where you were a direct observer of the recorded event and the other is the secondary source, where you are reporting the observations of others. In most cases, you have to depend upon the data observed by others rather than by yourself. At the same time, you must also be aware that inappropriate and biased information results in faulty conclusions and findings.

Characteristics

- Good historical data result from painstaking detective work which analyzes the authenticity, accuracy, and significance of source material.
- Historical research must be rigorous, systematic and exhaustive.
- Historical research depends upon two kinds of data: primary sources where the author was a direct observer of the recorded event and secondary sources where the author is reporting the observations of others and is one or more times removed from the original event. Of the two primary sources carry the authority of firsthand evidence and have priority in data collection.
- This critical evaluation of the data is what makes true historical research so rigorous - in many ways, more demanding than experimental methods.
- While historical research is similar to the, "reviews of the literature" which precede other forms of research, the historical approach is more exhaustive, seeking out information from a larger array of sources.

b. Descriptive Research

Descriptive research is a fact-finding operation searching for adequate information. It is a type of study, which is generally conducted to assess the opinions, behaviors, or characteristics of a given population and to describe the situation and events occurring at present. Descriptive research is a process of accumulating facts. It does not necessarily seek to explain relationships, test hypotheses, make predictions or get at meanings and implications of a study. Hence, a descriptive research is an extension of an exploratory research.

Descriptive research can be either quantitative or qualitative. This research involves gathering data that describes events and then organizes, tabulates, depicts, and describes the data collection. Descriptive statistics is used to reduce the data to manageable form. Descriptive research is unique in the number of variables employed:

- Descriptive research, like other types of research, can include multiple variables for analysis.
- Descriptive research might simply report the percentage summary on a single variable.

Descriptive studies thus simply portray an accurate profile of organizations, events, or situation. Investigators collect, classify, and correlate data to describe what exists. However, they do not fully analyze and explain why phenomena behave as they do. They do not put the relationships they describe to crucial experimental tests. Although descriptive research cannot predict and control conditions and events, it contributes to science primarily by building a foundation of facts upon which exploratory hypotheses may be constructed, by checking the validity of

existing theories, and by directing attention toward alternative hypotheses which better fit the facts (Van Dalen, 1973). In a descriptive research, it is necessary to have a clear picture of the phenomena on which you wish to collect data prior to the collection of data. Isaac (1978) identifies the characteristics and steps in a descriptive research as follows:

Characteristics

Nature

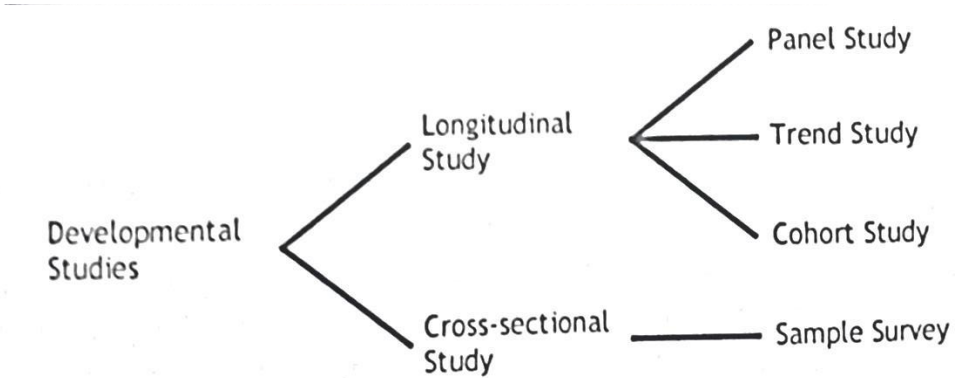
- Descriptive research is used in the literal sense of describing situation or events.
- It is accumulation of a database that is solely descriptive - it does not necessarily seek or explain relationship, test hypotheses, make predictions, or get at meanings and implications, although research aimed at these more powerful purposes may incorporate descriptive methods.

Purposes of Descriptive Studies

- To collect detailed factual information that describes existing phenomena.
- To identify problems or justify current conditions and practices.
- To make comparisons and evaluations.
- To determine what others are doing with similar problems or situations and benefit from their experience in making future plans and decisions.

Developmental research

Developmental research is conducted for the purpose of predicting future trends. It concentrates on the study of variables, their rates of change, direction, sequences and other inter-related factors over a period of time, there are several methods of developmental research. The nature of these different forms of developmental research design is shown in the following figure:



Longitudinal Study

It is a research where phenomena are studied over time either continuously or repeatedly. This type of study measures the nature and rate of change in a sample at different stages of development. This occurs when the data are collected in two or more points in time from the same group of individuals. Because data are gathered in two different points in time, it is not a cross-sectional or a one-shot study, but it is a study carried longitudinally across a period of time. The phenomena are studied over time either continuously or repeatedly. Longitudinal studies are mostly quantitative.

Trend Study

The trend study is probably the most common longitudinal study among others. When the data are collected at intervals spread over a period of time, it is called a trend study. This type of research samples different groups of people at different points in time from the same population. It is designed to establish patterns of change in the past in order to predict future patterns or conditions. A trend study thus provides information about net changes at an aggregate level .It can establish a pattern over time to detect shifts and changes in some event. Marketing firms, for example, compile trend studies that chart fluctuations in consumption level for a certain product. Trend studies do not have to be conducted by just one researcher or research project. You may combine data from several studies of the same population in order to show a trend. This type of

study is particularly used to obtain and analyze social, economic, and political data to identify trends and to predict what is likely to take place in the future. Frequently regression analysis is used for trend studies.

Cohort Study

A cohort is a group of people who share a common characteristic or experience within a defined period. Thus, cohort study is a study of a specific group, such as those born on a day or in the particular period, say in the year 2003. This group then forms a birth cohort or a kindergarten cohort. Similarly, a group of students graduating from college in a year form a student cohort. There are many other kinds of cohorts, including disease, education, employment, housing, family formation, and the like.

A sample of the selected cohort group is then studied at different points of time. A cohort study is thus a systematic follow-up of a group of people for a defined period of time or until a specified event. To form cohort studies, data are compiled for the same population over time. Such studies are therefore rare because of the difficulty of maintaining contact with members of the cohort from year to year.

Panel Study

A panel is a group of individuals that have agreed to provide information to a researcher over a period of time. In panel study we take the same people and study their attitudes towards a particular phenomenon over time. Panel studies are most useful when studying change. These studies allow the researcher to find out why changes in the population are occurring. They measure the same sample of respondents at different points in time. For example, if we were interested in finding out the general attitude towards single parenthood, we would take a group of people and interview them at periodic intervals on the same subject and over a number of years. Unlike trend studies, panel studies can reveal both net change and gross change in the dependent variable. Additionally, panel studies can reveal shifting attitudes and patterns of behaviour. Depending on the purpose of the study, You can use either a continuous panel (the same respondents reporting on a regular basis) or an interval panel (reporting only when the information is needed). Panel data are particularly useful in predicting long-term or cumulative effects which are difficult to analyze in a cross-sectional study.

Cross-sectional Study

This type of study is also known as cross-sectional analysis. It involves observation of some items of the population all at the same time. This study basically measures the rates of changes by drawing samples from a cross-section of society. It focuses on comparing and describing groups. In this study, data are gathered just once, perhaps over a period of time, in order to answer a research question. Such studies are also known as one-shot studies. Cross-sectional studies often employ the survey strategy. The fundamental difference between a cross-sectional

study and longitudinal study is that a cross-sectional study takes place at a single point of time and that a longitudinal study involves a series of measurements taken over a period of time.

Characteristics of Development Research

- Developmental research focuses on the study of variables and their development over a period of months or years. It asks, "What are the patterns of growth, their rates, their directions, their sequences, and the interrelated factors affecting these characteristics?"
- The sampling problem in the longitudinal method is complicated by the limited number of subjects it can follow over the years; any selected factor affecting attrition biases the longitudinal study.
- Once underway, the longitudinal method does not lend itself to improvements in techniques without losing the continuity of staff and financial support over an extended period of time and typically is confined to university or foundation centers that can maintain such an effort.
- Cross-sectional studies usually include more subjects, but describe fewer growth factors than longitudinal studies. While the latter is the only direct method of studying human development, the cross-sectional approach is less expensive and faster since the actual passage of time is eliminated by sampling different subjects across age ranges.
- Sampling in the cross-sectional method is complicated because the same children are not involved at each age level and may not be comparable. To generalize intrinsic developmental patterns from these sequential samples of children runs the risk of confusing differences due o development with other differences between the groups that are artifacts of the sampling process.
- Trend studies are vulnerable to unpredictable factors that modify or invalidate trends based on the past. In general, long-range predication is an educated guess while short-range prediction is more reliable and valid.

Survey Research

A survey is a means of gathering information about the characteristics, actions, or opinions of a large group of people, referred to as a population. A survey research is thus defined as "*the systematic gathering of information from respondents for the purpose of understanding and/or predicting some aspect of the behavior of the population of interest"* (Tull & Hawkins, 1997, p. 164). A survey study is perhaps the dominant form of data collection in social science, today. If conducted scientifically, this type of research can contribute to the advance of knowledge. A Survey study may be done in the field — an example would be a survey of employee attitudes toward a new compensation policy — or it may take place in a library, where a survey of secondary literature is conducted. In survey research, the researcher selects a sample of respondents from a population and administers a standardized questionnaire to them. Hence, using surveys, it is possible to collect data from large or small populations.

Types of Survey Research

- Exploratory survey research. This type of survey research takes place during the early stages of research. It provides the basis for more in-depth surveys. Sometimes, this kind of survey is carried out using data collected in previous studies.
- Confirmatory (theory-testing or explanatory) survey research. In this type of survey, data collection is done with specific aim of testing the theory or hypothesis.
- Descriptive survey research. This type of survey research describes the distribution of the phenomenon in a population. Through facts described, it can provide useful hints both for theory building and for theory refinement.

The aim of survey research is to measure certain attitudes or behavior of a population or a sample. The attitudes might be opinions about the services provided by a business firm or feelings about certain issues or practices. Most often respondents are asked for information. Surveys come in a wide range of forms and can be distributed using a variety of media: written surveys, oral surveys, or electronic surveys. The questionnaire, or survey, can be a written document that is completed by the person being surveyed, on online questionnaire, a face-to-face interview, a mail survey, or a telephone interview. Some forms of survey research by telephone or Internet may be completely automated. Using surveys, it is possible to collect data from large or small population.

The methods researchers use when designing, conducting, analyzing, and reporting the results of surveys vary according to research goals, the population surveyed, and the type of survey that is used. Survey research has also changed dramatically in the last ten or fifteen years. Today, we have automated telephone surveys that use random dialing methods. A whole new variation of group interview has evolved as focus group methodology.

Case Study Research

Case study research is an important approach to study the topics in social science and management. Case studies are written summaries or synthesis of real-life cases based upon data and research. A case study is thus defined as *"a strategy for doing research which involves an empirical investigation of a particular contemporary phenomenon within a real-life context using multiple sources of evidence"* (Robson, 1993, p. 54).

Yin (1994) defines the case study research as "*an empirical inquiry that investigates contemporary phenomenon within its real-life context"* (p.79). Rather than using samples to examine a limited number of variables, case study methods involve in in-depth, longitudinal examination of a single instance or event (case). This research thus views a social or study unit as a whole in its real-life context. This study phenomenon could be a person, a family, a social group, an institution, a community, or even an entire culture.

A case study not only uses the sources and techniques of historical story but also employs several techniques and sources of data for examining current aspects of the phenomenon under study. A case study could be conducted in the field as well as a non-field setting (Shah,1972, p.11).

The investigator gathers pertinent data about the present status, past experiences and environmental forces that contribute to the individuality and behavior of the unit. After analyzing the sequences and inter-relationships of these facts he or she conducts a comprehensive study of the social unit as it functions in society.

Case studies need to be both comprehensive and systematic. That is, as much data as possible need to be collected in a way that ensures as little as possible is missed. Jensen and Rodgers (2001) set forth a typology of case studies as follows:

■ *Snapshot case studies are the detailed study of one unit.*

■ *Longitudinal case studies are studies of the same unit at multiple time points.*

■ *Pre post ease studies are undertaken at two time points separated by a critical event.*

■ *Cross-cute studies are studies of multiple case studies for the purpose of comparison*.

When selecting a case for a case study, researchers often use information-oriented sampling, as opposed to random sampling. This is because the typical case is often not the richest in information. Information-oriented cases may be distinguished as: extreme cases, critical cases, and exemplar cases. Extreme cases reveal more information because they activate more basic mechanisms and more actors in the situation studied.

It is sometimes impossible for you to handle the whole social reality; at other times, the conceptual basis for understanding some aspects of social reality is not available. In such cases, you may first want to explore the social reality before you formulate and test specific hypotheses. But you must recognize that a case does not represent the total reality. It is just one example of the social reality. Hence, a case study may be an intensive, integrated and insightful method of studying the social phenomena. It can also be used to illustrate a theory by providing an example.

In recent years, there has been increased attention to implementation of case studies in a systematic manner which increases the validity of associated findings. However, although case study research may be used in its own right, it is more often recommended as part of a multi-method approach (triangulation).

Characteristics

■ Case studies are in-depth investigations of a given social unit resulting in a complete, well-organized picture of that unit. Depending upon the purpose, the scope of the study may encompass an entire life cycle or only a selected segment; it may concentrate upon specific factors or take in the totality of elements and events.

■ Compared to a survey study which tends to examine a small number of variables across a large sample of units, the case study tends to examine a small number of units across a large number of variables and conditions.

■ Because case studies are intensive, they bring to light the important variables, processes and interactions that deserve more extensive attention. They pioneer new ground and often are the source of fruitful hypotheses for further study.

■ Case study data provide useful anecdotes or examples to illustrate more generalized statistical findings.

■ Because of their narrow focus on a few units, case studies do not allow valid generalizations to the population from which their units came until the appropriate follow-up research is accomplished, focusing on specific hypotheses and using proper sampling methods.

.

Limitations of Case Study

- A case study is more expensive because of its exploratory nature.
- A generalization drawn from a single case cannot be applied to all cases in a given population
- There is some element of subjectivity. You must guard against permitting personal biases and standards to influence your interpretation.

Analytical Research Design

Analytical Research designs can be experimental or observational and each type has its own features. A study design is critical to the research study because it determines exactly how we will collect and analyze our data. If we aim to study the relationship between two variables, then an analytical study design is the right choice.

But how do we know which type of analytical study design is best for our specific research question? It's necessary to have a clear plan before we begin data collection. Lots of researchers, sadly, speed through this or don't do it at all .Analytical study designs can be experimental or observational and each type has its own features.

When are analytical Research designs used?

A study design is a systematic plan, developed so you can carry out your research study effectively and efficiently. Having a design is important because it will determine the right methodologies for your study. Using the right study design makes your results more credible, valid, and coherent.

Descriptive vs. analytical Research

Study designs can be broadly divided into either descriptive or analytical.

Descriptive studies describe characteristics such as patterns or trends. They answer the questions of what, who, where, and when, and they generate hypotheses. They include case reports and qualitative studies.

Analytical study designs quantify a relationship between different variables. They answer the questions of why and how. They're used to test hypotheses and make predictions.

Experimental and observational

Analytical study designs can be either experimental or observational. In experimental studies, researchers manipulate something in a population of interest and examine its effects. These designs are used to establish a causal link between two variables.

In observational studies, in contrast, researchers observe the effects of a treatment or intervention without manipulating anything. Observational studies are most often used to study larger patterns over longer periods.

Experimental Research Method

Introduction

The experimental method of research is used as the classical method in physical sciences. It is based on observation or experiments. It deals with actual experiments to determine the relationship between cause and effect of various experimental treatments. It is defined as 'the research method in which a researcher objectively observes phenomenon which is made to occur in a strictly controlled situation where one or more variables are varied and others are kept constant'.

The purpose of experimental research is to investigate cause and effect relationship by exposing one (or more) experimental groups to one (or more) treatment conditions & comparing the result to one (or more) control groups not receiving the treatments. In this method, the researcher undertakes control or manipulation (vary) of various variables under study. The usual approach is to hold all variables constant except one in controlled condition. By varying this one variable, the out puts (the effects) are studied and documented.

Actually, in social sciences, in natural sciences, in biological phenomena and to the human behavior control of variable is hardly possible. However, in physical sciences and experimental technology the investigation in controlled condition is highly acceptable.

Experiment is a test of a casual proposition, such as:

i) Do changes in variable 'A' cause changes in variable 'B' keeping other variables constant?

ii) ii) How the changes in the value of one variable (called independent variable) affect another variable (called dependent variable)?
The mathematical form of the experimental method is given below: If x_1, x2, x3, x4... xn are n independent variables taken as the inputs of the process and y is the output of the process (a dependent variable), then y is defined as a function x and denoted by,

$y = f(x)$

Where,

x means x_1, x_2, x_3, x_4..... x_n and f denotes the function

Suppose for an example, yield (y) of a product in an agricultural field is influenced by the following four different independent variables:

x_1 - seed quality (qualitative variable say, Si, S2),

x_2 - amount of fertilizer (quantitative variable, in kg),

x_3 - irrigation scheme (categorical variable say, I_1, I_2,I_3)

x4 - labor input (quantitative variable say, in number)

The production or yield, which depends upon these four variables, can be related mathematically as

$Y = f(x)$

Or $Y= f(x_1, x_2, x_3, x_4)$

By taking any three (say x_1, x_2, x_3) constant one can observe the effect of x_4 in Y; x_4 may vary as researchers will, so it is said to be a controlled variable.

The various factors in an experiment are divided into two groups: independent variables and dependent variables. The first sets of a factor or factors are called an experimental group and the second sets of factors are called control group. Control group is also known as a group of individuals, items or objects used as a standard for comparison or accepted norm. It is used to compare with or to evaluate the others among which a new process or method has been implemented.

An experimental or independent factor is such a cause, which is freely changed in an experiment; usually, an independent variable consists of a single factor only because an experimenter wishes to study the effect of a cause independent of all other influences. An experimenter tries to stabilize and control all the factors in an experiment; and then takes one of these for study. He varies it to see the effect of variations upon other factors. The variations of an independent variable produce responses in other factors and these responses are actually dependent variables. These responses are subject to the cause, which is being studied.

To make the experimental method of research effective and distinct from normal activity the method of local control (blocking) and statistical control methods are used. Control is necessary to reduce variations. In some experiments, some variables may be eliminated. Undesired variations can be reduced by standardization.

