RESEARCH METHODS IN ECONOMICS
(ECON -3052)

DISTANCE EDUCATION MODULE

ARBA MINCH UNIVERSITY COLLEGE OF BUSINESS AND ECONOMICS DEPARTMENT OF ECONOMICS

PREPARED BY: WONDIMAGEGN TAYE (MSC/MPH)

JANUARY, 2020

ARBA MINCH, ETHIOPIA

Chapter One

1. Introduction

Introduction

In this unit, we shall focus on the nature of research. Discussions on how human beings approach knowledge, types of research, and criteria of good research are also included. The learner should note that the ideas in this unit may look fragmented but later enjoyable when you start to see the complete picture of our discussion. You may wonder to see various definitions to (social) research. How do you perceive research in social sciences? Where and when do you think research is conducted? What involves a social research? When do we say a social research is good?

This unit is organized as:

Lesson 1: Definition of Research

Lesson 2: Characteristics of Research

Lesson3: Classification of Research

Lesson 4: Motivation in Research

Lesson 5: Types of Research

Objectives of Unit One

After you work through this chapter, you should be able to:

Understand what research mean

Explain the characteristics of research

Know the classification of research

Understand the motivations of research

List types of research

1.1. Definition of Research

Lesson 1: What is Research?

Description of Lesson 1

This lesson attempts to review some of the definitions by different scholars and identify common elements in their definitions in order to study the nature of research in general and social research in particular.

We are naturally curious to know about ourselves, nature, the environment, etc. Whenever we encounter problems, we try to find their solutions and make decisions accordingly. This may be done using either the arbitrary method or the scientific method. Research uses the latter method as opposed to the former. Dear learner, what do you think is the difference between the arbitrary method and the scientific method? Please write your answer hereunder:

_____

_____

Definition of Research

Well, we shall discuss about these at the latter part of this unit. Let us now try to see the definition of research (by different scholars). Not unfortunately, there is no consensus on the definition of research. However, the definition given by some scholars may give insight on the nature of research. Kerlinger (1973:11) defined research as 'systematic, controlled, empirical and critical investigation of hypothetical propositions about the presumed relations along natural phenomena.' Black and Champion (1976:6) defined [scientific] research as the process of systematically obtaining information through empirical observation that can be used for the systematic development of logically related propositions attempting to establish causal relation among variables. And, finally, Young (1977:30) defined it as "a scientific undertaking which, by means of logical & systematic techniques, aims to: (a) discover new facts or verify and test old facts; (b) analyze their sequences, interrelationships and causal explanations; and (c) develop new scientific tools, concepts and their theories which would facilitate reliable and valid study of human behavior.

Research is a systematic and objective process of collecting, recording, organizing and interpreting data for aid in decision-making. This suggests that:

The research information is neither perceptive nor haphazardly or spontaneously gathered; and

If the information generated is to be accurate, research endeavors must be objective.

You shall, nevertheless, note that all definitions share the fact that research is a diligent enquiry and careful search for new knowledge through systematic, scientific and analytical approach in any branch of knowledge. The learner may ask 'what characterizes research?'

Among other things, research is characterized by:

a systematic and critical investigation of a phenomenon

not a mere complication but a purposive investigation

use of the scientific method (to be discussed later)

objective and logical, applying possible tests to validate the measurement tools and conclusions

analysis based on empirical evidence and/or observable experience

directed towards finding answers to pertinent questions & solutions to

problems emphasizes the development of generalizations, principles or theories

Constant search for additional knowledge is the guiding factor of research which helps to discover new facts. The search of new knowledge also helps to accept, reject or modify existing facts or knowledge. It is a systematic effort to gain knowledge, truth or broad principles in a verifiable and objective way. According to Slessinger and Stevenson, *social research* aims to "extend, collect or verify knowledge; whether that knowledge aids in the construction of theory or in the practice of an art." Research will also help to ascertain existing theories or modifying or rejecting them and build or construct new theories on the basis of research. Young describes social research as "the systematic method of discovering new facts or verifying old facts, their sequences in their relationship, causal explanation and the natural laws which govern them".

Put differently, research is an investigation into relationships and hence is analogous to *detection*. The main aim of research is to search deeply and persistently and penetrate into the

reasons of occurrence and relationships. Thus, research as Young points out does not simply demand a search for facts, but thinking through of the issues and problems involved in every step of the research process. A wise researcher uses studies and principles thereof to unravel the facts intelligibly while craftsmen condemn such process. This is why research is said to originate as a matter of curiosity of human mind to know, to examine and to re-examine things. The main aim of research is to eliminate vague, superstitions and dogmatic idea by reasoned, objective and scientific ideas.

Arbitrary Vs. Scientific Methods

The scientific method refers to the ideas, rules, techniques, and approaches that the scientific community uses. It is the methodology employed that makes a given research scientific or non-scientific. In modern times, the scientific method has become the accepted way to gain knowledge of the [social] world.

There are three basic components of scientific method; viz: the use of empirical evidence, logical reasoning and skeptical attitude.

Use of empirical evidence: The goal of the scientific methods is to facilitate independent verification of scientific observation. Empirical evidence can be verified and determined reliable (replicable or repeatable under similar circumstances).

Logical reasoning (Critical thinking): Logical reasoning allows determination of truth through steps different from emotional and hopeful thinking. Scientists and critical thinkers always use logical reasoning. Often, the use of logical reasoning requires a struggle with the will, because logic sometimes forces one to deny one's emotions and face reality, and this is often painful. But you must not forget that emotions are not evidence, feelings are not facts, and subjective beliefs are not substantive beliefs.

Skepticism: this refers to possessing a skeptical attitude. This is the final key in science and critical thinking which involves constant questioning of your beliefs and conclusions. Good scientists and critical thinkers constantly examine the evidence, arguments, and reasons for their beliefs. Such an attitude is important in the scientific method because it:

- helps avoid self-deception and deception by others;

- encourages scientists to repeatedly examine the truth and reliability of knowledge claims by others;
- promotes the need to test to determine if current beliefs match objective reality as measured by empirical evidence; and
- motivates one to hold beliefs tentatively, and be open to new evidence and rational argument.

To sum up, the scientific method has proven to be the most reliable and successful method of thinking in human history, and it is quite possible to use scientific thinking in other human endeavors.