Types of Experiment

Experiment is the scientific investigation in which an investigator manipulates and controls one or more independent variables and observes the dependent variables for variation concomitant to the manipulation of the independent variable. There are four different types of experiments given as follows:

1. Positive and Negative Experiment If the subject of an experiment is such that i) the phenomenon and ii) its cause both are present, the experiment is said to be positive. For example, a bell rings in the air. Here both the sound and cause of its propagation are present; but if, on the other hand, a bell is rung in a vacuum, there will be no phenomenon of a sound because the cause of the propagation, namely air is absent. Such experiment is called negative experiment.

2. Natural Experiment These experiments are to be observed in natural phenomenon. In most of the natural state experiments, the controlling of variable is unnecessary (emphatically avoidable) to obtain. the real information about the phenomena. In such case, the whole phenomenon 15 divided into control group and experimental group to study the effects of seen and unseen variables**.**

3. Laboratory Experiment These are the experiments performed in physical sciences with full control of external conditions. A laboratory experiment is an artificially created situation in which the researcher controls one or more variables while manipulating other variable at will. The method of lab experiment is used in the experiments, mainly related to the physical, chemical, microbiological, clinical and such other sciences. If, it is difficult to conduct an experiment out-side or in the field or in the society then one tries to carry out it in the laboratories. In lab experiment, the controlling of the variable as required is possible**.**

4. Field Experiment Field experiments are the experiments conducted in the field or in natural setting. Research study in a realistic situation in which one or more independent variables are manipulated by the experimenter under carefully controlled conditions as the situation will permit. In the social, managerial, agricultural, environmental researches, the method of field experiment is widely used. In these experiments, the influence of apparently unconnected variables is minimized as possible and experiment is carried out to study the problem in its real setting or at its existence state. Some of the field experiments like agricultural or business field the controlling of the variable is possible but in the careful condition

Purposes of Experimental Method

The experimental researches are conducted for the following purposes.

a. To determine the effect of various treatments and to compare the differences of effects as significant or non-significant.

b. To estimate the interaction effects of various treatments and to compare them

c. To establish the mathematical relationship between various treatments and their effects.

Problem in Experimentation

1. **To single out one factor from the phenomena**

It is always difficult to single out one factor from a social phenomenon for the purpose of measurement, because in any event there may be many factors interacted**.**

2. **Controlling the factors**
   Control of factors sometimes is not possible, because some factors may be unknown and uncontrollable. It is better to select several random samples as experimental and control groups. One solution here is the adoption of the control group technique.
3. **To get data from the control groups**
   There are difficulties in getting data from the control groups. The remedy may be found in matching the control and experimental groups on as many points as possible.
4. **To assign the level of significance**
   The determination of the required level of significance of the differences between the experimental and control groups is also fraught with difficulty. What difference can be taken as fraught with difficulty? What difference can be taken as significant? There is the problem of value judgment. But the scientific criterion is the determination of the statistical test of significance. However, this requires a reliable and valid socio-metric scal**e.**
5. **Change in response of people**
   In field experiment related with human behaviors (society and clinical setup), when data collected through human interaction due to changes in time, situation, environment and types of questions to be asked people often changes their responses.
6. **Change in theme of trialing**
   Due to change in behavior of the respondents and unsatisfactory management of the investigator theme of trialing of the area under experiment (in social and clinical setup) may change at the end of the experiment from what it was started. Because of the changes made by experimentation may give different responses which may lead wrong conclusion.
7. **Problem of handling or operation**
   In social setup and to the medical trials, if the people under study area is not aware, attentive and responsive about the inquiry, true response cannot be possible; in such situation the problem of data handling (manipulation) or the problem of conducting the study (operation) is come to pass.

Steps in Experimental Methods

1. **Statement of the problem, research questions and the objectives**
   The first step in the application of field techniques is related in mentioning of problem, research questions and specific objective. The hypothesis, at this stage, should be stated explicitly in general terms.
2. **Examination of possible outcomes and events through literature**
   The second step consists in setting up the field experiment by thorough reading of the available literature. The factors to be controlled must be assessed; the cooperation

between the researcher and the subject must be set up; and scouting for information is required before any choice regarding setting is made.

3. **Design of experiment**

   The next step is the choice of experimental design regarding its size, material, control groups etc. The choice of material should be based on the criterion of maximum possible accuracy. The basic problem of design relates to control. Control and experimental groups should be matched on all important factors. In cases where conditions cannot be standardized, the significance of the factors can be deducted through various devices of measurement, such as trends, extrapolations etc.

4. **Performing experiment**

   The next step of this method is to performing experiment in predefined circumstances. The principles of randomization, replication and blocking should be implemented as can as possible. The sensitiveness of experiments can be augmented by neutralizing the biases through random choice, by increasing the replication, y improving the quantitative technique and by refinements of techniques.

5. **Analysis of experimental out comes statistically**

   The analysis of the experimental data should be done starting from stating the descriptive nature of the data, measuring relationship between them and modeling data into some mathematical models. The analysis of variance permits a study of complex interrelationship, which is not possible by simpler designs. It permits more reliable conclusions about more hypotheses with fewer cases than if hypotheses were tested in separate design. In the experimental designs, it is also possible to achieve a matching of the groups through an analysis of covariance.

6. **Drawing conclusions by measuring reliability**

   For an experimental research the conclusions are drawn based on the statistical significance testing. The tests can be performed as required level of design by the use of different statistical techniques. The results obtained then are put to test their reliability and the conclusions are made.

7. **Testing the validity of the conclusion**

   The validity of the results should be measured before disseminating the results and reports. The validity of the experimental results is checked by the comparing with other similar phenomenon or to the standards**.**

8. **Evaluation of the entire investigation through practice**

   The success of the experimental study can be measured only through putting into practice the experiments many times. If the repeated experiments give similar or better results, then the experimental results may be considered satisfactory**.**

Ex-Post Facto Research

According to Landman (1988) defined the term ex-post facto as an experiment in which researcher, rather than creating the treatment, examines the effect of naturally occurring

treatment after it has occurred. The literal meaning of Ex-post facto is 'from what is done afterwards'. It means 'something done or occurring after an event with a retroactive effect on the event.'

• The ex-post facto design attempts to trace out the prior cause from the present problem. It is an analysis from the present to the past. The design consists of the study of the same individuals at the present and at some prior period (or, the comparison of population of an area in the present or in the past). Here, the ordinary technique of induction is followed. The operations are simplified by control, through matching of several independent variables, which are related to dependent variables. However, in this design, the freedom of application in method and technique seems to be restricted because the facts are to be collected from the available past and present records only. Furthermore, the choice of causal and control factors also appears to be restricted because of the presence of only a handful of causal factors in this design.

Types of Ex-post Facto Study

1. **Retrospective studies**
   It is the method of an outcome linking with preceding phenomenon. Eg: A study relating cancer to cigarette smoking done by surgeon of Nepal. The study includes large numbers of people including those have lung cancer, who died of it & who did not have it. Cancer was the dependent variable & cigarette smoking was independent variable. The researcher observed that increased incident of lung cancer with increase of number of cigarettes smokes daily. The researcher concluded that cancer is related to cigarettes smoking with no assertion that the latter is the definite cause of the former.
2. **Prospective studies**
   It is the method in which, subjects are selected in the present situation & data is collected over a long period of time waiting until the presumed effect occurs. For the study of education upon future adjustability of the students, equal number of students from two groups should be selected keeping in the mind that other variable (age, economic status' intelligence) matches in both cases the study would start from now to the long time to get the result.

Procedures of Ex-post Facto Research

The following points give the procedure by which ex-post facto research is carried out.

- The Ex-post facto research may take a long duration because of its feature. It moves from present to future.
- For ex-post facto technique, a plan should be designed carefully to draw reliable and valid inferences.

- To compare the results of phenomena understudy two or more sub items or objects are set in the same circumstances of physical, social, economic and the other status.
- The different objects are left as it for a long time and external influences of the items or manipulation is restricted.
- After the duration completed the results are collected systematically and analyze to develop the inferences.

The useful result for the researcher is put down for the practice and follow.

Ethical Issues in Experimental Research Design

. The following practices are considered unethical:

- Putting pressure on individuals to participate in experiments through coercion, or applying social pressure.
- Deceiving subjects by deliberately misleading them as to the true purpose of the research.
- Exposing participants to physical or mental stress.
- Not allowing subjects to withdraw from the research when they want to.
- Using the research results to disadvantage the participants, or for purposes not to their liking.
- Not explaining the procedures to be followed in the experiment.
- Not debriefing participants fully and accurately after the experiment is over.
- Not preserving the privacy and confidentiality of the information given by the participants.
- Withholding benefits from control groups.

Research Guides

Research Guides are **librarian-curated pathways to information, videos, databases, and other resources for your discipline**. That is, they pull many different types of resources on a subject or topic together in one place.

Hand Book

A handbook is a compilation of miscellaneous information in a compact and handy form. It contains data, procedures, principles etc. Tables, Graphs, diagrams and illustrations are provided. Scientists and technologists use handbooks in their fields rather frequently.

A treaties on a special subject; often nowadays a simple but all-embracing treatment, containing concise information, and being small enough to be held in the hand; but strictly, a book written primarily for practitioners and saving for constant revision or reference. Also called a 'Manual'.
Example:

- Britain, 1948/49-, an official handbook, London, stationery Office, 1948-, Annual.
- Handbook of Chemistry and Physics: A ready reference book of Chemistry and Physical data, 52nd ed, Cleveland, Ohio, Chemical Rubber, 1971**.**

Academic database in the relevant Field

It is a collection of information that is commonly used for research and writing, including access to academic journals.

CITATION

A "citation" is **the way you tell your readers that certain material in your work came from another source**. It also gives your readers the information necessary to find the location details of that source on the reference or Works Cited page. A citation must include a set of parentheses.

For APA, IEEE and other reference style follow the web-link given below

https://pitt.libguides.com/citationhelp/apa7

Citation Index

Citation indexes allow researchers to trace the impact of an article upon later publications. Besides including the bibliographic information about an article (author, article title, journal title, date, etc.), citation indexes also provide each article's references or bibliography (the list of sources cited).

For citation Index follow the web link: **https://www.isko.org/cyclo/citation**

SCIFinder

Follow the link **: https://library.ulethbridge.ca/scifinder/overview#s-lg-box-wrapper-9172764**

SCOPUS

Follow the link: **https://www.aimlay.com/scopus-indexed-journals/**

Science direct:

ScienceDirect is **a website which provides subscription-based access to a large database of scientific and medical research**. It contains the world's largest electronic collection of full-text and bibliographic information on science, technology and medicine

Follow the link: **https://cscitconf.cikd.ca/an-introduction-to-sciencedirect/**

Impact Factor

Follow the link: **https://en.wikipedia.org/wiki/Impact_factor**

H-Index

Follow the link: **https://en.wikipedia.org/wiki/H-index#Calculation**

Data Analysis and Interpretations

MEASUREMENTS AND SCALES

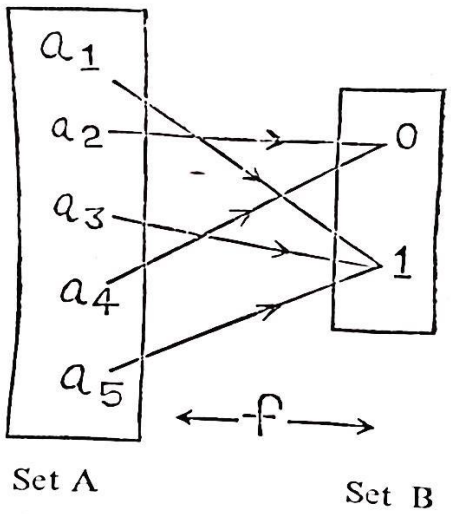
Definition of Measurement:

Measurement is the assignment of numerals to objects or events according to some rules. A numeral is a symbol used to distinguish objects from each other and has no quantitative meaning unless we give such a meaning. In other words, numerals are the symbols assigned to objects according to some prescribed rule. Numerals are usually the symbols 1. 2, 3...or I, II, III...etc ascribed to the objects such as players (Ex: Football players).

When numerals are assigned quantitative meaning, they become numbers. The term assignment means mapping of objects of **one set** onto the objects of **another se**t. A function, f, is the rule of correspondence.

Example 1: Suppose a family consists of five persons and we want to measure their sex. Assuming that we have a prior rule that allows unambiguously to determine the sex, the rule is to assign a symbol if the person is male and another symbol if the person is female. If the symbols used are 1 and 0 respectively, then we have two sets. A={a_1 , a_2 , a_3 , a_4, a_5} and B=(1, 0} where a_1 , a_2, a_3, a_4 and a_5 are the members of the family A.

If the members a_1, a_3 and a_5 are males, then the mapping of the sets will be as:



Thus, measurement may be defined as the map-ping of objects of one set to objects of other set.

The kind of measurement achieved is a function of the rule which assigns the numerals (symbols) to objects/observations. The rule is called a **scale**.

Indicants: Actually measurement is not of the objects, but the properties or characteristics of these objects. Strictly speaking, this is also not true. Actually measurements are done on the indicants of the properties of objects. Indicant is merely a convenient word used to mean something that points to something else. If a boy continually strikes other boys, his behavior is an **indicant** of his underlying hostility. The indicants from which properties of the objects are inferred are specified by **operational definitions** that specify the activities or operations necessary to measures variables (or constructs).

Constructs:

A construct is an invented name for a property such as **Authoritarianism, Achievement, Intelligence, Persistence,** etc. An operational definition is necessary in order to measure a property or a construct. This is done by specifying the activities or operations necessary to measure it.

Steps in Measurement Procedure:

The **first** step in any measurement procedure is to define the objects of the universe of discourse. **Second,** the properties of the objects must be defined. Then the universe `U' is partitioned into at least two mutually exclusive and exhaustive subsets, That is, each object must be assigned to one subset only and all objects must be so assigned.

After the objects of the universe have been classified into designated subsets, the member of the sets can be counted. When set members are counted in this fashion, all objects of a sub-set are considered to be equal to each other and unequal to the members of other sub-sets.

Example: Let U= All tenth grade pupils in a certain high school. Let the measurement characteristics be sex of the pupil. Then U is partitioned into two mutually sub-sets B and G where B is the subset of U which includes all males of U and G is the subset of U which includes all females of U so that U= B+ G

Postulates: There are three basic postulates of measurements. These postulates concern about the relationships between the objects being measured.

The postulates are:

(1) Either $x = y$ or $x \neq y$ but not both

(2) If $x=y$ and $y = z$, then $x= z$

(3) if $x > y$ and $y > z$, then $x>z$

It should be understood that when we say $x = y$, it is not necessary that x and y are same. It is meant to indicate that they are sufficiently the same to be classified as member of the same set if viewed with respect to some criterion.

In physical measurements, the postulate (3) is always true. But in case of behavior, it may not be so. For instance, if 'a' is longer than 'b' and 'b' is longer than 'c' then 'a' is definitely longer than 'c'. But suppose 'a' (Wife) dominates 'b' (Husband) and 'b' (Husband.) dominates 'c' (Child), it is not necessarily true that 'a' dominates 'c' instead 'c' may dominate 'a' (Mother). Thus the postulates should not be assumed to be true. It should be proved to be true.

Physical Scales: There are four types of physical scales. They are:

Nominal Scale

A nominal scale is one that allows the researcher to assign subjects to certain categories or groups. This is simplest and lowest form of data and it gives very basic information. This sale is usually used to obtain personal data, where grouping of individuals or objects is required.

EXAMPLE	Gender:	Male, Female
	Religion:	Hindu, Buddhist, Muslim, Christian
	Occupation:	Teacher, Manager, Doctor, Businessman, Civil servant
	Nationality:	Nepali, Indian, American, Japanese
	Department:	Sales, Finance, Personnel, Production

All categories are mutually exclusive. Every respondent has to fit into one of these categories. Therefore, one cannot rank these and say that a male is a higher value than a female, or that 3 teacher has a high value than a manager. Nominal data results from qualitative variables.

The information that can be generated from nominal scaling is to calculate the percentages and frequencies. For example, if you interviewed 150 students in a campus, and assigned a code number 1 to all male students and number 2 to all female students, then computer analysis of data might reveal that 100 were male and 50 were female students. This frequency distribution tells you that 66.6 percent of the respondents are male and 33.4 percent female. Nominal scale, thus, tells you nothing more than basic or gross information. Chi-square test is the most common test applicable to nominal data. Due to its limitations concerning statistical treatment, a nominal scale has the characteristic of exploratory research where the emphasis is on uncovering relationships rather than on specifying the form of relationship.

Ordinal Scale

A scale is ordinal when objects can be assigned order on some characteristic but they cannot be assigned values that represent degree of difference on that characteristic. An ordinal scale of

measurement, in addition to the function of classification, also allows cases to be ordered by degree according to measurements of the variable. This scale is usually used to rate the preferences of the respondents. It applies to data which can be ranked according to value but cannot be given a particular numerical value which actually is descriptive of the data. For example, one can rank drinks (coca-cola, tea, coffee, soda water, mineral water) in order of his or her preference from most preferred to least preferred. This means one likes coca-cola better than soda water or prefers coffee to mineral water.

EXAMPLE Rank the following occupations in terms of their social status.

Jobs.	**Rank of Social Status**
Manager	.....
Doctor	.....
Engineer	.....
Professor	.....
Lawyer	.....
Civil Servant	.....
Police Officer	.....

In this example, if the respondents give higher rank to doctors than engineers, we can say doctors have relatively higher social status than engineers. We cannot, however, conclude that the doctors have three times higher social status than the engineers.

EXAMPLE Rank the following cities in terms of their suitability to open a branch office of a commercial bank.

Cities	**Rank**
Biratnagar	.....
Hetauda	.....
Janakpur	.....
Butawal .	.....
Nepalgunj	.....
Dhangadi	.....

Ordinal scales represent numbers, letters, or any symbols used to rank items. The significant amount of business research relies on ordinal measures. The most common usage of ordinal scale is in obtaining preference measurements. For example, the employees of an organization maybe asked to rank their preferences for the newspapers they would like to read, or the games they would like to play. Ordinal scale thus provides more information than the nominal scale.

In ordinal data, median is an appropriate measure of central tendency. Percentile and quartile analysis are used for measuring dispersion. In most cases, rank-order correlations can be used Because of the nature of data; only non-parametric tests can be used.

Interval Scale

This scale assumes that the data have equal intervals. For example, there are five sisters who are all one year apart in age. Radha, Rambha, Reshma, and Rohini are 4, 3, 2 and 1 year older than the youngest sister Rabina. The important point is that we cannot say that the oldest sister Radha is twice as old as Reshma and four times older than Rohini. The reason is that we do not know what the age of the youngest sister, Rabina, is.

Interval scale is like ordinal but with constant intervals. The numbers tell both position and distance. Thus, the interval scale not only groups individuals according to certain categories and taps the order of these groups; it also measures the magnitude of the differences in the performances among the individuals. As such, it is more powerful scale than the nominal and ordinal scales.

	Strongly agree	Agree	Uncertain	Disagree	Strongly disagree
	1	*2*	*3*	*4*	*5*
I know exactly what tasks I am expected to do at all times.	1	2	3	4	5
Members of my team are fully cooperative.	1	2	3	4	5
My salary is regularly paid on time.	1	2	3	4	5
I am adequately trained for what I am doing.	1	2	3	4	5
Our norms of performance are quite realistic.	1	2	3	4	5
Our supervisor is very sympathetic.	1	2	3	4	5
I have the opportunity to work in a team environment.	1	2	3	4	5

Interval scales, or any other attempts at creating such scales, are found often in behavioral research. This is particularly true for measurement of attitudes and certain psychological characteristics, such as intelligence and learning. In analyzing interval data, many options are available. Mean can be appropriately used to measure central tendency. Standard deviation is

widely used for dispersion. Product moment correlation can be calculated, and t-test and F•test can be used for significance testing.

Ratio Scale

The ratio scale is the most powerful of the four scales because it has an absolute zero origin and subsumes all the properties of the other three scales. This allows the researcher to calculate the ratio of difference between the age of the individuals. For instance, one can say that the boy who is 8 years old is twice as old as the boy who is 4 years old. Also, we can say that a man who worked 40 hours, worked twice as many hours as the man who worked 20 hours. Some examples of ratio scales are actual age, income, the number of organizations an individual has worked for. The responses could range from 0 to any figure.

EXAMPLE

How many children do you have? .....

What is your annual household income? .....

How many workers are working in your factory? .....

How many workers are the members of the union? .....

What is the highest level of education you have completed? .....

How long have you lived in Kathmandu? .....

Ratio scales are found more commonly in the physical sciences than in the social sciences. Measurement of weight, length, time intervals, area, velocity, etc., all conform to ratio scales. In social sciences, we do find properties of concern that can be ratio scaled: money, age, years of education, etc. However, successful ratio scaling of behavioral attributes is rare. Various types of statistical analysis and mathematical operations can be done on ratio data. Geometric and harmonic means can be used. Likewise, coefficient of variation can be worked out.

Needs of Scaling

Scales are such which can be used in quantifying every measurable property of objects or the variables. To measure the physical phenomena different kinds of scales are to be needed. The scaling techniques are applied in ordering a series of items along some sort of continuum. The method of scaling is the technique of turning a series of qualitative facts into a quantitative series. The main needs of scaling techniques are:

(1) **For attaining scientific maturity**: The fundamental form of the movement in the direction of the greater precision is measuring graduation of the objects for a scientific study. Hence, the scaling technique is for the scientific maturity.

(2) **For objective measurement**: The scaling techniques are very useful in technical and social studies and the reliable inference about the technical/social phenomenon can be drawn by the use of the objective measurement.

(3) **For the improvement of more precise measuring device**: The existing measuring instrument and techniques can be improved by the use of different scales. More precise measuring devices are developed through the scientific use and practice of the existing scales.

Characteristics of a Scale

The essential characteristics of a good scale are:

a**) Continuum**

b) Reliability

c) Validity

d) Practicability

Continuum

It is the characteristics of scale that it should be in the form of continuous series and the factors to be measured interrelated. The continuum to be defined depends upon the nature of the phenomenon and the nature of the factors to be defined.

Reliability

To obtain consistent result the measuring instrument should be steadfast. The reliability of the measuring scale must be stable, consistent and the error function should be correctly defined. It means the stability, consistency, dependability, predictability and accuracy of the scale used. According to Ebil 'The term reliability means the consistency with which set of test scores measure whatever they do measure'. Reliability is necessary for validity but not sufficient. The

reliability of a test suffers to the consistency of score obtained of the some item/individual on different occasions or with different sets of equivalent.

Note:

a) If we get the same measuring result for the same object / observation in many times then the measure (scale) may be considered as reliable measure (scale).

b) The absence of measurement error in measuring instrument is the reliability of the measure.

c) The true measure of the property measure is called the reliable measure.

Validity

The most critical criterion of the scaling technique is the validity. A scale is said to be valid when it measures correctly what is expected to be measure. In other words, validity is the extent (degree) to which differences found with a measuring instrument reflect true difference among those being tested. It is thought as utility of the measure.

In experimentation, **reliability** is the extent to which the measurements resulting from a test reflect characteristics of the subject of measurement. An experiment is reliable if we are getting consistent results from the same measure. It is unreliable if repeat measurements of the same thing give different results.

In statistics a valid measure is one which is measuring what it is supposed to measure. **Validity** implies reliability (accuracy). A valid measure must be reliable, but a reliable measure may not be valid. Validity refers to getting results that accurately reflect the concept being measured.

Practicability

It is the characteristics of the measure that must be practicable to use. It should be reasonable economically, conveniently and interpretably. It should be easy to administer. The scales used must be supplemented by i) detailed instruction of handling ii) scoring techniques (keys) iii) guides for utilize and iv) evidence about reliability.

Reliability

Definitions of Reliability:

A measuring scale is said to be reliable if it provides consistent result. Prof. Goode and Hatt say 'a scale is reliable when it will consistently produce the same results when applied to the same sample'. 'Just as a ruler which shrank or expand materially when exposed to temperature, changes would be useless, so would it be a scale which yielded a different result upon each application?

To test the reliability of a measure (test) statistical method of analyzing (i) the scores obtaining in a single measure (test) or (ii) the correlation of the scores obtained in the more than one measure (tests) are used.

Methods of Estimating Reliability (Tests of Reliability)

To obtain the reliability of the different scores, the following four method of computing reliability, coefficient are used.

a) Test- retests Method

b) Parallel/Alternate/Equivalent forms Method

c) Split Half Method

d) Rational Equivalence Method (Kuder Richardson Method)

(a) **Test-retest Method**

In this method the same set of objects/ items is measured (tested) again and again by using the same or the comparable measuring instrument. The results so obtain are compared by computing correlation coefficient between the scores of the different tests (measures). If it is impossible to use such method due to the long space of time, it is considered whether the effects of causative factor in the period of two tests are present or not. To find the result related to this problem technique of control group (a team checking by trained and motivated persons) is applied.

(b) **Parallel Forms Method**

In this method two analogous forms of test-scales are constructed and alternatively applied to the same samples/items selected/objects selected. The analogous forms mean the alternate or parallel forms of the tests having of the same kind difficulty category and design. To obtain the reliability coefficient the correlation coefficient between the test-results is computed. If the results are in high degree of connection then the scale may be said to be reliable. The correlation of the forms is called self correlation and it becomes an index of equivalence of two sets. The parallel forms are mostly useful for the standard psychological and educational achievement tests.

(c) Split half Method

In this method the scores are divided in two equal parts randomly. For example, the scores on odd no. of test-set as a first half and the scores on even no. of test-set as the other half. Considering a part as complete, scaling procedure is applied for these two half-parts. If there exists a high degree of correlation between the scores of these halves, then the scaling is considered to be reliable. The split half method is generally used when it is difficult to construct the parallel form test. The coefficient of internal consistency or the self correlation coefficient of

the whole test is called Stepped up reliability and estimated by using Spearman-Brown formula. The general formula is

$$Rw = \frac{n * r_p}{1+(n-1)r_p}$$

Where, R_w is stepped up reliability coefficient; n the number of parts

r_p = the correlation coefficient between two parts.

Spearman- Brown formula for computing test reliability having two halves is

$R_{xy} = \frac{2\times r_{xy}}{1+r_{xy}}$

Where,

R_{xy} = Stepped up reliability coefficient of first and second half

2 is the number of parts and

r_{xy} = The correlation coefficient between two parts X and Y.

If the score are expressed in the ranks then, correlation coefficient is calculated by

$r_{xy} = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$, Which is called Spearman's Rank correlation Coefficient?

If the scores are in the numeric scale then, correlation coefficient is calculated by,

$$r_{xy} = \frac{n\Sigma XY - \Sigma X \times \Sigma Y}{\sqrt{nY^2-(\Sigma Y)^2}\sqrt{n\Sigma X^2 -(\Sigma X)^2}}$$

Which is called Karl Pearson's coefficient of correlation, where, X represent the scores of the first half and Y the scores on the second set.

Example

A test-score is divided in two halves as the scores on the odd numbered questions and the scores on the even numbered questions. The correlation coefficient between them is obtained as 0.72, what is the reliability coefficient of the whole test.

Solution

Here, Correlation coefficient (r_{xy}) = 0.72

n = 2

Reliability coefficient, $R_{xy} = \frac{2\times r_{xy}}{1+ r_{xy}} = \frac{2\times 0.72}{1+0.72} = 0.8272$

The reliability coefficient is 82.72%; the dependability of the whole score seems to be very good.

d) Rational Equivalence Method (Kuder-Richardson Method)

Two forms of a test are defined as equivalence when corresponding items are interchangeable and inter item correlation is same for both forms. Kuder-Richardson method is the method of

obtaining reliability by using the internal consistency between the measures (questions) of the same scaling (test). The reliability coefficient for this method is obtained by the following two formulae. Kuder-Richardson's first formula for reliability is denoted by KR_1, is computed as

$$KR_1 = R_w = \frac{n}{n-1}\left[1 - \frac{\Sigma pq}{\sigma^2}\right]$$

Where, R_w = The reliability coefficient of the whole test

n - The number of items in the test

σ -Standard deviation of the test score

p-the proportion of the answering test item correctly

q = (1 - p)

Note 1: If the values of p for each test is equal then Σpq = npq

2: If the values of p for 'n' tests are $p_1, p_2 \ldots\ldots p_n$, then $\Sigma pq = p_1q_1 + p_2q_2 + \ldots\ldots\ldots\ldots + p_nq_n$

Example

In a test there are 60 questions. The proportion of answering each question correctly is 70%, if the standard deviation is 10 what is the reliability coefficient?

Solution

Number of questions (n) = 60,

Proportion of correct answers (p) = 0.7 q = 1 - p = 1 - 0.7 = 0.3

Σ pq = npq

S.D. (s) = 10, therefore, the reliability coefficient is given by

$$KR_1 = R_w = \frac{n}{n-1}\left[1 - \frac{\Sigma pq}{\sigma^2}\right]$$

$$KR_1 = \frac{60}{60-1}\left[1 - \frac{60\times0.7\times0.3}{\sigma^2}\right]$$

Hence reliability coefficient is 88.88%

Validity

Introduction

A scale possesses validity when it actually measures what it claims to measure. In other words, a scale is said to be valid if it measures what is expected to measure.

Interpretation of test scores ultimately involves predictions about a subject's behavior in a specified situation. If a test is an accurate predictor, it is said to have good validity. Before validity can be demonstrated, a test must first yield consistent, reliable measurements. In addition to reliability, psychologists recognize three main types of validity

Types of Validity

There are mainly the following types of validity:

1. **Content Validity**

Content validity is the representativeness or adequacy of the unit selected of the content such as the substances, the matter, the topic of the measuring instrument etc. Content Validity is also known as logical validity. A test has content validity if the sample of items in the test is representative of all the relevant items that might have been used. Words included in a spelling test, for example, should cover a wide range of difficulty.

It is related to the objective of the study, it tests the objective of the course of action.

Let, U be the universe of the item; S is the subset of U i.e., $S \subset U$ and x be the item such that x is the element in U then it must be an element of S, i.e., $x \in U \Rightarrow x \in S$

The standardized achievement test is used for the content validity measure.

In Psychometrics, Content Validity refers to the extent to which a measure represents all facets of a given social concept. For example, a depression scale may lack content validity if it only assesses the affective dimension of depression but fails to take into account the behavioral dimension.

Content validity is related to face validity, although content validity requires more rigorous statistical tests than face validity, which only requires an intuitive judgment. &intent validity is most often addressed in academic and vocational testing, where test items need to reflect the knowledge actually required for a given topic area or job skill. In clinical settings, content validity refers to the correspondence between test items and the symptom content of a syndrome.

2. **Criterion-related Validity**

Criterion-related validity refers to a test's accuracy in specifying a future or concurrent outcome. A common approach, called criterion related validity is to correlate measures with a criterion measure known to be valid. for example, an art-aptitude test has predictive validity if high scores are achieved by those who later do well in art school. The concurrent validity of a new intelligence test may be demonstrated if its scores correlate closely with those of an already well-established test.

The criterion related validity is based on the four decisive factors: (i) an external criterion (ii) regular and future behavior (iii) logical analysis and (iv)empirical method.

The criterion-related validity is related to the ability to predict some outcomes or estimate the existence of some current conditions.

• It is used to predict the criterion on the basis of some measure.

• It reflects the success of measures for empirical estimating process

In this type of validity, the proposed criterion must possess the following quality:

- Freedom from bias (criterion should give each subject/matter an equal opportunity to score)

- Reliable (criterion should be stable and reproducible)

- Relevance (criterion should be defined in terms of proper measure). Availability (information specified by the criterion must be available)

The Criterion- related Validity broadly classified as: (a) Predictive validity and (b) Concurrent Validity

(a) Predictive validity:

It refers to the usefulness of the test in prediction some future performances on the criterion. It is concerned with how well the scale can forecast a future criterion. When the criterion measure is collected later the goal is to establish it is called the predictive validity.

(b) Concurrent Validity:

It refers to the usefulness of a test in closely relating to the other measure of known validity. It is concerned with the performances that how it can describe a present criterion. When the criterion measure is collected at the same time as the measure being validated the goal is to establish concurrent validity. Concurrent Validity is demonstrated where a test correlates well with a measure that has previously been validated. For example, if a test measuring job satisfaction gives similar results to those gathered using a job satisfaction which has been validated in past investigations the new measurement has concurrent validity.

The validity coefficient is measured in terms of the correlation coefficient (r) of the scores of the different tests. We say, for $0.9 \leq$ r ≤ 1 there is very high validity; for $0.8 \leq r \leq$ 0.9 a high validity; for $0.6 \leq r \leq$ 0.8 a satisfactory validity: for $0.4 \leq r \leq 0.6$ a moderate validity; for $0.0 \leq r \leq 0.4$ a poor validity; and for $r < 0.0$ a negative validity

3. Construct Validity In social science, construct validity refers to whether a scale measures the unobservable social construct that it claims to measure. It is the validity, most complex and abstract because of the complexity of the social parameter. A scale is said to posses construct

validity to the degree that it confirms to predicted association with other theoretical postulates. The essence of construct validity is its dependence on theory and the examination of observed association is a test of theory as valid scale. It is based on psychological trait and quality Construct validity is generally determined by investigating what psychological traits or qualities a test measures; that is, by demonstrating that certain patterns of human behavior account to some degree for performance on the test. A test measuring the trait "need for achievement," for instance, might be shown to predict that high scorers work more independently, persist longer on problem-solving tasks and do better in competitive situations than low scorers.

A construct is not restricted to one set of observable indicators or attributes. It is common to a number of sets of indicators. Thus, construct validity can be evaluated by statistical methods that show whether or not a common factor can be shown to exist underlying several measurements using different observable indicators. For determining construct validity, we associate a set of other proposition with the result received from the use of our measuring instrument. If measurements on our devised scale correlated (associated) in a predicted way with the other propositions, we can conclude that there is construct validity.

Scaling

Scores and Scales

Scores

The number of points somebody gets for correct answers in a test is said to be scores. In other words the value of parameter in the observed phenomenon is termed as score. From an experiment or from a test what we obtained as the observation is called the raw score. The raw score is the simple numerical count of responses such as the number of correct answers on an intelligence test. The usefulness of the raw score is limited however, because it does not convey how well someone does in comparison with others taking the same test. Suppose, one attempts to answer"20 IQs having 1 point each and answered 16 correctly then the raw score is 16. One tries to measure the length of 5 pencils and found to be the lengths of 10, 11, 10, 14, 12 in cm; then the raw scores of the measuring phenomenon are 10, 11, 10, 14, and 12 cm.

Scale

Scale is a predefined sequence of scores in ascending values that can map an item to it. Scale is a set of all the different levels of symbols or numerals or something so constructed, from the lowest to highest, that these can be assigned by rule to objects or to items or to the individuals or to their behavior to whom it is applied. In general concept, scale is also known as a quantifying appliance used to indicate the systematized numerals of the measuring instrument.

Scaling of the Scores

From a set of scores of the test we can construct a sequence of levels of the values that can be used as an extent for the test of that phenomena, this method of leveling is said to be scaling. For the purpose of scaling it is always desired to make the scores in an array. After that the scores are converted to the percentile points and then to a scale of required form.

The raw scores obtained in test can be converted to different auxiliary scores in relation to the distribution of the raw scores or according to the distribution of parent population. Such scores which are modified/improved/developed from the raw scores are called derived scores. From the derived scores of the same form in sequence can be used to create a continuous structure of the numerals which is the scale required to be constructed.

There are different types of derived scores widely used in the measurement of the phenomenon or the attitudes. The percentile scores, s-score (z-score) and T-score are such derived scores and are use to compare the strength and credibility of the measures. From these scores we can construct the standard scales namely percentile scale, sigma scale (z-scale) and T scale, respectively by arranging the scores lowest to highest.

Difficulties in scaling

In social phenomenon, following are the reasons that create difficulties of scaling in social sciences.

- Social complexity (intricacy) [social phenomena are complex and such complexity cannot be measured]
- Abstractness (nonfigurative) of the social phenomena
- Heterogeneity of the social values, customs and norms.
- Changing nature of human behavior Absence of universal measuring of social values
- Laboratory method cannot be applied in social phenomena.

Scales used in Social and Physical sciences

The main scales used in the measure of social/physical characteristics are:

(i) Point scale
(ii) (Social distance scale
(iii) Rating scale
(iv) Ranking scale
(v) Thurstone scale

(i) Point scale

In this type of scale words or situations representing the criteria are selected and one point (marks or number) is given for each criteria. Attitude of a person can determined by the use of all the three following method effectively.