The learner recalls that we have emphasized research makes use of the scientific method as opposed to the arbitrary method. What is the distinction between these two methods? Where as the arbitrary method of seeking answers consists of imagination, opinion, blind belief or impression, the scientific method is a systematic approach of seeking solutions objectively and precisely based on empirical evidence.

Activity 1.1

Circle the ways that may suffer from individual difference and hence lack objectivity.

A. Opinion
B. Blind Belief
C. Imagination
D. Empirical evidence
E. Logical reasoning
F. Observableexperience

I hope you circled the first three (A, B and C) because these methods result in different findings from one person to the other for the same observation, not to mention. Besides, they are usually vague and inaccurate. The last three (D, E and F), on the other hand, are objective and hence do not bring results that differ among people. Based on this difference, which method do you recommend a research use for a research? Very good, you must have said the *scientific method*.

This is because the scientific method is based on 'some articles of faith' including its reliance on empirical evidence, use of relevant concepts, commitment to objectivity, ethical neutrality, generalization, verifiability and logical reasoning process. The major components this method involves include defining a problem, formulating hypothesis/tentative assertion, data collection & analysis, testing hypothesis & inference, and making claims logically. Dear learner, do you notice research involves the use of theories, facts and methods? Good, let's hold on here and see the relationship between theories, facts and methods.

Check Your Progress

1.1 Distinguish between the scientific method and the arbitrary method.
1.2 List at least two main elements in the definition of research.
1.3 If one deduces on an issue based on her personal belief, do you categorize it to the scientific method or the arbitrary method? Why?

Lesson Summary

Research is a systematic and objective process of collecting, recording, organizing and interpreting data for aid in decision-making.

The scientific method refers to the ideas, rules, techniques, and approaches that the scientific community uses. There are three basic components of scientific method – viz. the use of empirical evidence, logical reasoning and skeptical attitude.

Lesson 2: Characteristics of Research

Description of Lesson 2

This lesson establishes the characteristics of research. It attempts to show different characteristics of research which is common in all scientific researches.

Research is a procedure by which we attempt to find systematically and with the support of demonstrable fact. This will eventually lead to the answer to a question or the resolution of a problem. Research has eight distinct characteristics:

- Research originates with a question or problem in the mind of the researcher;
- Research requires a clear goal;
- Research requires a specific plan or procedure;
- Research usually divides the principal problem into more manageable sub problems;
- Research is guided by the specific research problem, questions or hypothesis;
- Research accepts certain critical assumptions;
- Research requires the collection and interpretation of data in an attempt to resolve the problem that initiated the research; and
- Research is, by its nature, cyclical or more exactly helical (continuous process).

Check Your Progress

2.1 What are the common characteristics of research?

_____

_____

_____

_____

_____

_____

2.2 clearly list out 8 characteristics of research?

_____

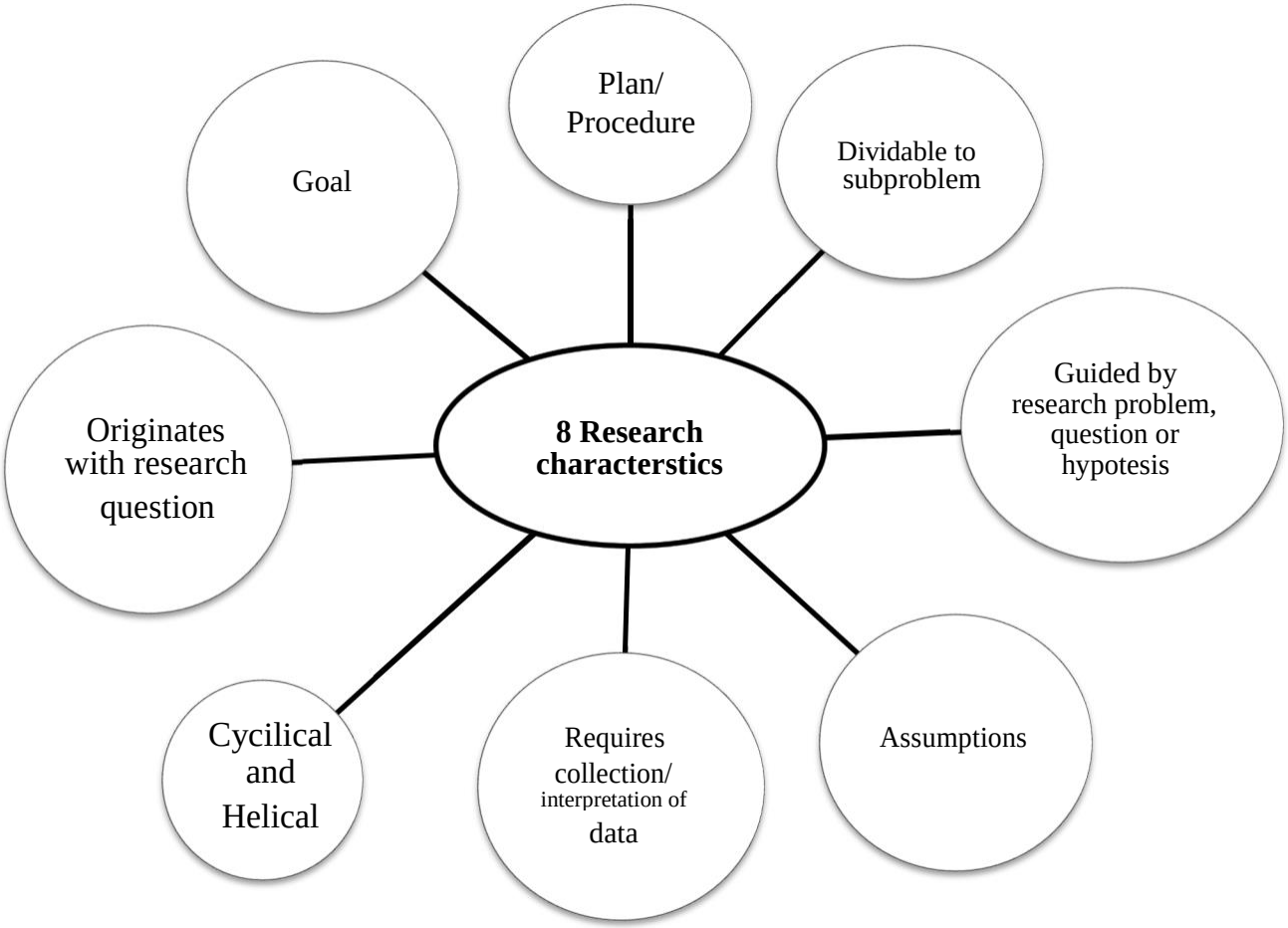
_____

_____

_____

_____

Figure 2: *Clearly lists out the 8 research characteristics.*



Lesson Summary

Research begins with the question in the mind of the researcher;

Research demands the identification of a problem-stated in clear & unambiguous terms; Research requires a plan of action;

Research deals with the general objectives through the appropriate sub-objectives;

Research seeks direction through the appropriate hypothesis;

Research deals with facts and their meanings; and

Lesson 3: Classification of Research

Description of Lesson 3

This lesson establishes the different classification of research. It attempts to show organogram structures of different researches.