Method 1: The respondent is asked to tick one that is representing or favorable to him /her. The scores are counted and result is derived**.**

Method 2: In the second method the respondent is asked to cross the one point or situation which is not favorable to him /her. A point is given to each and every word that has not been crossed. The attitude of the respondent is then determined by counting no of points.

Method 3: In the third method of point scale the respondent is asked to cross, on which points he /she is agree or not.

(ii) Social distance scale

The social distance scale is developed by Emary S. Borgadus to measure the social distances (it is commonly known as Borgadus scale). The social distance may be defined as the proximity and favoritism; for example the cultural distance from one race to the other, custom from one ethnic group to the other etc. To measure a person's (respondent's) attitude how far from the given cause situation the Borgadus scale can be used. Borgadus developed a scale to measure the nearness of liking between two social groups using several items or statements which show the varying relationship of social distance of Americans with other races as English, Korean, Swedish and poles.

(iii) ***Rating scale***

When the character to be measured is not dichotomous in this case the rating scale is used. Rating scale consists of a set of figures that can match to the individual or items to be measured. The response or the opinions of the respondent's attitude is rated in three to six points in continuum (range). The intensity of the attitude is measured by using equal or unequal type intervals. An example of three point rating scale is:

Very good .....satisfactory.....poor

A five points rating scale is

Strongly-Agree.....Agree .....Not-Decided. .....Dis-agree.....Strongly-Disagree

(iv) Ranking scale

The ranking scale is similar to rating scale applied to a set of objects or individuals with the preference or liking. In this scale the situations are placed in such a way that, everybody who inspects it knows that one likes the one better than the other. Ranking scales is determined in comparison to a few cases known as stimuli. The item

obtaining first preference scores 1, the second as 2, third as 3 and so on. 'The smaller the score the greater the preferences' is the principle of ranking scale.

(v) Thurstone scale

American psychologist Louis L. Thurstone proposed that intelligence was not one general factor but a small set of independent factors of equal importance. He called these factors primary mental abilities. To identify these abilities, he developed a plan to conduct study amongst 250 college students, identified factors and developed a scale of measuring aptitude using factor analysis. The scale so developed is known as Thurstone scale. In educational and psychological experiments, it is used as a main type of scale used to measure the attitude. The statements are collected and arranged in continuum from most favorable to least favorable with neutral point (zero). It is one type of point scale having neutrality point at the central location.

Most- Favored.....neutral .....Least— Favored

Introduction

Sampling is an essential part of any research investigation. Almost all research studies involve sampling. It is, therefore, essential that we understand the main concepts of sampling and are familiar with the sampling methods.

Research studies assume that the people selected for studies are representative of a large group about whom generalizations are to be made. We normally cannot survey everyone in the population; but through sampling techniques, we can be confident that only a small part of the total population can fairly represent the total population. Sampling, then, is a technique that saves the time and trouble of questioning 100 percent of the population.

What is Sample?

A sample is a collection of items or elements from a population or universe. Hence, a sample is only a portion or subset of the universe or population. It comprises some observations selected from the population. For instance, if 50 students are drawn from a population of 500 students of a college, these 50 students form the sample for the study.

Population or universe refers to the entire group of people, events, or things of interest that the researcher wishes to investigate. For example, if you are interested in investigating the smoking habits of employees in a chemical factory, then all employees in that factory will form the population.

Sometimes, the entire population will be sufficiently small, and you can include the entire population in your study. If the total items are studied, that is called a census study. However, it is not always possible to study every items or elements in a universe. Usually, the population is too large. Hence, a small, but carefully chosen sample can be used to represent the population. The sample thus selected reflects the characteristics of the population from which it is drawn. For thesis or project work to be undertaken by you, the study of the total population is neither possible nor necessary. Making a census study of the entire universes is not possible on account of limitations of time and money. Hence, sampling becomes inevitable.

Population may be finite or infinite. A finite population is one containing a fixed number of elements. The number of books in a library is an example of finite population. An infinite population is one without limits of any kind and is therefore indeterminate. The production of cigarettes by Surya tobacco Cigarette Factory is an example of infinite population.

Sample Design and Related Terminologies

Sampling design or strategy is the way in which you design your sample plan and select your samples from the population. In designing a sample, you must consider three things: sampling frame, selection of sampling items, and sample size. These terminologies can be explained as follows.

■ Sampling *frame* is the list identifying each unit in the study population. All the elements in a sampling population constitute its sampling frame. Thus, it may be all the students at Trichandra Multiple Campus, all names in the telephone directory, or all persons having their bank accounts with the Nepal Bangladesh Bank. After determining the sampling frame, the researcher will decide how sample will be selected.

■ *Sampling item or unit* is an element (person, institution, etc) of your study that becomes

■ The basis for selecting your sample.

■ The *size of the sample* must be determined. What should be the sample size? Should fifty or eighty employees be interviewed? Though accuracy is greater with large samples, so are costs.

■ *Sample statistics* are the information obtained from the respondents selected for your study. Your sample statistics become the basis of estimating the prevalence of the characteristics in the study population.

■ *Population parameters* or population mean are the characteristics of the population estimated from the sample statistics. If you measure the entire population and calculate a value like a mean or average, this is called a population parameter.

Thus, the basic components of a sample design are: (a) choosing the sample units (who are to be surveyed), (b) choosing the sample size (how many to be surveyed), (c) choosing the sampling procedure (how to ensure that those who are to be interviewed are included in the sample), and (d) choosing the media (how to reach respondents in the sample? - through mail survey, personal interview, or telephone interview).

THE PRINCIPAL STEPS IN A SAMPLE SURVEY

The main steps involved in the planning and execution of a sample survey may be grouped somewhat arbitrarily under the following heads.

1. **Objectives of the Survey**.
   The first step is to define in clear and concrete terms, the objectives of the survey. It is generally found that even the sponsoring agency is not quite clear in mind as to what it wants and how it is going to use the results. The sponsors of the survey should take care that these objectives are commensurate with the available resources in terms of money, manpower and the time limit required for the availability of the results of the survey.
2. **Defining the Population to be Sampled**. The population, i.e., the aggregate of objects (animate or in-animate) from which sample is chosen should be defined in clear and unambiguous terms. For example, in sampling of farms clear-cut rules must be framed to define a farm regarding shape, size., etc., keeping in mind the border-line cases so as to enable the investigator to decide in the field without much hesitation whether or not to include a given farm in the population.

But practical difficulties in handling certain segments of the population may point to their elimination from the scope of the survey. Consequently, for reasons of practicability or convenience the population to be sampled (the sampled population) is different, in fact more restricted, than the population for which results are wanted (the target population).

3. **The Frame and Sampling Units**.

The population must be capable of division into what are called sampling units for purposes of sample selection. The sampling units must cover the entire population and they must be distinct, unambiguous and non-overlapping in the sense that every element of the population belongs to one and only one sampling unit. For example, in socio-economic survey for selecting people in a town, the sampling unit might be an individual person, a family, a household or a block in a locality. In order to cover the population decided upon, there should be some list, map or other acceptable material, called the frame, which serves as a guide to the population to be covered. The construction of the frame is often one of the major practical problems since it is the frame which determines the structure of the sample survey. The lists which have been routinely collected for some purpose, are usually found to be incomplete or partly illegible or often contain an unknown amount of duplication. Such lists should be carefully scrutinised and examined to ensure that they are free from these defects and are up-to-date. If they are not up-to-date, they should be brought up-to-date before using them. A good frame is hard to some by and only good experience helps to construct a good frame.

4. Data to be collected.

The data should be collected keeping in view the objectives of the survey. The tendency should not be to collect too many data some of which are never subsequently examined and analysed. A practical method is to chalk out an outline of the tables that the survey should produce. This would help in eliminating the collection of irrelavent information and ensure that no essential data are omitted.

5. The Questionnaire or Schedule. Having decided about the type of the data to be collected, the next important part of the sample survey is the construction of the questionnaire (to be filled in by the respondent) or schedule of enquiry (to be completed by the interviewer) which requires skill, special technique as well as familiarity with the subject-matter under study. The questions should be clear, brief, corroborative, non-offending, courteous in tone, unambiguous and to the point so that not much scope of guessing is left on the part of the respondent or interviewer. Suitable and detailed instructions for filling up the questionnaire or schedule should also be prepared.

6. Method of Collecting information

The two methods commonly employed for collecting data for human populations are : (i) Interview Method. In this method, the investigator goes from house to house and interrviews the individuals personally. He asks the questions one by one and fills up the schedule on the basis of the information supplied by the individuals. (ii) Mailed Questionnaire Method. In this method,

the questionnaire is mailed to the individuals who are required to fill it up and returns it duly completed. Whether the data should be collected by interview method or mail questionnaire method or by physical observation has to be decided keeping in view the costs involved and the accuracy aimed at. Although mail surveys are less costly, there is scope for considerable non-response. Moreover mail method is practicable only among the educated people who are really interested in the particular survey being conducted. On the other hand, interview method costs more and there are interviewer errors also but without investigators the data collected may be worthless. In cases where data are to be collected by observations, the method of measurement, the type of measuring equipment or instrument, etc., are to be decided.

7. Non-respondents.

Quite often (due to practical difficulties), the data cannot be collected for all the sampled units. For example, the selected respondent may not be available at his place when the investigator goes there or he may fail or even refuse to give certain information when contacted. This incompleteness, called non-response, obviously tends to change the results. Such cases of non-response should be handled with caution in order to draw unbiased and valid conclusions. Procedures will have to be devised to deal with those who do not furnish information. The reasons for non-response should be recorded by the investigator.

8. Selection of Proper Sampling Design.

The size of the sample (n), the procedure of selection and the estimation of the population parameters along with their margins of uncertainty are some of the important statistical problems that should receive the most careful attention. A number of designs (plans) for the selection of a sample are available and a judicious selection will guarantee good and reliable estimates. For each sampling plan, rough estimates of sample size n can be obtained for a desired degree of precision. The relative costs and time involved should also be considered before making a final selection of the sampling plan.

9. Organization of Field Work. It is absolutely essential that the personnel should be thoroughly trained in locating the sample units, recording the measurements, the methods of collection of required data before starting the field work. The success of a survey to a great extent depends upon the reliable field work. It is very necessary to make provisions for adequate supervisory staff for inspection after field work. From practical point of view a small pretest, (i.e., trying out the questionnaire and field methods on a small scale) has been found to be immensely useful. It always helps to decide upon effective method of asking questions and results in the improvement of the questionnaire. Moreover, it might disclose certain problems and troubles that will otherwise be quite serious on a large-scale survey such as "the cost and the time may far exceed the availablee money and stipulated period.”

10 Summary and Analysis of the Data. The analysis of the data may be broadly classified into the following heads :

a)Scrutiny and editing of the data: An initial quality check should be carried out.by the supervisory staffs while the investigators are in the field.Accordingly, the schedule should be thoroughly scrutinized to examine the plausibility and consistency of the data obtained. The scrutiny or editing of the completed questionaire will help in amending recording errors or in eliminating data that are obviously erroneous and inconsistence.

about the (b) Tabulation of data: Before carrying out the tabulation of the data, we must decide about the procedure for tabulation of the data which are incomplete due to non-response to certain items in the questionnaire and where certain questions are deleted in editing process. The method of tabulation, viz., hand tabulation or machine tabulation, will depend upon the quantity of the data. For large-scale survey, machine tabulation will obviously be much quicker and economical. For a large-scale sample survey, the use of code numbers for qualitative variables is essential for machine tabulation. With simple questionnaire, the answers can sometimes be pre-coded, i.e., entered in a manner in which they can be conveniently or routinely transferred to mechanical equipment such as personal computers, etc. Finally, the tables that lead to the estimates are prepared.

(c) Statistical analysis. After the data has been properly scrutinized, edited and tabulated, a very careful statistical analysis is to be made. Different methods of estimation may be available for the same data. Appropriate formulae should then be used to provide final estimates of the required information. Efforts should be made to keep the procedure free from errors.

(d) Reporting and conclusions. Finally, a report incorporating detailed statement of the different stages of the survey should be prepared. In the presentation of the results, it is good practice to report the technical aspect of the design, viz., the types of the estimators used along with the amount of error to be expected in the most important estimate.

11. Information gained for Future Surveys. Any completed survey is helpful in providing a note of caution and taking lessons from it for designing future surveys. The information gained from any completed sample in the form of the data regarding the means, standard deviations and the nature of the variability of the principal measurements tougher with the cost involved in obtaining the data serves as a potential guide for improved together sampling. Moreover, in any complex survey, the things usually do not go exactly as planned. Any completed sample may serve as a lesson to the organisers for future surveys recognising and rectifying the mistakes committed in the execution of the survey.

PARAMETER AND STATISTIC

In order to avoid verbal confusion with the statistical constants of the population, viz., mean, variance, etc., of the population which are usually referred to as parameters, statistical measures computed from the sample observations alone, e.g., mean, variance, etc., of the sample have been termed as statistic.

In practice parameter values are not known and their estimates based on the sample values are generally used. Thus statistic which may be regarded as an estimate of the parameter, obtained from the sample, is a function of the sample values only. It may be pointed out that a statistic, as it is based on sample values and as there are multiple choices of the samples that can be drawn from a population, varies from sample to sample. The determination or the characterization of the variation (in the values of the statistic obtains from different samples) that may be attributed to chance or fluctuations of sampling is one of the fundamental problems of the sampling theory.

(Unbiased Estimate)

A statistic $t = t(x_1, x_2, ..., x_n)$, a function of the sample values $x_1, x_2.....,x_n$ is an unbiased estimate of population parameterθ if $E(t) = \theta$, i.e., if E(Statistics) = Parameter, then statistic is said to be an unbiased estimate of the parameter.

S**ampling Distribution**. The number of possible samples of size n that can be drawn from a finite population of size N is NCn. (If N is large or infinite, then we can draw a large number of such samples.) For each of these samples we can compute a statistic, say't' ....e.g., mean, variance, etc., which will obviously vary from sample to sample. The aggregate of the various values of the statistic under consideration so obtained (one from each sample), may be grouped into a frequency distribution which is known as the sampling distribution of the statistic. Thus, we can have the sampling distribution of the sample mean $\bar{x}$, the sample variance, etc.

Standard Error.

The standard deviation of the sampling distribution of a statistic is known as its Standard Error. The standard errors (S.E.) of some of the well-known statistics are given in Table, where n is the sample size, σ^2 the population variance, P the population proportion and Q = 1- P.

Utility of Standard Error.

S.E. plays a very important role in the large sample theory and forms the basis of the testing of hypothesis. If t is any statistic, then for large samples

$Z = \frac{t - E(t)}{\sqrt{V(t)}} \sim N(0,1)$

$\Rightarrow \quad Z = \frac{t - E(t)}{S.E.(t)} \sim N(0,1)$

: Thus, if the discrepancy between the observed and the expected (hypothetical) values of the statistic is greater then 1.96 times the S.E the hypothesis is, rejected at 5% level of significance. Similary,

if t - E(t) $\leq$ 1.96 x S.E. (t),

the deviation is not regarded significant at 5% level of significance. In other words the deviation' t - E(t), could have arisen due to fluctuations of sampling and the data do not provide us any evidence against the null hypothesis which may, therefore , be accepted at 5% level of Significance. Similarly we can discuss the significance of the difference at 1% level of significance

The .magnitude of the standard error gives an index of the precision of the estimate of the parameter. The reciprocal of the standard error is taken as the measure of reliability or precision of the sample.

STANDARD ERRORS OF STATISTIC

S. No.	Statistic	Standard Error
1.	$\bar{x}$	$\sigma / \sqrt{n}$
2.	Observed sample proportion 'p'	$\sqrt{PQ/n}$
3.	Sample standard deviation s	$\sqrt{\sigma^2/2n}$
4.	s^2	$\sigma^2\sqrt{2/n}$
5.	Quartiles	$1{\cdot}36263\ \sigma/\sqrt{n}$
6.	Median	$1{\cdot}25331\ \sigma/\sqrt{n}$
7.	'r' = sample correlation coefficient	$(1-\rho^2)/\sqrt{n}$
8.	μ_3	ρ, population correlation coeff. $\sigma^3\sqrt{96/n}$
9.	μ_4	$\sigma^4\sqrt{96/n}$
10.	Coefficient of variation (V)	$\frac{V}{\sqrt{2n}}\sqrt{\left(1+\frac{2V^2}{104}\right)} \cong \frac{V}{\sqrt{2n}}$

SAMPLING AND NON-SAMPLING ERRORS

The errors involved in the collection, processing and analysis of a data may be broadly classified under the following two heads:

(**i) Sampling Errors, and (ii) Non-sampling Errors**.

(i) Sampling Errors.

Sampling errors have their origin in sampling and arise due to the fact that only a part of the population (i.e., sample) has been used to estimate population parameters and draw inferences about the population. As such the sampling errors are absent in a complete enumeration survey. Sampling biases are primarily due to the following reasons:

1. Faulty selection of the sample. Some of the bias is introduced by the use of defective sampling technique for the selection of a sample, e.g., purposive or judgment sampling in which the investigator deliberately selects a representative sample to obtain certain results. This bias can select a representative sample to obtain certain results. This bias can be overcome by strictly adhering to a simple random sample or by selecting a sample at random subject to restrictions which while improving the accuracy are of such nature that they do not introduce bias in the results.

2. Substitution.

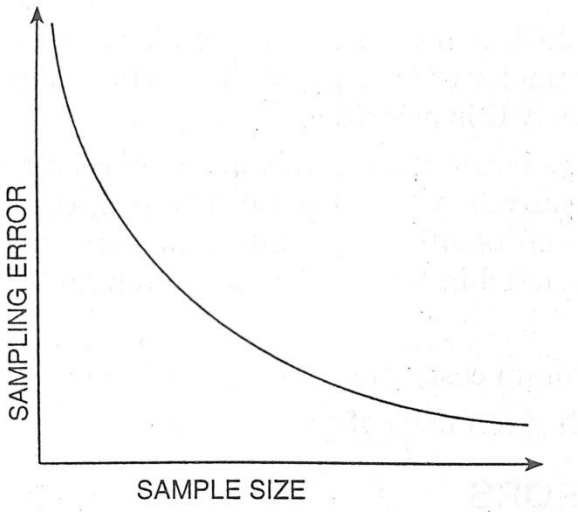
If difficulties arise in enumerating a particular sampling unit included in the random sample, the investigators usually substitute a convenient member of the population. This obviously leads to some bias since the characteristics processed by the substituted unit will usually be different from those possessed by the unit originally included in the sample.

3. Faulty demarcation of sampling units.

Bias due to defective demarcation of sampling units is particularly significant in area surveys such as agricultural experiments in the field or crop cutting survey, etc. In such surveys, while dealing with border line cases, it depends more or less on the discretion of the investigator whether to include them in the sample or not,

4. Constant error due to improper choice of the statistics for estimating the population Parameters.

For example, if $x_1, x_2, \ldots\ldots\ldots, x_n$ is a sample of independent observations, then the sample variance $s^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 / n$ as an estimate of the population variance σ^2 is biased whereas the statistic $\frac{1}{n}\sum_{i=1}^{n}(x - \bar{x})^2$, is an unbiased estimate of σ^2.



Remark: Increase in the sample size (i.e., the number of units in the sample) usually results in the decrease in sampling error. In fact, in many situations this decrease in sampling error is inversely proportional to the square root of the sample size as illustrated in Figure

. **(ii) Non-sampling Errors**. As distinct from sampling errors which are due to the inductive process of inferring about the population on the basis of a sample, the non-sampling errors primarily arise at the stages of observation, ascertainment and processing of the data and are thus present in both the complete enumeration survey and the sample survey. Thus, *the data obtained in a complete census, although free from sampling errors, would still be subject to non-sampling errors whereas data obtained in a sample survey should be subject to both sampling and non-sampling errors.*

Non-sampling errors can occur at every stage of the planning or execution of census or sample survey. The preparation of an exhaustive list of all the sources of non-sampling errors is a very difficult task. However, a careful examination of the major phases of a survey (complete or sample) indicates that some of the more important non-sampling errors arise from the following factors :

1. ***Faulty Planning or Definitions***
   The planning of a survey consists in explicitly stating the objectives of the survey. These objectives are then translated into (i) a set of definitions of the characteristics for which data are to be collected, and (ii) into a set of specifications for collecting, processing and publishing. Here the non-sampling errors can be due to:

   (a) Data specification being inadequate and inconsistent with respect to the objectives of the survey.

   (b) Error due to location of the units and actual measurement of the characteristics, errors in recording the measurements, errors due to ill-designed questionnaire, etc.

(c) Lack of trained and qualified investigators and lack of adequate supervisory staff.

2. Response Errors. These errors are introduced as a result of the responses furnished by the respondents and may be due to any of the following reasons:

(i) Response errors may be accidental. For example, the respondent may misunderstand a particular question and accordingly furnish improper information un-intentionally.

(ii) Prestige bias. An appeal to the pride or prestige of person interviewed may introduce yet another kind of bias, called prestige bias by virtue of which he may upgrade his education, intelligence quotient, occupation, income, etc., or downgrade his age, thus resulting in wrong answers.

(iii) Self-interest

. Quite often, in order to safeguard one's self-interest, one may give incorrect information, e.g., a person may give an underestimate of his salary or production and an over-statement of his expenses or requirements, etc.

(iv) Bias due to interviewer.

Sometimes the interviewer may affect the accuracy of the response by the way he asks questions or records them. The information obtained on suggestions from the interviewer is very likely to be influenced by interviewer's beliefs and prejudices.

(v) Failure of respondent's memory.

One source of error which is common to most of the methods of collecting information is that of 'recall'. Many of the questions in surveys refer to happenings or conditions in the past and there is a problem both of remembering the event and associating it with the correct time period.

3. Non-response Biases.

Non-response biases occur if full information is not obtained on all the sampling units. In house-to-house survey, non-response usually results if the respondent is not found at home even after repeated calls, or if he/she is unable to furnish the information on all the questions or if he/she refuses to answer certain questions. Therefore, some bias is introduced as a consequence of the exclusion of a section of the population with certain peculiar characteristics, due to non-response.

4. Errors in Coverage.

If the objectives of the survey are not precisely stated in clear cut terms, this may result in

(i) the inclusion in the survey of certain units which are not to be included, or

(ii) The exclusion of certain units which were to be included in the survey under the objectives. For example, in a census to determine the number of individuals in the age group, say, 20 years to 50 years, more or less serious errors may occur in deciding whom to enumerate unless particular community or area is not specified and also the time at which the age is to be specified.

5. Compiling Errors.

Various operations of data processing such as editing and coding of the responses, tabulation and summarizing the original observations made in the survey are a potential source of error. Compilation errors are subject to control though verification, consistency check, etc.

6. Publication Errors.

Publication errors, i.e., the errors committed during presentation and printings of tabulated results are basically due to two sources. The first refers to the mechanics of publication—the proofing error and the like. The other, which is of more serious nature, lies in the failure of the survey organization to point out the limitations of the statistics.

Sampling Frame

The sampling frame (also known as the "sample frame" or "survey frame") is indeed the actual collection of units. A sample has now been taken from this. A basic random sample gives all units in it an equal probability of being drawn and appearing in the sample. In the ideal scenario, the sample frame should match the sample of people.

A complete list or collection from which our sample participants will be drawn in a predetermined manner. The list will be organized in some way. That is, each member of a population will have an individual identity and a contact mechanism. This allows you to categorize and code known information about segmentation features.

Collecting the sample indicates that we have a supply or list of all the individuals of the target population from which to take a sample, as well as a process for selecting the sample. Any resource that has the information needed to reach every individual in the targeted group qualifies as a source.

One of the first steps in creating a research study is to define all the modules (also known as cases) we want to investigate. People, organizations, and existing records might all be considered units. The population of research interest is made up of various units. It is critical to be as detailed as possible when describing the population.

Characteristics of a Good Sampling Frame

Be assertive when selecting lists! Make sure the sample frame is large enough for our requirements. A decent sample frame for research on living conditions, for example, might include:

Everyone is in the target demographic.

Exclude everyone who isn't part of the target group.

A file containing factual information that may be used to reach specific people.

Other considerations:

Each member has a unique identification. This might be a short number code (e.g., from 1 to 3000).

Make sure the frame doesn't have any duplicates.

The list should be well organized. Sort them alphabetically for better access

Information should be up to date. This might need to be examined regularly (e.g., for address or contact number changes).

Examples of the Sampling Frame

The issue is that studying every individual in a population is not always practical or practicable.

Suppose we might be curious to learn about the opinions of American bankers about vehicle ownership, for example. Gathering data from every bank in the United States would be too time-consuming and expensive. You can investigate a sample of the population in situations like these.

The process of picking a sample should be intentional, and you can utilize various sampling strategies based on the research's aim.

It would help if we first constructed a sampling frame, which would be a list of all the units in the population of interest before we can choose a sample. Our study findings can only benefit the population identified by the sample frame.

Again, consider a survey to determine the number of prospective clients for digital programs in the New Jersey population. The research team selected 1,000 random numbers from a local telephone directory, made 200 calls daily from 9 a.m. to 6 p.m., and asked specific questions.

The sample frame comprises just those New Jersey residents who meet all the following criteria:

Owns a phone.

The number is listed in the directory.

Is present at home Monday through Friday from 9 a.m. to 6 p.m.

Is not a user who refuses to take part in any telephone surveys.

In this situation, the sample frame is distinct from the population. For example, it under-represents groups that do not have a telephone (e.g., the poorest), have an unlisted number, were not at home at the time of calls (e.g., employed individuals), or do not want to engage in phone surveys (e.g., more busy and active people). Such disparities between the sample frame and the target population are the most common limitations in surveying and other random sampling procedures.

Conclusion

A sampling frame is a researcher's list or device to specify the population of interest. A basic random sample gives all units in this an equal probability of being drawn and appearing in the sample.

People, organizations, and existing records might all be considered units. It is critical to be as detailed as possible when describing the population.

A decent sample frame for research on living conditions might be to include everyone in the target demographic. Exclude everyone who isn't part of the target group.

Issues of choosing appropriate sampling technique(s) while selecting samples

Choosing a sampling strategy is an essential step in the capture phase of the data journey and will ensure that, data is reliable and reflects the characteristics of target group. In this blog, we'll take step by step through the process by outlining the ways in which primary data is collected using an example in which a survey on characteristics (tax, education levels, etc) is collected on residents in five towns. The towns are of different sizes and have a total of 3,200 households. These 3,200 households make up the target population for survey.

Step one: Define sample and target population

At times, the survey may require covering the entire target population, as is the case in mapping or population studies. That's usually referred to as a census survey. However, target populations are generally large and expensive to survey. In our example, it may not be feasible to visit all 3,200 households of the five towns. Instead, we want to choose a smaller sample that would be representative of the population and reflect its characteristics.

A survey that is done on a smaller number of the target population is referred to as a sample survey. we can infer our findings for the entire population based on this representative sample. In the following sections, we'll describe the different terminologies that are associated with sample surveys, such as sample size and sampling technique. These concepts will enable us to determine the number of surveys needed to accurately reflect the true characteristics of a population and to choose the best method of selecting a sample from that population.

Step two: Define sample size

The first step in sampling exercise will be decided on an appropriate sample size. There are no strict rules for selecting a sample size. We can make a decision based on the objectives of the project, time available, budget, and the necessary degree of precision.

In order to select the appropriate sample size, we will need to determine the degree of accuracy that we want to achieve. For this, we'll need to establish the confidence interval and confidence level of our sample.

The confidence interval, also called the margin of error, is a plus or minus figure. It is the range within which the likelihood of a response occurs. The most commonly used confidence interval is +/- 5. If we wish to increase the precision level of our data, we would further reduce the error margin or confidence interval to a +/- 2. For example, if our survey question is "does the household pay tax?" and 65% of our sampled households say "yes," then using a confidence

interval of +/- 5, we can state with confidence that if we are asked the question to all 3,200 households, between 60% (i.e. 65-5) and 70% (i.e. 65+5) would have also responded "yes."

The confidence level tells how sure we want to be and is expressed as a percentage. It represents how often the responses from our selected sample reflect the responses of the total population. Thus, a 95% confidence level means we can be 95% certain. The lower the confidence level, the less certain we will be.

Most surveys use the 95% confidence level and a +/- 5 confidence interval. When we put the confidence level and the confidence interval together, we can say that we are 95% sure that, if we had surveyed all (3,200) households, between 60% and 70% of the households of the target population would have answered "yes," to the question "does the household pay tax?".

The size of sample may be determined using any standard sample size calculator such as Survey Monkey or Raosoft. Using a standard sample size calculator (as can be seen in table one below) for our example of 3,200 households in five towns, we can examine the difference in sample sizes based on different confidence levels and intervals.

Option A
If we decide on a 5% confidence interval and want to achieve a 95% confidence level, the sample size will be 345 households.

Option B
If we wish to have higher accuracy and increase the confidence level to 99%, the recommended sample size would be 551.

Option C
For an even higher accuracy we could choose a 2% confidence interval and 99% confidence level and arrive at a sample size of 1807.

If time and resources permit, we could opt for larger samples and choose option C, to survey 1807 households. However, the quality of our findings are likely to only be marginally better than with option A or B, as the rate of improvement of accuracy gradually diminishes with the increase in sample size. The size of sample should therefore be decided by the objectives of the study and resources available.

Table 1: Calculate your sample size

Factors	Factors description	Option A	Option B	Option C
Population (no.)	The total population that your sample will present	3,200	3,200	3,200
Confidence level (%)	The probability that your sample accurately represents the characteristics of your population	95%	99%	99%
Confidence interval (%)	The range that your population's responses may deviate from your samples	5	5	2
	Sample size calculated	*345*	*551*	*1807*

Step three: Define sampling technique

Once we've chosen the sample size for survey, we need to define which sampling technique to select sample from the target population. The sampling technique that's right depends on the nature and objectives of project. Sampling techniques can be broadly divided into two types: random sampling and non-random sampling.

Random sampling

As the name suggests, random sampling literally means selection of the sample randomly from a population, without any specific conditions. This may be done by selecting the sample from a list, such as a directory, or physically at the location of the survey. If we want to ensure that a particular household does not get selected more than once, you can remove it from the list. This type of sampling is called simple random sampling without replacement. If you choose not to remove duplicate households from the list, you would do a simple random sampling with replacement.

Systematic sampling is the most commonly used method of random sampling, whereby you divide the total population by the sample size and arrive at a figure which becomes the sampling interval for selection. For example, if you need to choose 20 samples from a total population of 100, your sampling interval would be five. Systematic sampling works best when the population is homogeneous, i.e. most people share the same characteristics. In our example, the sampling interval would be nine (3200/345= 9 for a 95% confidence level and 5% confidence interval). Thus we will select every ninth household in a town.

However, populations are generally mixed and heterogeneous. To ensure sufficient inclusion of all categories of the population, we need to identify the different strata or characteristics and their

actual representation (i.e. proportion) in the population. In such cases, we can use the stratified random sampling technique, whereby we first calculate the proportion of each strata within the population and then select the sample in the same proportion, randomly or systematically, from all the strata.

If we take our earlier example of five towns, to calculate a stratified random sample, you will need to calculate the proportion of each town within the sample size of 345 as shown in table two below. Column three gives the proportion of each town of the total population (3,200). In column four, the sample size (345) is proportionately divided across the five towns. For example, town three, which is 25% of the total population, will select 86 households with a sampling interval of nine (i.e. 800/86) in the same manner as was done for systematic sampling.

Table 2: Calculate stratified random sample

Location	Population size	Proportion (%) of population	Stratified sample size
Town 1	1200	38%	129
Town 2	900	28%	97
Town 3	800	25%	86
Town 4	180	6%	19
Town 5	120	4%	13
Total	*3200*		*345*

Non-random sampling

In non-random sampling, the sample selection follows a particular set of conditions and is generally used in studies where the sample needs to be collected based on a specific characteristic of the population. For example, you may need to select only households which own a car, or have children less than six years of age. For this, you would consciously select only the 345 or 551 households that have those characteristics. Also termed purposive or subjective sampling, non-random sampling methods include convenience, judgment, quota and snowball sampling.

Step four: Minimize sampling error

It’s normal to make mistakes during sample selection. Our efforts should always be to reduce the sampling error and make the chosen sample as representative of the population as possible. The robustness of sample depends on how we minimize the sampling error. The extents of errors during sampling vary according to the technique or method you choose for sample selection.

For samples selected randomly from a target population, the results are generally prefixed with the +/- sampling error, which is the degree to which the sample differs from the population. If our study requires to know the extent of sampling error that is acceptable for the survey, you can select a random sampling technique. In random sampling, you will be able to regulate the survey design to arrive at an acceptable level of error. In a non-random sample selection, the sampling error remains unknown.

Thus, when your sample survey needs to infer the proportion of a certain characteristics of the target population, you can select a random sampling method. But if you want to know the perceptions of residents regarding taxation laws or the school curriculum, you would want to capture as many perceptions as possible, and therefore select a non-random method in situations where sampling errors or sampling for proportionality are not of concern. Non-random sampling techniques can be very useful in situations when you need to reach a targeted sample with specified characteristics very quickly.

If you don’t have a sampling strategy in place, you may collect data which is biased or not representative, rendering your data invalid.

Follow the link for snow ball sampling (https://www.scribbr.com/methodology/snowball-sampling/)

Sample Size

It is usually a compromise between what is needed to satisfy the statistical requirement and what can be realistically be accomplished. The size of the sample is an important factor. It has direct bearing on the accuracy, estimation, cost and administration of the survey. Large sample has low sampling error where as small sample have higher sampling error. To avoid unnecessary cost small sample should be selected. Hence optimum sample size should be selected to fulfill the requirement of efficiency, representativeness, reliability and flexibility. Some of the factors affecting the sample size are nature of study, nature of reaction of respondent towards the subject under study, nature of population i.e. composition of population under study, number of classes in the population, types of sampling used during study etc.

Factors affecting Sample Size

Size of sample depends upon different factors. These are

i. Nature of population
ii. Number of classes
iii. Nature of the study
iv. Types of sampling used
v. Degree of accuracy

Nature of population

If the population under study is homogeneous then small sample size is sufficient, but in case of heterogeneous population large sample size is required to make sample size representative of the population.

Number of classes

For the classification with large number of classes, large sample size is required.

Nature of study

If the study takes long time then small sample size is better from the financial and analysis point of view.

Types of sampling used

The sample size depends upon the type of sampling used. For simple random sampling large sample size is required but for the case of stratified sampling small sample size is sufficient.

Degree of accuracy

If the greater degree of accuracy is required then large sample should be selected.

Testing Reliability of the Sample

If the selected sample is representative of the population then sample is called reliable. The selected sample is reliable or not can be tested using following methods;

i. Drawing parallel sample.
ii. Comparing sample with population.
iii. Drawing sub sample from main sample.

Drawing parallel sample

Draw sample parallel to the drawn sample from the population and compare various measures such as average, dispersion, skewness, kurtosis etc. between the samples. If the comparison measures are alike then the sample is reliable otherwise unreliable.

Comparing sample with population

Different measures computed from samples are compared with that of population. If the measures are identical then the selected sample is reliable.

Drawing sub sample from main sample

The different measures computed from sub sample arc compared with main sample. It cannot be used to find the sample is representative of population or nut but can be used to find if any error occurred due to faulty selection of sample.

Method of Estimating Sample Size

Estimation of sample size by using mean

Let $\bar{x}$ be the sample mean from a random sample of size n drawn from population with mean E($\bar{x}$) and standard deviation σ.

Now,

$$Z = \frac{\bar{x} - E(\bar{x})}{SE(\bar{x})} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

at α level of significance and $(1-\alpha)$ confidence limit is

$$P\left(\left|\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| \leq Z_{\alpha/2}\right) = 1-\alpha$$

$$P(|\bar{x} - \mu| \leq {}^{\sigma}/_{\sqrt{n}}) \; Z_{(}\alpha/2) = 1 - \alpha$$

Now, x - μ = d (margin of error) then

d = $\frac{\sigma}{\sqrt{n}} Z\ \alpha/2$

$\sqrt{n}$ = $\frac{\sigma}{d} Z\alpha/2$

n = $\frac{\sigma^2\ {Z_{\alpha_2}}^2}{d^2}$

In case of σ is not known take $\sigma = s$

For the finite population of Size N, sample size = $\frac{\sigma^2\ {Z_{\alpha_2}}^2}{d^2\ + \frac{\sigma^2\ {Z_{\alpha_2}}^2}{N}}$ = $\frac{n}{1+\frac{n}{N}}$

Estimation of sample size by using proportion

Let p be sample proportion from random sample of size n drawn from population with proportion P

Now,

Z=$\frac{p-E(p)}{SE(p)} = \frac{p-P}{\sqrt{\frac{PQ}{n}}}$

At α level of significance (1 - α) confidence limit is

P $\left(\left| \frac{p-P}{\sqrt{\frac{PQ}{n}}} \right| \leq Z\alpha/_2 \right)$ = 1 - α

or P ($|p - P| \leq \sqrt{\frac{PQ}{n}}\ Z\alpha/_2$ = 1 - α

Now, p - P = d (margin of error) then

d= $\sqrt{\frac{PQ}{n}}\ Z\alpha/_2$

$\sqrt{n}$ =$\frac{Z\alpha/_2\ \sqrt{PQ}}{d}$

n= $\frac{PQ\ Z_{\alpha/2}{}^2}{d^2}$

In case of P is not known take P = p.