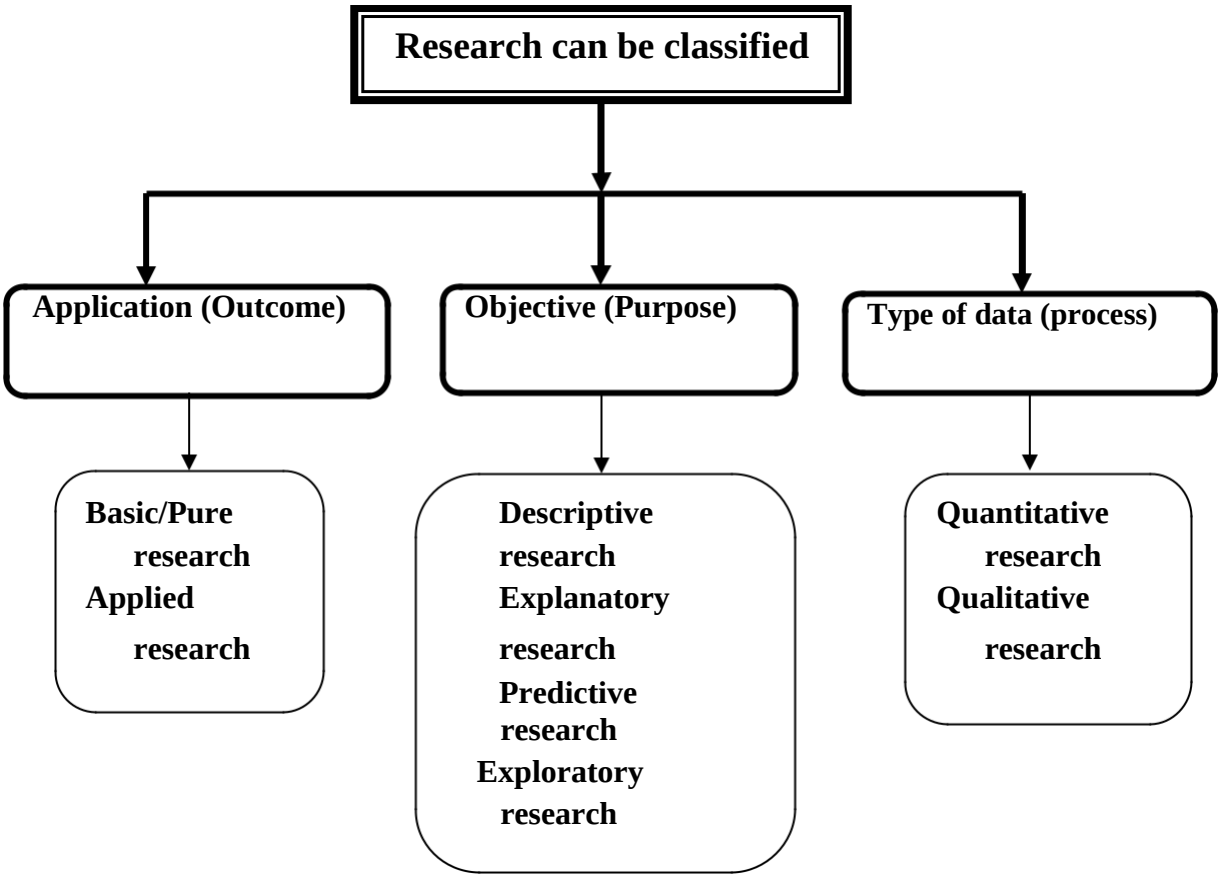


Figure 3: classification and type of research

Progress check:

Dear learner could you briefly show how the classification of research?

_____

_____

_____

Lesson Summary

Classification of research may assume different forms. For our purpose, research can be classified from various perspectives like application of the research study, objectives in undertaking the research, and the type of data sought.

1.1. Motivation in Research

Description of Lesson 4

This section is an attempt to define conditions for a [proposed] research to be feasible. It tries to account for the important factors that determine the feasibility of research. It helps us judge when a research has to be conducted.

Hello students, what measures when a research has to be conducted? The question refers to the *feasibility* of research project. Below are the most common factors to be considered for feasibility of [proposed] research:

Dear learner, what makes people to undertake research? This is a question of fundamental importance. The possible motives for doing research may be either one or more of the following:

- Desire to get a research degree along with its consequential benefits;
- Desire to face the challenge in solving the unsolved problems, i.e., concern over practical problems initiates research;
- Desire to get intellectual joy of doing some creative work;
- Desire to be of service to society;
- Desire to get respectability.

However, this is not an exhaustive list of factors motivating people to undertake research studies. Many more factors such as directives of government, employment conditions, curiosity about new things, desire to understand causal relationships, social thinking and awakening, and the like may as well motivate (or at times compel) people to perform research operations.

Progress checking:

What are the motivational factors for conduction a good research?

4.1 Criteria of Good Research

Whatever may be the types of research works and studies, one thing that is important is that they all meet on the common ground of scientific method employed by them. One expects scientific research to satisfy the following criteria:

- The purpose of the research should be clearly defined and common concepts be used. The
- research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.
- The procedural design of the research should be carefully planned to yield results that are as objective as possible.
- The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.
- The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be appropriate. The validity and reliability of the data should be checked carefully.
- Conclusions should be confined to those justified by the data of the research and limited to those for which the data provide an adequate basis.
- Greater confidence in research is warranted if the researcher is experienced, has a good reputation in research and is a person of integrity.