For the finite population of size N, sample size = $\frac{PQ\ Z_{\alpha/2}{}^2}{d^2+\frac{PQ\ Z_{\alpha/2}{}^2}{N}} = \frac{n}{1+\frac{n}{N}}$

Example 1

Determine the minimum sample size required so that the sample estimate lies within 10% of the true value with 95% level of confidence when coefficient of variation is 60%.

Solution

Here,

C.V. = 60% = 0.6

P ($|\bar{x} - \mu| \leq 0.1\ \mu$) = 0.95

(i) Confidence level (1 - α) = 95% = 0.95 then α = 0.05

Now,

$P\left(\left|\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\right| \leq Z_{\alpha/2}\right)$ = 1- α

$P(|\bar{x} - \mu| \leq {}^{\sigma}/_{\sqrt{n}})\ Z_{(}\alpha/2) = 0.95$

$P(|\bar{x} - \mu| \leq 1.96 \times \frac{\sigma}{\sqrt{n}})$ =0.95

From equation (i) and (ii)

$0.1\mu = 1.96 \times \frac{\sigma}{\sqrt{n}}$

$\sqrt{n} = \frac{1.96}{0.1} \times \frac{\sigma}{\sqrt{n}}$

$n = \left(\frac{1.96}{0.1} \times \frac{\sigma}{\sqrt{n}}\right)^2$

$n = 384.16 \times CV^2$

$n = 384.16 \times (0.6)^2$

or $n = 138.29 = 138$

Hence required sample size is 138.

Example 2 In measuring reactions time, a psychologist estimates that the standard deviation is 0.05 seconds. How large a sample of measurement must be taken in order to be 99% confident that the error of his estimate will not exceed 0.01 seconds?

Solution

Here,

Sample size (n) = ?

Standard deviation (s) = 0.05

Confidence interval $(1 - \alpha) = 99\% = 0.99$

or $\alpha = 0.01$ $Z_{\alpha/2} = 2.58$

Error (d) = 0.01

Here $\sigma = s$

$n = \frac{\sigma^2 \ {Z_{\alpha_2}}^2}{d^2} = \frac{(0.05)^2(1.96)^2}{(0.01)^2} = 166.4 \approx 167$

Hence required sample size is 167.

Example 3

A researcher wants to conduct a survey of disabled at Kathmandu valley. What should be the sample size of the prior estimate of population of disables in the population is 10% and the desired error is estimation is 2% and level of significance is 5%.

Solution

Here

Sample size (n) = ?

Population proportion (p) = 10% = 0.1

q = 1 - p = 0.9

Error (d) = 2% = 0.02

Level of significance (α) = 5%

Here P = p

$n = \frac{PQ\ Z\alpha_{/2}^{\ 2}}{d^2} = \frac{(1.96)^2 \times 0.1 \times 0.9}{(0.02)^2} = 864.36 \approx 865$

Hence required sample size is 865**.**

Example 4

For p 0.2, d = 0.05 and z = 2 find n. Also find n if N = 1000.

Solution

Here P= p

Now

$n = \frac{PQ\ Z\alpha_{/2}^{\ 2}}{d^2} = \frac{4 \times 0.2 \times (1-0.2)}{0.05^2} = 256$

When N = 1000

Sample size = $\frac{n}{1+\frac{n}{N}} = \frac{256}{1+\frac{256}{1000}} = 203.82 \approx 204$

Example 5

The mean systolic blood pressure of a certain group of people was found to be 125 mm of Hg with standard deviation of 15 mm of Hg. Calculate sample size to verify the result at 5'y level of significance if error do not exceed 2. Also find sample size if sample is selected from population of size 500.

Solution

Here,

Standard deviation (s) = 15

Level of significance (α) = 5%

Sample size (n) = ?

Error (d) = 2 Here σ = s

Now,

n = $\frac{\sigma^2 \ {Z_{\alpha_2}}^2}{d^2} = \frac{(15)^2(1.96)^2}{(2)^2} 216.09 \approx 216$

When N = 500

Sample size = $\frac{n}{1+\frac{n}{N}} = \frac{216}{1+\frac{216}{500}} = 150.83 \approx 151$

YAMANE FORMULA

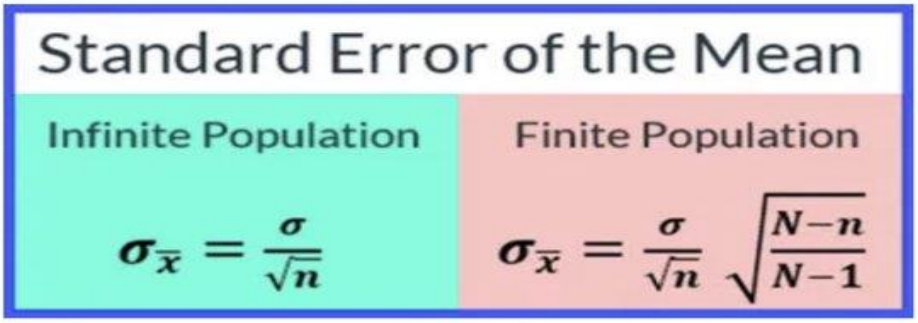
n=$\frac{N}{1+N\times e^2}$

Where n= the sample size

N= the population size

e= the acceptable sampling error

95% confidence level and p= 0.5 are assumed



Working with a finite population and if the population size is known, the **Yamane formula for determining the sample size is given by:**

n=N/(1+Ne2)

Where

n= corrected sample size, N = population size, and e = Margin of error (MoE), e = 0.05 based on the research condition.

Let's assume that the population is 10,000. At 5% MoE., the sample size would be: 10000(1+10000(.052)

1000026

= 384.61 ~ 385

In a finite population, when the original sample collected is more than 5% of the population size, the corrected sample size is determined by using the Yamane's formula.

In the example above, 5% of 10,000 is 500 and hence the corrected size is 385 although for research purposes, even 385 is a big number (for handling and collection point of view) and the researcher has to make a decision to collect even smaller number in order of ease of handling, costing but he has to ensure that the sample is representative.

Statistical Analysis

THE RESEARCH analyst breaks down data into constituent parts to obtain answers to re-search questions and to test research hypotheses. The analysis of research data, however, does not in and of itself provide the answers to research questions. Interpretation of the data is necessary. To interpret is to explain, to find meaning. It is difficult or impossible to explain raw data; one must first analyze the data and then interpret the results of the analysis.

Analysis means the categorizing, ordering, manipulating, and summarizing of data to obtain answers to research questions. The purpose of analysis is to reduce data to intelligible and interpretable form so that the relations of research problems can be studied and tested. A primary purpose of statistics, for example, is to manipulate and summarize numerical data and to compare the obtained results with chance expectations. A researcher hypothesizes that styles of leadership affect group-member participation in certain ways. He plans an experiment, executes the plan, and gathers data from his subjects. Then he must so order, break down, and manipulate the data that he can answer the question: How do styles of leadership affect group-member participation? It should be apparent that this view of analysis means that the categorizing, ordering, and summarizing of data should be planned early in the research. The researcher should lay out analysis paradigms or models even when working on problem and hypotheses. only in this way can he see, even if only dimly, whether his data and its analysis, can and will answer the research questions.

Data Editing

Editing is the process of examining errors and omissions in the collected data and to make necessary corrections. Data should be edited after getting the filled up questionnaire or schedule and before entering in to the step of data processing. It is done to assure the data are accurate, consistent, uniformly entered, complete and well arranged. Editing is carried out in two stages:

(i) field editing
(ii) central editing

Field editing is the review of reporting forms by enumerator or investigator for completing what the signs and symbols have written in abbreviated form in the time of recording respondents' response. Central editing is editing obvious errors such as entry in wrong place, missing replies etc. by editor when all schedules or forms have been completed and returned to office.

Data Coding

It is process of assigning numerals or other symbols to answers so that response can be put into a limited number of class or categories. The quantitative data collected using questionnaire or schedule is numeric so that no need of coding. For the data which is qualitative in nature the numeric codes are to be used before the analysis. For the statistical treatments qualitative responses are to be converted into numerical figures which satisfy, all the rules of arithmetic operation. Different social scales are used on assigning numerical figures to the qualitative response. For example for male and female code 1 and 0 are used.

Classification of Data

The data contained in questionnaire or schedule will not enable us to see quickly all possible characteristics. In order to make data easily understandable the classification is adopted. Classification is the process of arranging the related facts or data into different groups or classes according to their similarities. Facts differ from class to class with re characteristics which is the basis of classification.

The classification should be

(i) according to research problem
(ii) exhaustive
(iii) mutually exclusive
(iv) independent

The main objectives of classification are

i) to condense mass of data
ii) to facilitate comparison
iii) to pinpoint feature of data at a glance
iv) to enable statistical treatment

Types of classification: Statistical data are classified in respect of their characteristics. Broadly there are four basic types of classification namely

a) Chronological classification
b) Geographical classification
c) Qualitative classification
d) Quantitative classification

a) **Chronological Classification**
In chronological classification the collected data are arranged according to the order of time expressed in years, months, weeks, etc.. The data is generally classified in ascending

order of time. For example, the data related with population, sales of a firm, imports and exports of a country are always subjected to chronological classification. For example, The estimates of birth rates in India during 1970 - 76 are

Year	1970	1971	1972	1973	1974	1975	1976
Birth rate	36.8	36.9	36.6	34.6	34.5	35.2	34.2

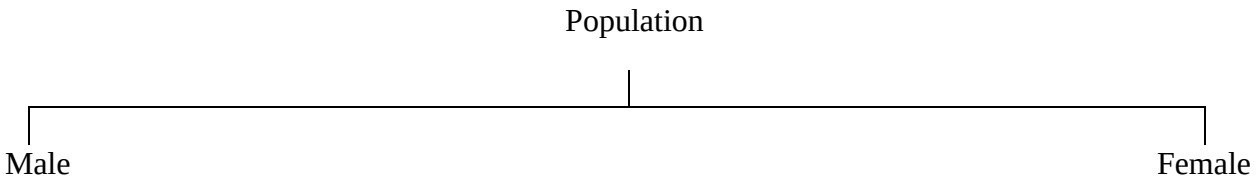
b) **Geographical Classification**

In this type of classification the data are classified according to geographical region or place in For instance, the production of paddy in different states in India, production of wheat different countries etc. For example;

Country	America	China	Denmark	France	Nepal
Yield of wheat kg/acre	1925	893	225	439	862

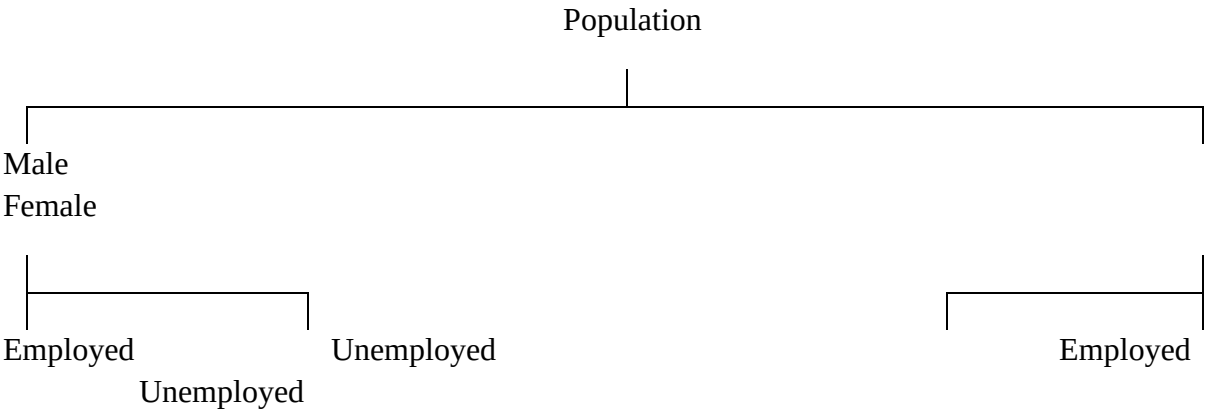
c) **Qualitative Classification:**

In this type of classification data are classified on the basis of same attributes or quality like: sex, literacy, religion, employment etc. Such attributes cannot be measured along with a scale.



(i) The classification, where two or more attributes are considered and several classes are formed, is called a manifold classification. For example, if we classify population simultaneously with respect to two attributes e.g., sex and employment then population are first classified with respect to 'sex' into 'males' and 'females'. Each of these classes may then be further classified into 'employment' and 'unemployment' on the basis of attribute 'employment' and as such population are classified into four classes namely. (i) Male employed (ii) Male unemployed (iii) Female employed ((iv) Female unemployed

Still the classification may be further extended by considering other attributes like marital status etc. This can be explained by the following chart



(d) Quantitative classification

Quantitative classification refers to the classification of data according to some characteristics that can be measured such as height, weight, etc.

Data Entering into Spreadsheet

A spreadsheet is an interactive computer application program for organization and analysis of data in tabular form. Spreadsheets developed as computerized simulations of paper accounting worksheets. The program operates on data represented as cells of an array organized in rows and columns. Each cell of the array is a model-view-controller element that can contain either numeric or text data or the results of formulas that automatically calculate and display a value based on the contents of other cells. The user of the spreadsheet can make changes in any stored value and observe the effects on calculated values. This makes the spreadsheet useful for "what-if" analysis since many cases can be rapidly investigated without tedious manual recalculation. Modern spreadsheet software can have multiple interacting sheets and can display data either as text and numerals or in graphical form.

Management of Missing and Inconsistent Information

Generally, in data management activity of research work a researcher faces the threats of

i) Missing data
ii) Impossible values
iii) Inconsistencies and
iv) Transcription errors.

Missing and inconsistent data (information) are a part of almost all research and all the researcher have to deal with it from time to time. There are various alternative ways of dealing with missing data. To get data with less missing and inconsistent information attention should be given from the designing questionnaire to data entry. Most of the

missing are arisen in the survey field because of the imperfection of the field workers and the less skill of the person who involved in data entry work.

There are a number of strategies for handling missing and inconsistent data and common will be described here. These methods can be accomplished with standard statistical software packages (SAS, SPSS). Special issues arise when dealing hods are data that are categorical and treatment of these issues is very difficult. Deletion met sog not as much a strategy for handling missing data as they are approaches to ignoring nu data. We just introduce two deletion methods. with 5651

List wise deletion

In this method, cases with any missing values are deleted from an analysis. This method is sometimes called *complete case analysis* because only cases with complete data are retained. This is the default procedure for many statistical programs but it is generally not an advisable method.

Pairwise deletion

In this method, the maximum amount of available data is retained and so this method is sometimes referred to as *available case analysi*s. Cases are excluded from only operations in mussing which data are missing on a variable that is required. In a correlation matrix, for example a case that was data on one variable would not be used to calculate the correlation coefficient between that variable and another but would be included in all other correlations.

Ways of managing Missing and Inconsistent data

The quality of the data can be kept up by careful and systematic method of data cleaning. The following are the steps by using which we can reduce the inconsistent observation and problems of missing data.

- Develop a plan for data management
- Make a an intensive training before data entry
- Make strategy of getting quality data
- Maintain the question that can cross check the responses
- Adopt checking system of impossible values
- Record the variables and if possible create composite variables
- Preferably do not make a change in raw data set, if changed logically it should be documented
- Coding system should be preferably used to reduce such errors
- Use standard methods of data cleaning using software
- Use the method of labeling the values and merge cells if possible

- Once the data set is cleaned, the next step is to format it for analysis
- Data formatting should be done using code
- Maintain a master dataset that is distributed to everyone conducting analyses

Descriptive Statistical measure

Types of average

I. Arithmetic Mean
II. Geometric mean
III. Harmonic mean
IV. Median
V. Mode

Measure of Dispersion

Absolute and relative measure of dispersion

Types of dispersion

Range

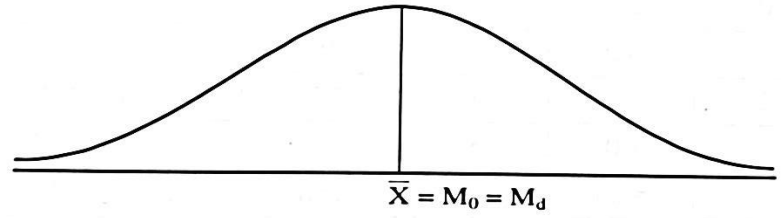
Quartile deviation

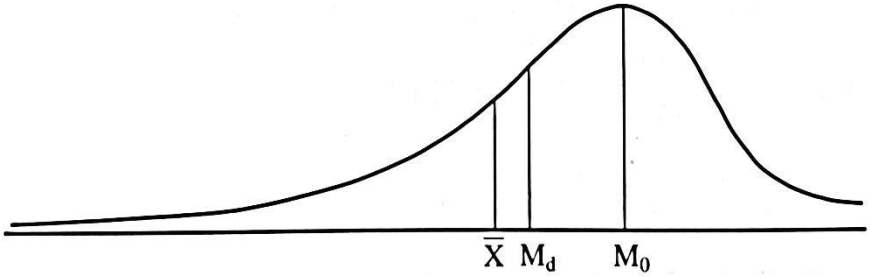
Mean deviation

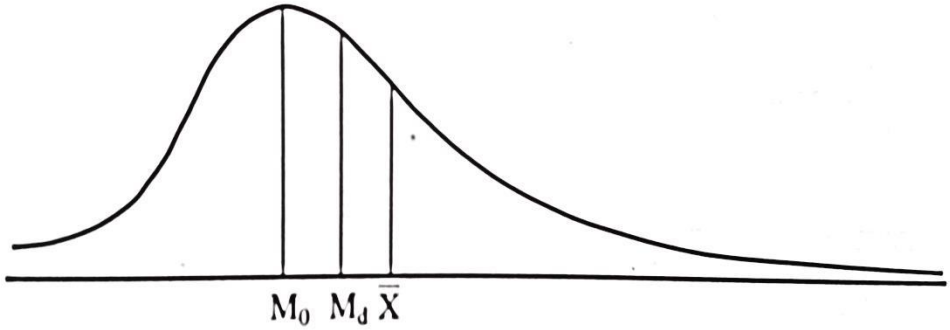
Standard deviation

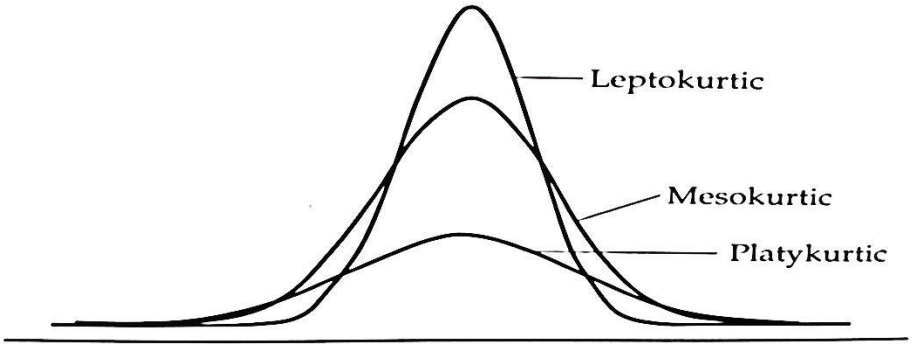
(Coefficient of Variation)

Skewness **kurtosis** $K=\frac{Q_3 - Q_1}{2(P_{90}-P_{10}}$ **K=0.263**









Correlation and Regression

Inferential statistics

Testing of Hypothesis

Z test

It is important parametric test based upon the normality assumption. Traditionally Z test is used, when the samples are selected from population of known parameter with sample size more than 30. We consider that if sample size is more than 30 then sample selected from non normal population is also approximately normal distributed.

Z test is defined as the ratio of difference between t and E(t) to the S.E.(t)

$$Z = \frac{t - E(t)}{SE.(t)} \sim N(0, 1),$$

where t = statistic, E(t) = Expected value of statistic and S.E.(t) = Standard error of the statistic.

Z test is used to test

- Significance of single mean.

- Significance of difference between two means.

- Significance of single proportion.

Significance of difference between two proportions

- Significance of difference between sample correlation and population correlation.

- Significance of difference between independent sample correlations

Test of significance of a single mean

Let us consider sample of size n (n > 30) has been drawn from the normal population N (μ, σ^2) then the sample mean $\bar{x} \sim N\ (\mu, \sigma^2)$.

Different steps in the test are;

Problem to test

Ho:$\mu = \mu_0$ (sample is drawn from population with mean μ_0)

H1: $\mu \neq \mu_0$ (Two tailed test)

or H1: $\mu > \mu_0$ (One tailed right)

or H1: $\mu < \mu_0$ (One tailed left)

Test statistic

For the sample selected from the population of unknown size

$Z = \frac{\bar{X} - E(\bar{X})}{SE.(\bar{X})} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ for known variance

$\frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$ for unknown variance(for large sample size($\hat{\sigma} = s$)

For the sample selected from the population of known size

$Z = \frac{\bar{X} - E(\bar{X})}{SE.(\bar{X})} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$ for known variance

$\frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$ for unknown variance

Where $\bar{X}$ = sample mean, μ = population mean, σ = population s.d. s = sample s.d.,

N = population size, n = sample size

Level of significance

Let α be the level of significance. Usually we take α = 0.05 unless we are given.

Critical value

Critical or tabulated value of Z is obtained from table according to the level of significance and alternative hypothesis.

Decision

Reject Ho at α level of significance if I z I > $Z_{tabulated}$, accept otherwise.

Example

A sample of 400 students is found to have mean height of 170 cm. Can it be reasonably regarded as a sample from a large population with mean height 169.5 cm and standard deviation 3.5 cm?

Solution

Here,

Sample size (n) = 400

Sample mean ($\bar{X}$) = 170

Population mean (μ) = 169.5

Population S.D. (σ) = 3.5

Problem to test

Ho : Mean height of students is 169.5 cm (μ = 169.5)

H_1 = Mean height of student is not 169.5 cm ($\mu \neq 169.5$) (Two tailed)

Test statistic

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{170 - 169.5}{\frac{3.5}{\sqrt{400}}} \quad \frac{0.5 \times 20}{3.5} = 2.857$$

Critical value

Let 5% be the level of significance then critical value is Ztab = $Z\alpha/2$ = 1.96 Decision Here Z= 2.857 > Z_{tab} = 1.96, reject Ho at 5% level of significance.

Conclusion

The sample of 400 students cannot be regarded as sample from large population with mean height 169.5 cm and standard deviation 3.5 cm.

Test of significance difference between two means

Let us consider two independent samples of size n_1 and n_2 be drawn from population having means means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 respectively. Let $\bar{X}_1$ and $\bar{X}_2$ be the sample means.

For large n_1 and n_2.

$$\bar{X}_1 \sim N(\mu_1, \frac{\sigma_1^2}{n_1}$$

$$\bar{X}_2 \sim N(\mu_2, \frac{\sigma_2^2}{n_2}$$

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$$

Different steps in the test are

Problem to test

Ho: $\mu_1 = \mu_2$ There is no significant difference between two population mean.

H1: $\mu_1 \neq \mu_2$ (two tailed)

or H1 : $\mu_1 < \mu_2$ (one tailed left)

or H1 : $\mu_1 > \mu_2$ (one tailed right)

Test statistic

$$\boldsymbol{Z} = \frac{\bar{X}_1 - \bar{X}_2 - E(\bar{X}_1 - \bar{X}_2)}{\boldsymbol{S.E.}(\bar{X}_1 - \bar{X}_2)}$$

$$\boldsymbol{Z} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

When population means and variances are known

$$\boldsymbol{Z} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

When population variances are known

$$\boldsymbol{Z} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

When population variances are unknown

for large sample size $\widehat{\sigma_1^2} = S_1^2$ and $\widehat{\sigma_2^2} = S_2^2$

$\bar{X}_1$ = sample mean of size n_1,

$\bar{X}_2$ = sample mean of size n_2

σ_1^2 = population variance of first population

σ_2^2 = population variance of second Population

S_1^2 = sample variance of first sample

S_2^2 = sample variance of second sample.

Level of significance

Let a α be the level of significance. Usually we take α = .05 unless we are given.

Critical value

Critical or tabulated value of Z is obtained from table according to the level of significance and alternative hypothesis.

Decision

Reject Ho at a level of significance if I z I > $z_{tabulated}$, accept otherwise.

Example

In a random sample of 500 the mean is found to be 20. In another independent sample of 400 the mean is 15. Could the samples have been drawn from the same population with S.D. 4? Solution Here,

Sample size of first sample (n_1) = 500

Sample mean of first sample ($\bar{X}_1$) = 20

Sample size of second sample (n_2) = 400

Sample mean of second sample ($\bar{X}_2$) = 15

Population S.D. of first (σ_1) = 4

Population S.D. of second (σ_2) = 4

Problem to test

Ho : $\mu_1 = \mu_2$ (both the populations are same)

H1: $\mu_1 \neq \mu_2$ (population are different)

Test statistic

$$\boldsymbol{Z} = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$Z = \frac{(20-15)}{\sqrt{\frac{16}{500}+\frac{16}{400}}} = 18.51$$

Critical value

Let α = 5 % be the level of significance the critical value is $_{Ztabulated}$ = Z α /2 = 1.96. .

Decision

Z =18.51 > Ztabulated = 1.96, reject Ho at 5% level of significance.

Conclusion

We cannot conclude that the samples have been drawn from the same population.

Test of significance difference between two proportions:

Let P_1 and P_2 be the two population proportions possessing a certain characteristic. Let two independent samples of sizes n_1 and n_2 be drawn from the two population. Also p_1 and p_2 be the proportion of units possessing certain characteristic in the two samples.

For large sample size

$p_1 \sim N(p_1, \frac{P_1Q_1}{n_1})$

$p_2 \sim N(p_2, \frac{P_2Q_2}{n_2})$

Then

$p_1 - p_2 \sim (p_1 - p_2, \frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2})$

Different steps in the test are;

Problem to test

$H_0 : P_1 = P_2$

$H_1 : P_1 \neq P_2$ (Two tail test)

$H_1 : P_1 > P_2$ (One tail right)

$H_1 : P_1 < P_2$ (One tail left)

Test statistic

$$\boldsymbol{Z} = \frac{(p_1 - p_2) - E((p_1 - p_2)}{\boldsymbol{S.E.}((p_1 - p_2)}$$

$$= \frac{(p_1 - p_2 - (P1 - P2)}{\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}}$$

$$= \frac{(P1 - P2)}{\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}}$$

If population proportion are given

$$= \frac{(p_1 - p_2}{\sqrt{\boldsymbol{PQ}(\frac{1}{n_1} + \frac{1}{n_2})}}$$

If population proportion are not given

Where P_1 = population proportion of first population

P_2 = population proportion of second population

p_1= sample proportion of first sample of size n_1

p_2 = sample proportion of second sample of size n_2

Level of significance

Let α be the level of significance. Usually we take α = .05 unless we are given.

Critical value Critical or tabulated value of Z is obtained from table according to the level of significance and alternative hypothesis.

Decision

Reject Ho at a level of significance if I z I > $Z_{tabulated}$, accept otherwise.

Example

A machine puts out 21 defective articles in a sample of 500 articles. Another machine gives 3 defective articles in a sample of 100 are the two machines significantly different in their performance? Use p value method at 1% level of significance.

Solution

Here Defective articles by a machine (x_1) = 21

Number of articles by a machine (n_1) = 500

Defective articles by another machine (x_2) = 3

Number of articles by another machine (n_2) = 100

Sample proportion of defective article by a machine (p_1) = $\frac{x1}{n1} = \frac{21}{500}$ = 0.042

Sample proportion of defective article by another machine (p_2) = $\frac{x2}{n2} = \frac{3}{100}$ =0.03

Let

P_1 = Population proportion of defective from a machine

P2 = Population proportion of defective from another machine

P = $\frac{n_1p_1+n_2p_2}{n_1+n_2} = \frac{500\times0.042+100\times0.03}{500+100} = \frac{24}{100} = 0.04$

Level of significance (α) = 1%

Problem to test

Ho : $P_1 = P_2$ (There is no significance difference in performance of machines)

H_1 : $P_1 \neq P_2$ (There is significance difference in performance of machines)

Test statistic

$$Z = \frac{(p_1 - p_2}{\sqrt{PQ(\frac{1}{n_1}+\frac{1}{n_2})}}$$

$$\frac{0,042-0.03}{\sqrt{0.04\times0.96(\frac{1}{500}+\frac{1}{100})}} = 0.571$$

Now prob(Z≥ Zcalculated) $= \text{prob}(Z \geq 0.571) = 0.5 - \text{prob}(0 \leq Z \leq 0.571)$

$$= 0.5 - 0.2175 = 0.284$$

For two tailed test, p value = 2 Prob ($Z \geq Z_{calculated}$) = 2 x 0.284 = 0.568

Here α = 1% = 0.01

Decision P value = 0.568 > α = 0.01, accept H_0 at 1% level of significance.

Conclusion

There is no significant difference in performance of two machines.

t test

When the sample size is small (traditionally it is assumed less than or equal to 30), then the sampling distribution of the sample mean is assumed to follow student's t distribution. The t distribution is also similar to normal distribution having shape as in normal distribution but little bit flatter. As the sample size increases the shape of t distribution is more likely to normal curve. Whatever be the sample size the statistical software uses the t test for all sample size instead of Z test, since it can compute the tail area of the curve (p value) or to compare with the pre-assigned value of a.

t test is based upon the assumption that

- Sample size small
- Sample is selected from normal population.
- Population standard deviation is not known.
- Samples are independent.

It is used to test

- **Significance of single mean**.
- **Significance of difference between means.**
- **Significance of correlation coefficient.**
- **Significance of regression coefficient**.

Some other test

i)Chi- square test **ii) ANOVA (Analysis of Variance)**

iii) Run test **iv) Sign test**

v) Mann Whitney U test **vi) The Kruskal- Wallis test**

Preparation of Research Report

Research Report

A report is simply a statement or description of things that have already occurred. It is concise, clear communication of the important findings of the research work. Reports communicate information which has been compiled as a result of research and analysis of data and of issues. Reports focus on transmitting information with a clear purpose, to a specific audience. Good reports are documents that are accurate, objective and complete. They should be well-written, clearly structured and expressed in a way that holds the reader's attention and meets their expectations.

A research report has a special feature in that it conveys information to the evaluator about the entire activities you had undertaken during the research process. It is the only truly effective way in which you can communicate with your examiners or evaluators about the interesting findings of your work and the new knowledge you have generated.

Purpose and Importance of Research Reports

■ It is a means whereby the data, analysis and conclusions are placed in an organized form. These information can be used both for academic and application purposes.

■ Your research work will be judged mainly by the quality of the report. The examining committee may not see your effort while at field. Your research report is the "only tangible product" of hundreds of hours of work. Therefore, you should show your performance, skills, and thoughts in your report, as these are vital to its assessment and grading.

■ The effectiveness of the report may be judged by its use. The organizations, professors, researchers and students are using good reports for different purposes. Report writing may benefit you in any of the following ways, although the relative importance of each will change as the program of work progresses:

■ To see whether you are on target with your work, so that any problems can be spotted in time to be attended to.

■ To provide an opportunity for you to reflect on progress, consolidate arguments and identify any gaps in knowledge, data or methodology.

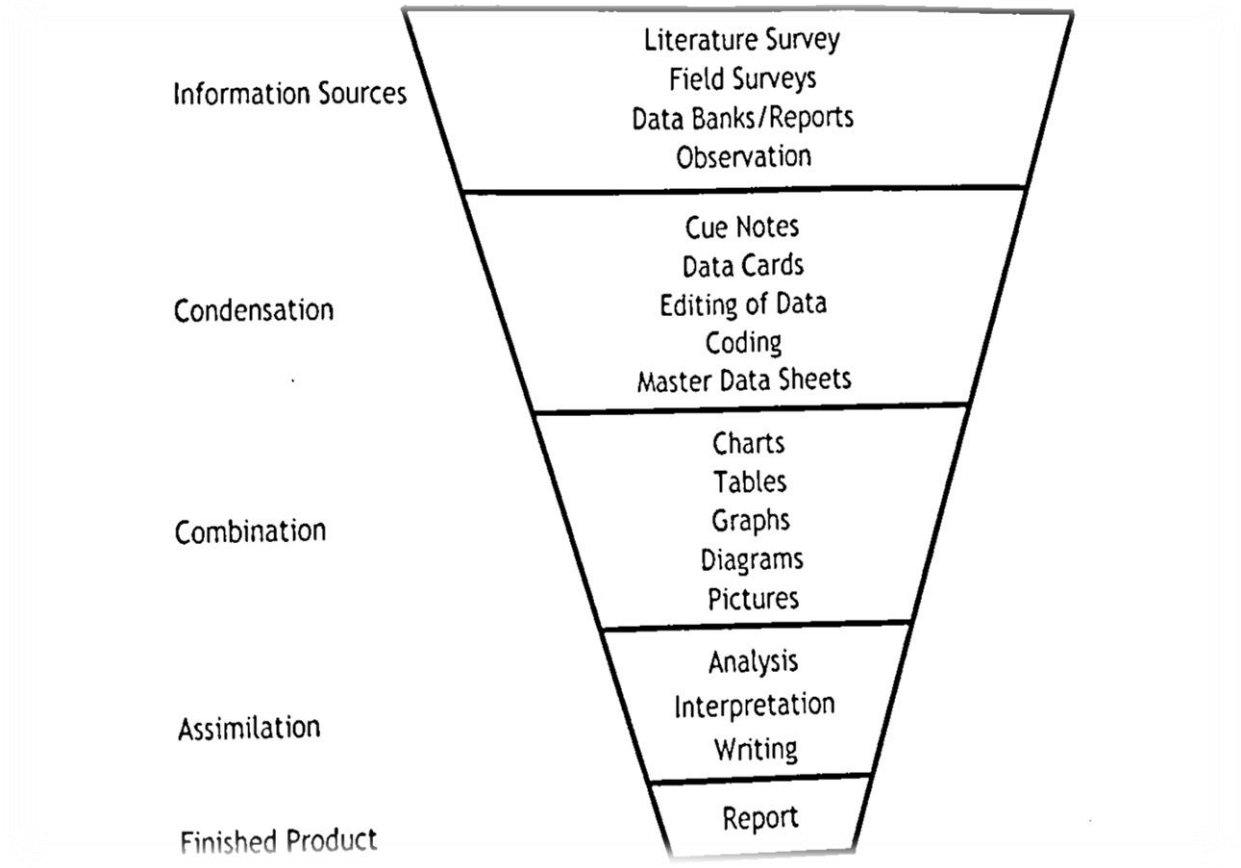
■ To help you to develop an appreciation of standards and hence to learn to monitor your own progress.

■ To provide practice in academic report writing and academic discourse, so that any additional training which may be necessary in this report can be supplied at an early stage.

■ To form a basis, in due course, for your project work and possibly a journal article.

Research Report Process

There are five steps in the research reporting process. At the top, we have raw data collected from different sources. The data is then passed through different stages of compression. Given figure highlights this report process.



Types of Research Reports

Research reports prepared for a business situation may be of various types. These reports could be classified as: (a) formal-informal, (b) written-oral, (c) internal-external, (d) short-long, (e) informational-analytical, and (f) technical-popular. All of these types of reports have their own formats and conventions. For the purpose of thesis/dissertation, reports can be classified into two functional groups: (a) descriptive, and (b) analytical.

Descriptive Reports

Descriptive reports are description of facts, trends or opinions gathered by you in course of your investigation. The presentation and analysis of facts in an organized way may be of real value in properly understanding the situation. These reports indicate the current situation and the nature of the problems facing the organization under study and also indicate the reforms required to overcome the problems.

Analytical Reports

Analytical reports go one step further than descriptive reports. These reports, in addition to presenting facts and statistics, interpret this information in relation to the problem under consideration. Focused on a single or limited area of the problem, these reports follow the process of scientific investigation and reporting. These reports also recommend the actions to be taken for improvements in the situation. You may choose any one of these types of reports. The choice of the types would depend on the nature of your investigation..