In other words, we can state the qualities of a good research as under:

1. *Good research is systematic:* It means that research is structured with specified steps to be taken in a specified sequence in accordance with the well defined set of rules. Systematic characteristic of the research does not rule out creative thinking but it certainly does reject the use of guessing and intuition in arriving at conclusions.
2. *Good research is logical:* This implies that research is guided by the rules of logical reasoning and the logical process of induction and deduction are of great value in carrying out research. Induction is the process of reasoning from a part to the whole

whereas deduction is the process of reasoning from some premise to a conclusion which follows from that very premise. In fact, logical reasoning makes research more meaningful in the context of decision making.

3. *Good research is empirical:* It implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.
4. *Good research is replicable:* This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

Lesson Summary:

We can state the qualities of a good research as under:

1. Good research is systematic:
2. Good research is logical:
3. Good research is empirical
4. Good research is replicable:

Lesson 5: Types of Research in Economics

Description of Lesson 5

Lesson 6 tries to briefly see the types of research with a focus on economic research. It is so brief it doesn't treat the details of each sub-division in the classification.

Classification of research may assume different forms. For our purpose, research can be classified from various perspectives like application of the research study, objectives in undertaking the research, and the type of information sought, etc.

However, there are two broad/general categories of research; viz., *basic* research and *applied* research. While basic research (also known as pure research) is usually conducted to develop principles, theories & models and is used to expand the limits of knowledge (used to evaluate theories and models), applied research (also known as action research, development research, or empirical research) is conducted when decisions must be made about a specific real life problem.

Basic research is usually too challenging as it requires the accumulation of knowledge and is mostly used for elder scientists & professors. It cannot be used for solving pragmatic (concerned with actual circumstances rather than general theories) issues. Applied research, on the other hand, is conducted when there is a desire to formulate developmental strategies and policies; when you need action to combat the existing or possible problems. In a nutshell, therefore, basic research tries to extend the knowledge in a given field and applied research tries to solve problems.

The two types of research, nevertheless, have similarities in that the procedures & techniques each utilizes do not differ substantially. Both employ the *scientific method* to answer the research question. Scientific method here refers to techniques and procedures that help the researchers know & understand about the problem of the research.

In basic research, scientific method is used to test prior concepts and hypotheses and then making inferences and conclusions about the phenomenon. In applied research scientific method is used to assure objectivity in gathering facts and test creative ideas for alternative decision strategies. Applied research is very common among research in Economics and other disciplines too.

Types of Applied Research:

Applied research can be classified on the basis of various variables: technique (method), purpose (function), the role of time and so on.

a) Technique or method

Applied research by their *technique* (or method) can be further classified as experimental, survey and observation.

a.i) Experimental Research (Controlled Experiment)

A researcher would at first create two matched groups as similar as possible to test his/her hypothesis. One group is considered as the experimental group, while the other would be designated as the control group. The experimental group is the one on which the tests are conducted. The role of the control group is very critical since it serves to isolate the

experimental stimulus as the single source of change (as the two are identical in all other regards) in the experimental group.

It is extremely important to make sure the two groups are identical in all other regards than the experiment of interest in order to deduce that changes following the experiment on the experimental group and hence population of interest are due to the stimulus from the experiment.

a.ii) Survey Method

It represents a comprehensive description and explanation of many complements of a given social system. The researcher seeks to collect and examine as much data as possible regarding the subject of her study.

a.iii) Observation

The researcher seeks to become a member of the social event or group under study. He may or may not reveal his research role. By impressing himself in the actual social events in progress, the researcher will be in a position to obtain a far greater depth of knowledge. This could be done with or without the knowledge of the target population.

b) Purpose or function

Another basis of classification of applied (social) research is its purpose or function. On the basis of *purpose/function* applied social research can be classified as exploratory, descriptive and causal.

b.i) Exploratory Research

It is an initial research conducted to clarify and define the nature of a problem – ambiguous problem. It is conducted to gain familiarity with a phenomenon or to achieve new insights into it. The researcher aims at formulating more precise questions that future research can answer, and discovering new insights. It may be the first stage in the sequence of studies. Its result may not be used as conclusive evidence to determine a particular course of action. It is usually conducted with the assumption that subsequent research will be required to provide

conclusive evidence. It uses qualitative data and tends to be more open using range of evidence and discovering new issues.

Exploratory research may employ the techniques of:

b. i.i) Experience Survey

It is an exploratory research technique in which individuals who have knowledge about a particular research problem are surveyed or asked. This is about the experience of people who are experts or well-familiar to a given issue.

b.i.ii) Secondary Data Analysis

This refers to literature review or literature survey and deals with preliminary review of data collected for another purpose to clarify issues in the early state of a research effort. This technique does not make any attempt to collect primary data.

It merely assesses secondary data and usually considered as a stepping stone for other techniques of analysis and not a research technique per se. It is not a suggested way for student research like yours as it merely compares and contrasts already done research by other people and hence fails to show your contribution.

b.i.iii) Case Study

This involves intensive investigation of one or a few situations similar to the researcher's problem situations such that the case is fairly representative or related to the population of interest.

b.i.iv) Pilot Survey

It is a brief preliminary survey often using a small, convenience sample conducted to test the survey instruments and data collection method, before the project details are finalized and the formal survey conducted. Most, but not certainly all, exploratory research provides qualitative information or clue information; we cannot say this much or this percentage.

Generally the purpose of exploratory research is to clearly and precisely define the problem. Generally speaking, exploratory researches have three integrated purposes:

to diagnose the situation i.e., to identify the nature of the problem

to screen the alternatives

to discover new ideas

b.ii) Descriptive Research

The major purpose of descriptive research is to describe the characteristics of a population or phenomenon. It attempts to give answer to questions like who, what, when, where and how, but not why. The answer to questions to why is not given by descriptive research. Unlike exploratory research, descriptive studies are based on some previous understanding of the nature of the problem. Its purpose is to describe the state of affairs, as it exists at present. The researcher has no control over the variables, but only reports what has happened or what is happening. This is relevant in positive economics and less important in normative economics.

b.iii) Causal Research

Such research aims at identifying the cause-effect relationship between variables where the research problem has already been narrowly defined. For example, the relationship between fertilizers application and land productivity, the impact of packaging and advertising on the volume of sales of a product, the effect of training on labor productivity.

c) Other Criteria

On the basis of other criteria research is also classified into:

c.i) Descriptive Vs. Analytical Research

This classification considers the degree of analysis of a research and whether or not a research is interested in critical analysis of the issue. In analytical research, the researcher has to use facts or information already available, and analyze these to make a critical evaluation of the material while descriptive analysis tries to deal with the state of affairs or a phenomena or the population of interest.

c.ii) Active Research vs. Evaluation Research

Active research treats knowledge as a form of power and abolishes the line between research and social action where as evaluation research makes social impact analysis that attempts to estimate the likely consequences of planned change. It is mainly used for planning and

making choice among alternative policies. Social impact analysis often includes cost benefit analysis; hence, evaluation research is expected to address the question, '*Did it work?*' For example, questions like 'Has the bonus system increase productivity of workers?'

c.iii) Quantitative Vs Qualitative Research

Based on the [non-]measurability of the variables of interest in a given study, research may be grouped as quantitative or qualitative. Quantitative research is based on the measurement of quantitative figures, quantities or amounts; i.e., it is applicable to phenomenon that can be expressed numerically. On the other hand, qualitative research is concerned with qualitative phenomenon; i.e., a phenomenon relating to or involving quality or kind. For example, questions like 'why do people think or do certain things?' All kinds of motivation, attitude research or opinion research are typical qualitative research studies. It is particularly important in behavioral sciences where the underlying motive is to discover the underlying motive of human behavior.

c.iv) Conceptual Vs Empirical research

Based on the motive of research as to whether it is related to theories & disciplines or real world experience, it is grouped as either conceptual or empirical. Conceptual research is related to some abstract or theory and is conducted by philosophers to develop new ideas or reinterpret the existing ones. Empirical research, on the other hand, relies on experience or observation alone, without due regard to the system or theory – it is data-based research or also sometimes called experimental research.

c.v) Cross-section Vs Longitudinal Research

Based on the role of time, research may be grouped as cross-sectional or longitudinal. This is particularly valid for quantitative applied research. While the former deals with the problem at a single point of time, the latter considers the issue at multiple points of time.

Cross Sectional research takes a snap-shot approach to the social world. It is the simpler and less costly research approach as compared to longitudinal research. However, it cannot capture social process or changes.

Longitudinal research, on the other hand, examines feature of people or other units at more than one time – extended period of time. It is usually more complex and costly than cross-sectional research but usually more powerful especially with respect to social changes. Based on the way it treats the subject of interest, longitudinal research may assume a time-series, panel or cohort analysis. Broadly, longitudinal research may be sub-grouped into time-series research, panel study or cohort analysis based on the role of time in the unit of analysis. A brief introduction of each of these is made in what follows. Hence, I advise you to read more for further information.

c.v.i) Time-Series Research

A time series research is a longitudinal study on a group of people or other units across multiple periods. For example, a time series data on exports of coffee in Ethiopia may cover the period 1970 up to 2007. If the data set is annual, we would have 38 observations only.

c.v.ii) Panel Study

In a panel study, researchers observe exactly the same people, group or organ across time periods but each time each time taking a snap-shot approach. In a panel study, the focus is on individuals or *who*? For example, interviewing exactly the same people in 1991, 1995, 1999, 2003, etc and observe the change is an example of panel data set. It, however, is usually costly but may yield better outcome.

c.v.iii) Cohort Analysis

Cohort analysis is similar to panel study. But, in a cohort analysis, rather than observing exactly the same people, a category of people who share similar life experience in a specified period are studied. Hence, the focus is on groups of individuals, and not on specific individuals or who? This is less costly than panel study.

Lesson Summary

Research may be classified based on so many variables: nature of data, role of time, purpose of the study, etc.

Review Questions

1 What is research?
2 List some of the characteristics of research (in social science).
3 Distinguish between the scientific method and the arbitrary method. State some of the ways people gain knowledge and group them in either category.
4 What are the motivational factors for doing good research
5 What is the (ultimate) goal of research?
6 A distance learner wants your help to evaluate whether his proposed topic on ‘food security challenges in Africa: Evidence from Ethiopia’ is researchable. What criteria may make this topic not researchable for a student researcher (for term-paper)?
7 What is the difference between applied and basic research?
8 Using the techniques employed, list the different types of applied research?
9 What are the criteria of good research?

UNIT TWO:

2. FORMULATION OF RSEARCH PROBLEM AND HYPOTHESIS

Dear learners, know that research is a process! This process may not be distinct between different research projects. Yet, all research projects share certain major steps in common. This unit discusses some of these steps. Can you guess the major steps in the space provided below and compare your results in the discussions below please?

_____

_____

_____

This unit is organized as:

Objectives of Unit Two

After you study this unit, you should be able to:

Appreciate the various common steps in social research;

Distinguish between the different phases of the research process;

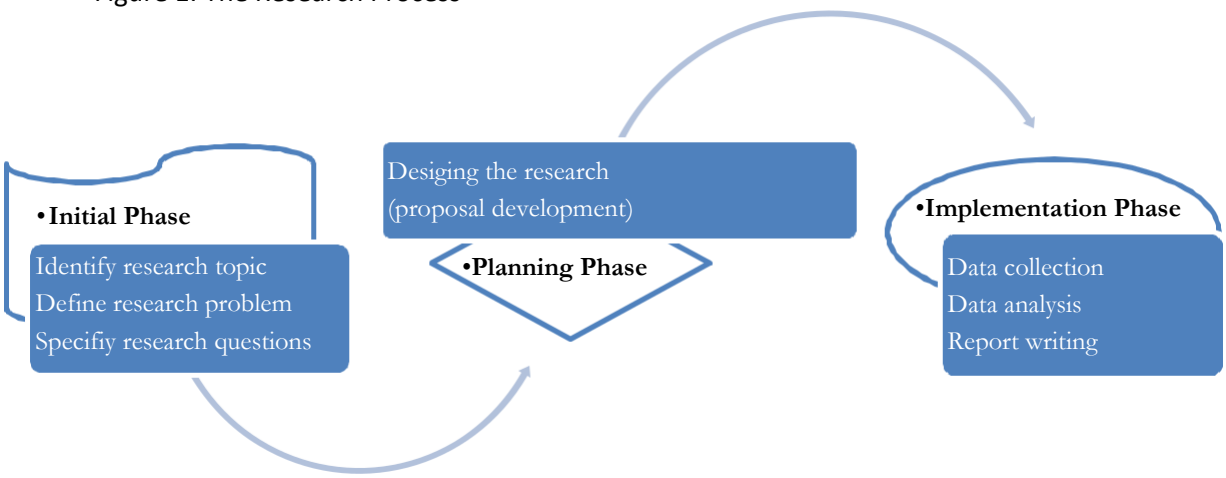
and Elucidate the importance of each step in the research process;

Explain the steps in the initial phase of the research process

Introduction

The research process consists of a series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps. However, not only these steps may overlap one another, but also may not always follow a single sequence. Yet, the following flow chart will illustrate the general research process:

Figure 1. The Research Process



More specifically, an eight steps model may be developed to illustrate the major steps in the research process that are the concern of the discussions to come. The steps are grouped in each phase of the research process and then discussed in a nutshell in this and subsequent units later.