Procedures for Report Writing

There are no set rules for writing a research report. A procedure that works well for one person may not work so well for another.

Analyze the Task

As with any assignment task, you might first analyze what is expected of you. This careful reading of the assignment task as outlined in your courses of instruction. You may find the following questions useful when analyzing the task:

I. What is the purpose of the report?
II. What should I look up?
III. Who will read my report?
IV. What is the word limit?
V. What is the expected format of the report?

Prepare Outline

In view of your task analysis and the data that you have collected, you can draw an outline of reporting prior to actually starting the write up. An outline is a roadmap to keep you from getting lost when you start to write. This will help you in arranging ideas in a better way. This will also enable you to see broadly which information you want to communicate and how the various points are related to one another. Thus, for effective report writing, it is essential to plan its contents well.

Plan Your Time

The campus/college decides the date of report submission. Therefore, you may not have enough time to work at ease. As you have to meet the time schedule, it is always better for you to plan the available time accordingly. Delays in submission of the research report may cause problems for the student as well as the campus/college.

Arrange Data

You may have collected the data from different sources. The raw data need to be processed and tabulated first. After some editing work is done, the data would be somewhat revealing. You then have to arrange the data in some sequential order for meaningful presentation. The better you organize and arrange the data, the better the data will be revealing.

Start Writing

In writing, the beginning is often difficult. You may begin with the introduction and proceed through to the conclusion. But it is always better to start writing the sections, which appear to be especially easy. This would increase the enthusiasm for the task and enhance your level of confidence.

Prepare the First Draft

You should never expect to produce the excellent report in the first attempt. This may not even be possible for experienced writers. Therefore, it is better-to prepare the draft first. Then you can rewrite where necessary.

Put the Report aside for a Day or Two

After preparing the first draft, you can keep the report aside for a day or two and can do some other work. Forgetting about the report for a day or two can be very beneficial. Then you review the report afresh, almost as objectively as another person. If you find weaknesses, you can remove or correct them.

Review and Rewrite

In any report, there is always some scope for improvement. You must review your work in terms of the format and style of reporting. You must ensure that you have followed the format and style prescribed by your campus/college. You should also review your work to identify any errors in grammar, spelling, punctuation, sentence structure, etc. Revision or editing is not the same as re-writing the whole thing from scratch. What you are doing is taking a close and careful look at each word, sentence, and paragraph to make sure you have made the best choices.

Style of Report Writing

We have discussed some points, which should be kept in mind while preparing the report. At the same time, attention should also be given to style of writing. Writing style is much more an art than a matter of rules. We can, therefore, suggest only some broad guidelines.

Write Clearly

The sentence must be as simple as possible. Two or three relatively simple sentences may convey an idea more clearly than one complicated sentence. Similarly, you should not write long paragraphs. You should divide the material into separate paragraphs and use sub-headings, if necessary, to highlight important separate points. The words used must express precisely what you want to say. Finally, you must make sure that the report has uniform style and format.

Adhere to the Study Objectives

You should focus on the research problem. The main purpose of the research investigation is to answer the questions that derive from the statement of objectives. The findings or a simple piece of writing without reference to the objective may be a futile exercise.

Be Careful of Terminology, Grammar and Spelling

The fundamental medium of communicating one's findings is words. Therefore, you should give adequate attention to the correct use of grammar and to the correct spellings of all wordy Similarly, appropriate punctuation marks must also be used according to the standard rules' and not arbitrarily.

Be Selective

Research reports cannot include everything known on the given subject. If an attempt is made to include too much, there is always danger that the important points will be lost in the detail. Unnecessary accumulation of materials may distort the focus. Hence, you should make good judgment as to what materials to include in the report and what to exclude.

Be Objective

Objectivity is an important determinant of the quality of a research report. Hence, objectivity should be maintained in your research methods and interpretations. You must at all times retain your objectivity. In other words, facts should determine conclusions. Objectivity is essential because it is necessary that others be able to understand and replicate a finding before it is considered dependable. Hence, extra caution should be taken to keep the research reporting process bias-free.

Draw Conclusions Students often confuse conclusions with research findings. Findings state facts. Conclusions represent inferences drawn from the findings. Findings are just like threads in

a handloom. These threads need to be woven, intermixed and converted into a finished product called conclusion..

Conventions of Academic Writing

There are certain conventions of academic writing. Hence, it is suggested that the academic and project report writing should follow these basic norms and conventions. Moreover, a written research report is the means by which you present your work to the world. It is therefore worth taking a lot of care over its design and writing.

Make Direct and Positive Sentences

You should not use unessential words and phrases. You should avoid unnecessarily long, technical, or unusual words or phrases. Thus, well-constructed sentences are a mark of skill in writing. In writing a report, you should write naturally and directly using familiar words, short sentences with simple constructions.

Presentation

Charts, sections, sub-sections, tables etc. should be labeled adequately. The system of headings and subheadings should be kept simple. The report must be a coherent whole; it must be a tightly woven fabric of facts and ideas - of sections, subsections, paragraphs and sentences. The sequence of the sections and subsections should be logical and clear. The introduction, data analysis and findings are to be in the past tense; conclusion in the tense; and recommendations in the future tense.

Use of the First Person

The report is to be written in the third person. The pronouns such as I, my, mine, our, ours, We, us and me should be eliminated from the report. If required to refer to yourself, you use the Words "the writer" or "the investigator".

Use Gender-neutral Language

When writing a research report, it is very important to select terminology which treats both genders equally. You should not make assumptions about one gender as opposed to the other. When reporting empirical data in a project work, however, there may often be instances where it is necessary to refer to the gender of a respondent. In such instances, the use of the relevant pronoun, he or she, may be appropriate and even necessary for the reporting of the research.

Avoid Emotional Terms

Instead of writing "Sales increased tremendously", "The increase was fantastic" or "The amazing increase was attributed to... ", simply state the percentage of increase in sales. By so doing, you are neither passing judgment nor trying to make the reader pass judgment.

Label Opinions

Generally, facts are strongly preferred over opinions. Sometimes, however, opinions add conviction. The opinions of specialists may be available when facts from research are not available. When presenting opinions, you can reveal the background and identity of the person presenting the opinion. Sometimes, data are such that no solid conclusion can be drawn, or a variety of explanations could exist. In such cases, opinions can strongly substantiate your explanations and conclusions.

Use of Notes and Footnotes

It is sometimes necessary to include in a project report additional material which it is felt to be inappropriate to place in the main text. This kind of material may be information which is supplementary to the principal arguments of the text. Such material may be included in footnotes at the bottom of the relevant page, or in notes at the end of each chapter.

Non-English Terms and Expressions

Non-English terms are sometimes used in project research reports. The terms which are used regularly in English should not be italicized. However, an expression such as chakka jam would probably be italicized. When using an expression from another language, it is worth considering whether it is likely to be understood by most readers.

English and American Spellings The issue which occurs most commonly in academic writing is the selection of the form of verbs which can end in either Ise' or `ize'. In English academic writing, the latter form is the norm. Examples of such regularly used words in research include hypothesize, synthesize, and socialize. However, both these formats are equally acceptable. You should use one spelling format consistently.

Abbreviations

Abbreviations should be used sparingly. When inserted in the text to any great extent, they have the effect of breaking up the text, and making it more difficult to read. The reader may forget the meaning of some of the abbreviations and need to turn back regularly to consult the glossary. When the abbreviation is used for the first time, then the full form of the term should be given first, followed by the abbreviation in brackets afterwards. From that point onwards, the abbreviation may be used on its own.

Confidentiality and Anonymity

As far as possible, you should try to maintain confidentiality and anonymity. If the research involves any type of case study, then it is possible to use fictional names. The main advantage in this is that it gives the report an air of reality.

Consistency

While writing, consistency in the spellings, abbreviations, style, etc. should be maintained. It is therefore to make a choice of format, and then to abide by that choice throughout the report. Even slight variations in practice, can be an irritant for the evaluators of your thesis.

Typing the Research Report

Writing a report requires special writing skills and knowledge of format. Now that the word-processing equipment is becoming common, reports can be typed quickly and efficiently from relatively rough drafts. Word processors can make the typing work easy and fast. The following guidelines are given for typing the report:

Paper

For the purpose of typing white Xerox paper may be used. The size of the paper should be 8.6 by 11 inches. Only one side of the paper is to be used.

Chapter Page

The chapter number is centered about two inches from the top of the page. Following to spaces below should be the title of the chapter in capital letters. The first line of the text should begin four spaces below that title.

Margins

Margins indicate the boundaries of the text. APA specifies 1-inch margins all around (top, bottom, left, and right).

Spacing

The text of the report should be double-spaced. Indented, quotations and footnotes should be single-spaced. Same style and size of font should be used throughout the report.

Page Number

Page number should come at the top right hand corner of the page, one inch from the top edge and one inch from the right-hand edge of the page. The first line of the text should be two spaces below the page number.

Pagination

Pages should be numbered consecutively in Arabic numerals from the first page of the text to the end of the manuscript (including the appendices). The pages in the introductory sections (preface, table of contents etc.) should be numbered with small Roman numerals i, ii, iii, iv, v, etc., one inch from the bottom of the page. All page numbers should stand alone without periods, hyphens or. dashes.

Proofreading

The manuscript should be read critically, searching for inaccurate statements, wrong entries' omissions and inconsistencies. After verifying and locating errors in quotations, footnotes, tables, figures, paragraphing, sentence structure, headings, spellings, style, bibliography, mark the copy to provide the typist with necessary directions for providing a satisfactory transcript.

LAYOUT OF THE RESEARCH REPORT

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 means as to what the research report should contain. A comprehensive layout of the research report should comprise (a) preliminary pages; (b) main text; and(c) the end matter. Let us deal with them separately

(a) Preliminary Pages

In its 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 illustrations so that the decision maker or anybody interested in reading the report can easily locate the required information in the report.

(b) 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 (The questionnaire or interview schedule is usually given in an appendix).If measurements were based on observation, then what instructions were

given to the observers? Regarding the sample used in the study the reader should be told: Who were the subjects? How many were there? How were they selected? All these questions are crucial for estimating the probable limits of 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 limitations, 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 easily 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. The result section of the report should contain statistical summaries and reductions of the data rather than the raw data. All the results should be presented in logical sequence and splitted into readily identifiable sections. All relevant results must find a place in the report. But how one is to decide about what is relevant is the basic question. Quite often guidance comes primarily from the research problem and from the hypotheses if any, with which the study was concerned. But ultimately the researcher must rely on his own judgement in deciding the outline of his report.

(***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 to apply in similar circumstances.

(b) The conditions of the present study which may limit the extent of legitimate generalizations of the inferences drawn from the study.

(c) The relevant questions that still remain unanswered or new questions raised by the study along with suggestions for the kind of research that would provide answers for them.

It is considered a good practice to finish the report with a short conclusion which summarises and recapitulates the main points of the study. The conclusion drawn from the study should be clearly related to the hypotheses that were stated in the introductory section. At the same time, a forecast of the probable future of the subject and an indication of the kind of research which needs to be done in that particular field is useful and desirable.

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

(c) End Matter

At the end of the report, appendices should be enlisted in respect of all technical data such as questionnaires, sample information, mathematical derivations and the like ones. Bibliography of sources consulted should also be given. Index (an alphabetical listing of names, places and topics along with the numbers of the pages in a book or report on which they are mentioned or discussed) should invariably be given at the end of the report. The value of index lies in the fact that it works as a guide to the reader for the contents in the report.

Research Proposal

In a research proposal, the goal is to present the author's plan for the research they intend to conduct. In some cases, part of this goal is to secure funding for said research. In others, it's to have the research approved by the author's supervisor or department so they can move forward with it. In some cases, a research proposal is a required part of a graduate school application. In every one of these circumstances, research proposals follow the same structure.

In a research proposal, the author demonstrates how and why their research is relevant to their field. They demonstrate that the work is necessary for the following:

- Filling a gap in the existing body of research on their subject
- Underscoring existing research on their subject, and/or
- Adding new, original knowledge to the academic community's existing understanding of their subject

A research proposal also demonstrates that the author is capable of conducting this research and contributing to the current state of their field in a meaningful way. To do this, your research proposal needs to discuss your academic background and credentials as well as demonstrate that your proposed ideas have academic merit.

But demonstrating your research's validity and your personal capability to carry it out isn't enough to get your research proposal approved. **Your research proposal also has to cover these things**:

a. The research methodology you plan to use
b. The tools and procedures you will use to collect, analyze, and interpret the data you collect
c. An explanation of how your research fits the budget and other constraints that come with conducting it through your institution, department, or academic program

If you've already read our post on literature reviews, you may be thinking that a research proposal sounds pretty similar. They're more than just similar, though—a literature review is part of a research proposal. It's the section that covers which sources you're using, how you're using them, and why they're relevant. Think of a literature review as a mini-research proposal that fits into your larger, main proposal.

Length of Proposal

Generally, research proposals for bachelor's and master's theses are a few pages long. Research proposals for Ph.D. dissertations and funding requests, are often longer and far more detailed. A research proposal's goal is to clearly outline exactly what your research will entail and accomplish, so including the proposal's word count or page count isn't nearly as important as it is to ensure that all the necessary elements and content are present.

Research proposal structure

A research proposal follows a fairly straightforward structure. In order to achieve the goals described in the previous section, nearly all research proposals include the following sections:

Introduction

Your introduction achieves a few goals:

- Introduces your topic
- States your problem statement and the questions your research aims to answer
- Provides context for your research

In a research proposal, an introduction can be a few paragraphs long. It should be concise, but don't feel like you need to cram all of your information into one paragraph.

In some cases, you need to include an abstract and/or a table of contents in your research proposal. These are included just before the introduction.

Background significance

This is where you explain why your research is necessary and how it relates to established research in your field. Your work might complement existing research, strengthen it, or even challenge it—no matter how your work will "play with" other researchers' work, you need to express it in detail in your research proposal.

This is also the section where you clearly define the existing problems your research will address. By doing this, you're explaining why your work is necessary.

In your background section, you'll also outline how you'll conduct your research. If necessary, note which related questions and issues you won't be covering in your research.

Literature review

In your literature review, you introduce all the sources you plan to use in your research. This includes landmark studies and their data, books, and scholarly articles. A literature review isn't merely a list of sources (that's what your bibliography is for); a literature review delves into the collection of sources you chose and explains how you're using them in your research.

Research design, methods, and schedule

In this section, make sure you cover these aspects:

- The type of research you will do. Are you conducting qualitative or quantitative research? Are you collecting original data or working with data collected by other researchers?
- Whether you're doing experimental, correlational, or descriptive research
- The data you're working with. For example, if you're conducting research in the social sciences, you'll need to describe the population you're studying. You'll also need to cover how you'll select your subjects and how you'll collect data from them.

The tools you'll use to collect data.

- Sampling frame
- Sampling method
- Use of descriptive statistics and inferential statistics
- Will you be running experiments?
- Conducting surveys?
- Observing phenomena?

Note all data collection methods here along with why they're effective methods for your specific research.

Beyond a comprehensive look at your research itself, you'll also need to include:

- Your research timeline
- Your research budget
- Any potential obstacles you foresee and your plan for handling them

Suppositions and implications

Although you can't know your research's results until you've actually done the work, you should be going into the project with a clear idea of how your work will contribute to your field. This section is perhaps the most critical to your research proposal's argument because it expresses exactly why your research is necessary.

In this section, make sure you cover the following:

- Any ways your work can challenge existing theories and assumptions in your field
- Your work will create the foundation for future research
- The practical value of your findings will provide to practitioners, educators, and other academics in your field
- The problems your work can potentially help to fix
- Policies that could be impacted by your findings
- How your findings can be implemented in academia or other settings and how this will improve or otherwise transform these settings

In other words, this section isn't about stating the specific results you expect. Rather, it's where you state how your findings will be valuable.

Conclusion

This is where you wrap it all up. Your conclusion section, just like your conclusion paragraph for an essay, briefly summarizes your research proposal and reinforces your research's stated purpose.

Bibliography

Yes, you need to write a bibliography in addition to your literature review. Unlike your literature review, where you explained the relevance of the sources you chose and in some cases, challenged them, your bibliography simply lists your sources and their authors.

The way you write a citation depends on the style guide you're using. The three most common style guides for academics are MLA, APA, and Chicago, and each has its own particular rules and requirements. Keep in mind that each formatting style has specific guidelines for citing just about any kind of source, including photos, websites, speeches, and YouTube videos.

Sometimes, a full bibliography is not needed. When this is the case, you can include a references list, which is simply a scaled-down list of all the sources you cited in your work. If you're not sure which to write, ask your supervisor.

How to write a research proposal

Research proposals, like all other kinds of academic writing, are written in a formal, objective tone. Keep in mind that being concise is a key component of academic writing; formal does not mean flowery.

Adhere to the structure outlined above. Your reader knows how a research proposal is supposed to read and expects it to fit this template. It's crucial that you present your research proposal in a clear, logical way. Every question the reader has while reading your proposal should be answered by the final section.

Common mistakes to avoid when writing a research proposal

When you're writing a research proposal, avoid these common pitfalls:

Being too wordy

As we said earlier, formal does not mean flowery. In fact, you should aim to keep your writing as brief and to-the-point as possible. The more economically you can express your purpose and goal, the better.

Failing to cite relevant sources

When you're conducting research, you're adding to the existing body of knowledge on the subject you're covering. Your research proposal should reference one or more of the landmark research pieces in your field and connect your work to these works in some way. This doesn't just communicate your work's relevance—it also demonstrates your familiarity with the field.

Focusing too much on minor issues

Including too many questions and issues in your research proposal can detract from your central purpose, weakening the proposal. Save the minor issues for your research paper itself and cover only the major, key issues you aim to tackle in your proposal.

Failing to make a strong argument for your research

This is perhaps the easiest way to undermine your proposal because it's far more subjective than the others. A research proposal is, in essence, a piece of persuasive writing. That means that

although you're presenting your proposal in an objective, academic way, the goal is to get the reader to say "yes" to your work.

This is true in every case, whether your reader is your supervisor, your department head, a graduate school admissions board, a private or government-backed funding provider, or the editor at a journal in which you'd like to publish your work.

Polish your writing into a stellar proposal

When you're asking for approval to conduct research—especially when there's funding involved—you need to be nothing less than 100 percent confident in your proposal. If your research proposal has spelling or grammatical mistakes, an inconsistent or inappropriate tone, or even just awkward phrasing, those will undermine your credibility.