Dear learners, how do you think are these phases complementing each other? Do you think they are distinct? Please write your views on the space provided below:

_____

_____

_____

Very good, I hope, you realized that sometimes the first step determines the last step to be undertaken. If subsequent procedures have not been taken into account in the early stages, serious difficulties may arise that may even prevent the completion of the study. Nevertheless, the above phases are so complementary; it is usually difficult to put distinct cut-off points for

each activity. For instance, not only for selecting your topic but also to define your problem, you may need to review the literature; you need to consider issues related to data when you select your research topic; you may need to evaluate the relevance of potential hypothesis to address your research questions; etc. That is, the various steps involved are not mutually exclusive; nor are they separate and distinct. They hardly necessarily follow each other in any specific order.

Despite this nature of activities in a research, the above sequence of activities may help us as a general guide. Let's see some of the initial steps and the planning and implementation phases shall be considered in separate chapters later. The approach in this module is we treat the initial phase in this unit and the other phases discussed in the subsequent units and only briefly touched in this unit.

The Initial Phase

This phase involves:

1st Defining the research problem

2nd Identifying topic – Working title

3rd Review of literature – Theoretical and/or empirical

4th Formulating hypothesis, if any

The Planning Phase

5th Develop the research design – research proposal [Unit 3]

Implementation Phase

6th Collect Data – primary or secondary [Unit 4]

7th Data Processing [Unit 5]

8th Report Writing [Unit 5]

Lesson 1: Topic Identification

Description of Lesson 1

This lesson sheds light on how to select a working title for research. It tires also to appreciate the factors one has to consider when selecting a research topic.

A topic may be defined as an area within a given field of study that you would like to explore in your research. It is the subject discussed or to be discussed. To do any research, a topic or a research problem must be identified. So, the first step in the research process is the choice of researchable problem for investigation. Research usually starts with a feeling of difficulty. It takes place when there is a problematic situation and a need to solve the problem. The identification of a research problem is the most difficult but the most important phase of the entire research process.

Potential Sources of Research Topic

Usually a research topic springs from the researcher's mind. Nevertheless, some important sources may be suggested:

- Own personal/professional experience of the researcher of research team is the most important source. In addition, contacts and discussions with research-oriented people, attending conferences, seminars and listening to learned speakers, media broadcast, etc may be good sources.
- Inferences from theory and professional literature: research reports, bibliographies of books and articles, periodicals, research abstracts and research guides may suggest areas for research.
- Technological and social changes: New developments bring forth new development challenges for research. New innovations and changes need to be carefully evaluated through the research process.

Dear learner, you may be surprised as to how one arrives at a working title. There is no consensus on the way to identify a research topic; but some guidelines are suggested by scholars. It is easy to try one! Follow the steps given below and check if your topic is reasonably acceptable:

a) Identify the general area of interest (E.g. agriculture, industry, trade etc) based on: your interest

your expertise

a degree to be acquired

potential funding institution

a grant to compete for

organization you work with/for

Please write the broad area of your interest _____

b) Draw up a shortlist of possible topics that have emerged from your reading or from your own experience that look potentially interesting. In listing topics for research, you may rely on the following:

Personal experience: based on the situation that happens to you or those you know

Curiosity based on situation in the media: Ex: watching TV program on street children may dictate a topic

The state of knowledge in the field: your knowledge of the field suggests a topic; basic research is driven by new research findings and theories that push at the frontiers of knowledge

Solving a problem: appropriate topics for applied research are those that often begin with a problem that needs a solution

Personal values: some people are highly committed to a set of religious, political or social values

Re-examining your course materials

When you are selecting broad area of interest and list potential topics, you must carefully revisit your professional interest and experience.

Please list three or more of your potential topics (from 5-20 words)

i) _____

ii) _____

iii) _____

c) Discuss your list with your tutor and anyone who has expertise in this area and might be able to offer advice or insight considering the researchability of the topics, and

Dear learner, can you imagine what determines the researchability of your potential topics?

_____

_____

_____

d) Select a topic from the list that interests you and that you are fairly certain can be investigated with the resource at your disposal.
e) Formulate the research question: Usually the topics chosen are broad at first and lacks focus. Therefore, it should be narrowed properly and focused into specific questions.
f) The next step is to narrow down the topic. Techniques that help us to narrow a research topic into research question:

Examining past and present literature

Discussing the idea with other people

Applying the topic to specific context (specific time period, society, geographical unit, sub-groups, categories, etc)

Defining the aim of the outcome

[Student] research topics are usually written in 4-20 words and may assume the following forms: indicative, hanging and question type.

Indicative titles are by far the most common type: they state the subject of the proposal rather than expected outcomes. An example is 'The role of agricultural credit in alleviating poverty in a low-potential area of Ethiopia'. If a title written in this form becomes too long, however, consider rewriting it using the hanging format.

Hanging titles have two parts: a general first part followed by a more specific second part separated by a colon (:) or hyphen (-). This title type may be useful in rewording an otherwise long, clumsy and complicated indicative title. An example of a hinging title is 'Alleviation of poverty in low-potential area of Ethiopia: the impact of agricultural credit'. Such a title format may help a researcher emphasize on the technique of the study, the case under scrutiny, the geographical coverage, the time period under investigation or any other emphasis she would like to make.

Question-type titles are used less commonly than indicative and hinging titles. However, they are acceptable where it is possible to use few words – say less than 15. Example could be 'Does agricultural credit alleviate poverty in low-potential areas of Ethiopia?'

Check Your Progress

2.1 Why is a topic identified at the initial stage of your research process considered only a working title?

2.2 List at least two potential sources of research topic for a (distance) learner? What exactly is the source of the research topic you identified above for your term-paper?

Well, I hope you now are able to arrive at a certain working title. Good, note, however, that the title you have just arrived at is not final. It is subject to modifications and hence is termed as working title.

Now, write your working title in the space given below please:

_____

_____

Rewrite your topic in other topic types, compare the topics and make your choice of a working title once again:

_____

_____

Write your final working title in the space provided below please:

_____

_____

Lesson Summary

Topic identification is an important first step in the research process that effectively determines almost all the remaining steps of the process. Hence, a researcher needs to give due attention to identification of research topic.

A [student] research may be written in 4-20 words commonly in three forms: indicative, hanging and question types.

Lesson 2: Defining Research Problem

Description of Lesson 2
Lesson 2 deals with the need to define a research problem, how it is defined and the factors to be considered in problem definition.

Once a topic is identified, the next step is to clearly define the problem in unambiguous manner, i.e., explaining the problem in a clear and unambiguous term. What is a research problem? It refers to some difficulty, which a researcher experiences in the context of either a theoretical or a practical situation and wants to provide a solution for. Generally, a research problem exists when there is/are:

Certain part of society who has some difficulty or problem

Some objectives to be attained

Alternative measures for a problem

Unanswered questions

Some environment to which the difficulty pertains

Points to be observed when selecting a problem:

A subject that is overdone should be avoided;

Controversial subjects are not recommended to be studied (at least, for beginners);

Too narrow or too broad or vague problems should not be selected;

The subject selected should be familiar and feasible – It should be within the researchers reach/scope;

The importance of the subject shall be considered;

The qualifications and training of the researcher (team) should be taken into account; and

The cost involved and the time factor should be affordable given the resource and/or budget of the research.

A problem statement describes the context for the study and it also identifies the general analysis approach. A clearly defined is a problem half-solved. A problem statement should be presented with in a context and that context should be provided and briefly explained, including a discussion of the conceptual or theoretical framework in which it is embedded. You are expected to clearly and succinctly identify and explain the problem within the framework of the theory or line of enquiry that undergirds the study.

Effective problem statements answer the question why does this research need to be conducted? If a researcher is unable to answer this question clearly and succinctly, then it will come off as ambiguous and diffuse.

If the problem is well defined it helps the researcher to:

Discriminate relevant data from irrelevant ones

To lay boundaries within which to study

To be on the right track

Recommended steps in defining a research problem:

i) Statement of the problem in a general way;
ii) Understand the nature of the problem more clearly i.e., by discussing it with other more acquainted or experienced people;
iii) Survey of the available literature; this would lead to review both empirical and theoretical literature. This enables the researcher to:

Find out what data are available for operational purpose,

Find out if there is a gap in theories,

Find out whether the existing theory is applicable to the problem under study.

iv) Developing the ideas through discussion (also called experience survey): It produces useful information, and

It sharpens his focus of attentions on a specific aspects of the study

v) Rephrasing the research problem: This helps to put the research problem in specific terms, as possible so that it may become operationally viable and may help in developing new hypothesis; For example, a research problem in broad terms may look as 'why is productivity in Ethiopia lower than in other countries?' This has many ambiguities as to which set of productivity: land, capital, labor…? With what industries is it related? With what period of time the productivity is being considered? Hence, such presentations need rephrasing and further formulation.

Formulation is translating and transforming the problem into a scientifically researchable question. It is concerned with specifying exactly what the research problem is and why it is studied. Additional considerations necessary when describing a problem:

Technical terms or phrases with special meanings used in the statement of the problem should be clearly defined

Basic assumptions or postulates relating to the research problem should be clearly stated

The suitability of the time period and the source of data available must be considered in defining the problem

The scope of the investigation within which the problem is to be studied must be mentioned explicitly.

At the cost of repetition, a problem well-defined is half-solved. The problem to be investigated must be defined unambiguously for not only that it helps to discriminate relevant data from irrelevant ones and techniques of analysis, but also enables the researcher to be on the right track. Defining a research problem properly is, thus, a prerequisite for any study and is a step of highest importance in the research process. Formulation of a problem is often more essential than its solution.

Lesson Summary

Defining the problem is the most crucial step of the research process.

The statement of the problem should answer the question why does the research need to be conducted [the scientific justification of the study]? It is concerned with specifying exactly what the research problem is and why it is studied

The problem statement should make a convincing argument that there exists not sufficient knowledge to explain the problem and its possible alternative solutions, or for the need to test what is known and taken as fact.

Activity 2.1

What is a research problem? Elucidate how a research problem is defined. List the importance of defining a research problem.

_____

_____

_____

Check Your Progress

2.3 Why is it the first most important task for a researcher to define the problem statement?

2.4 Illustrate how one can make a convincing argument as the need for a research. In your problem statement, how do you think can you convince readers, evaluators or supervisors on the need of the study?

Lesson 3: Objectives

Description of Lesson 3
This lesson attempts to briefly see what characterizes research objectives, how it is formulated and how it is related to the research problem.

The next step after the statement of the problem is setting up the objective(s) – the goals that you set out to attain in your study. It informs the reader what the researchers wants to achieve through the study. It must be stated clearly and specifically under the headings of major and specific objectives.

The major objective is an overall statement of the thrust of the study – statement of the main associations and relationships that you seek to discover or establish. Where as the specific objectives are about certain aspects of the topic you want to investigate with in the main framework of the study. The specific objectives, therefore, should emanate from the general objective and may be stated in action-oriented words such as 'to determine', 'to find out', 'to assess', 'to measure', 'to explore', etc. The way these objectives are worded determines how the research is classified (descriptive, experimental, exploratory, etc). Put differently, the wording of your objective determines the type of research design you need to adopt to achieve them.

It is here that the research can exactly determine what data to be collected. Thus:

Set an overall/general objective of the research
Split the overall objectives into smaller segments known as specific/sub-objectives

State the sub-objectives clearly and precisely/concisely in turn think of the overall objectives through those segments and conquer than in turn rather a frontal attack, upon the entire problem.

Make sure that your objectives are SMART (specific, measurable, achievable, relevant and time-bounded). Even the general objectives of a (student) research shall not be very general as such. They need to be specific enough to be addressed in your research. Besides, one must be able to measure (using available tools) whether the stated objectives (general and specific) are achieved or not. For this, all the stated objectives must be achievable given the resources and expertise of the researcher. Moreover, the objectives of any study shall have a direct connection to the

research problem to quality the criteria of relevance. Finally, they have to be achievable in the available time for the research and hence are supposed to be time-bounded.

Lesson Summary

Objectives may be general and specific – the latter drives from the former

Objectives have to be SMART (specific, measurable, achievable, relevant and time-bounded)

Objectives serve as bases of evaluation for the success of a given research and hence have to be carefully stated.

Check Your Progress

2.5 Distinguish between general and specific objectives.

2.6 How do we measure the quality of objectives of a study?

Activity 2.2

If a research topic reads Farm Inputs Technology Adoption in Gamo Gofa Zone, write two or more SMART objectives for this study. Can you please your objectives as general and specific?

_____

_____

_____

Lesson 4: Literature Review – Theoretical and Empirical

Description of Lesson 4

Lesson 4 tries to show what we mean by literature survey and how and when it is done. It discusses the purpose and sources of literature review.

This is an important exercise probably throughout the research as it helps identify potential working title, defining researchable problem, hypotheses formulation, report writing and so on. This exercise may come here and there as deemed necessary. For certain studies having exploratory nature, survey of literature may be less important. Survey of literature (also called document review or literature review) has to be directly related to the problem to be investigated. The materials to be reviewed may include academic journals, conference proceedings, dissertations, and government report, and policy report, publications of international organizations, books, newsletters and sometimes unpublished sources on the pipeline of publishing. The best place for the survey is LIBRARY.

Literature review is of immense importance as it helps to know the data collection rule, to select appropriate method of data collection and to know whether it is up to date. Broadly, it

i) helps to gain a background knowledge of the research topic; to know whether the problem has been studied thoroughly or not and see if the problem is new; get and summarize what is known in the area and guides identification of knowledge gap. Hence it helps to avoid an intentional duplication of well investigated problems;
ii) helps to know about the recommendations of the previous research for further study which they have listed in the studies.
iii) shows the path of prior research and how contemporary research is linked and its relationship to your study; helps to identify the concepts relating to it, potential relationships between them and to formulate research hypotheses,
iv) to identify data sources, appropriate methodology, research design, methods of measuring concepts, tools of data collection and techniques of analysis used by other (previous) researchers as well as other supportive ideas for discussion;
v) to learn how others structured their reports.

In general, literature survey is used to familiarize oneself with the issue and establish credibility, to learn from others and stimulate new ideas. A preliminary literature survey is important in order to find out: what other researchers have to say about the topic, ensure that no one else has already exhausted the questions that you aim to examine, see how the topic has been discussed within the competing contemporary theoretical framework, and make sure there is enough material available for you to work with productively. This step is important because a 'researcher' who is not fully conversant with what has gone and is going has little chance of making a worthwhile contribution. One must keep herself updated in the field of study.

Literature consists of published and unpublished materials. It should be noted, however, that it is not a mere reading for reading sake; rather, it is purposive, selective, and critical. Common sources of literature include books, journals, reports, dissertation and theses, newspapers, microforms, documentations, bibliographies, governmental reports, research abstracts, book reviews, etc. Usually one source of literature guides to other sources.

Dear learners, I would advise you to plan your literature survey in the following fashion:

decide what information is useful and what is not,
determine how you should record what is gathered,
set up an orderly recording or note-taking system,
It is advisable to take note of the necessary details for your reference/bibliography after you take note of the necessary information from the materials you referred to. This basically involves taking note of the name of the author(s), year of publication, title of the material, place of publication, volume and number, if any. If you wish to consult secondary sources, you may have a look at encyclopedias, research reviews, books, journals and others.

When going over the relevant literature it should never be just a list of papers where you summarize each paper with one paragraph per paper – no this is not critical enough to quality a literature review. This kind of summarizing papers may serve for your own use, to keep track of what was done where, but the idea of a literature survey is that you should identify the issues relevant for the problem. You can be writing: ...x (1980) and y (1985) argues that the microfinance in rural Ethiopia contributes towards poverty reduction, but z (1990) and q (1995) argued for no contribution of these institutions.

I must warn you not to write as: x (1980) says...., y (1985) says..., z (1990) says... and q (1995) says.... Besides, being extremely boring, this method of writing does not serve to sum up the economic problem you are trying to analyze in a good way, and it does not show your ability to take information and aggregate it.

Lesson Summary

Literature review helps in almost every decision you are making throughout the research process: helps identify researchable topic, discriminate relevant data from irrelevant ones, formulate appropriate hypothesis, select techniques of data collection, choose hypothesis tests techniques, write research reports appropriately, etc.

When you review the literature, try to be critical and don't merely state what others did so far; rather, compare and contrast them to identify a gap and potentialities of your research to fill the gap.

Check Your Progress

2.7 What is literature review?

2.8 What are the major sources of literature for your review?

Lesson 5: Formulate Hypothesis

Description of Lesson 5

This lesson discusses when and why we need to formulate research hypotheses. It also attempts to show how it is formulated and its types briefly.

Another step that is often useful in social research is hypothesis formulation, if any. Hypothesis is a tentative intelligent guess postulating form the purpose of directing the researcher towards the solution of problem. Alternatively, a hypothesis may be defined as a hunch, assumption, suspicion, assertion or an idea about a phenomenon, relationship or situation, the reality or truth of which you do not know. Or it is a statement which predicts the relationship between two or more variables. It is necessary link between theory and investigation, usually stated after an extensive survey of the literature.

Hypotheses bring clarity, specificity and focus to a research problem, but are not a must for every study. That is, you can conduct a valid investigation without constructing a single formal hypothesis. On the other hand, with the context of the study, one may formulate as many hypotheses as she considers appropriate.

The importance of research hypotheses lies in their ability to bring direction, specificity and focus to a study. It tells a researcher what specific data to collect and thereby provide greater focus. For most social research, hypotheses serve as:

an indicator of the type of data needed

direction to the research objectives

direction to data gathering techniques (procedures)

facilitator of the extension of knowledge because it links theory and investigation.

a guide to the researcher by delimiting the area of research and keep her on the right track.

The initial hypothesis is called working hypothesis because it is subject to modifications as the investigation proceeds.

Example: Is age at marriage associated with divorce? Variables age at marriage and divorce! Which one is independent (cause) variable? - Age at marriage! What is the direction of the relationship?

The hypothesis could be the lower the age at the time of marriage, the greater the chance that the marriage will end in divorce

Note:

A hypothesis should be simple, specific and conceptually clear. There is no place for ambiguity in its construction as it may make its verification almost impossible;
A hypothesis should be uni-dimensional – it should test only one relationship at a time; Several hypotheses can be developed for one research question;
All hypotheses should be tested statistically; otherwise, don't include a hypothesis in your study;
A hypothesis should be operationalisable – it can be expressed in terms that can be measured. If it cannot be measured, it cannot be tested and, hence, no conclusions can be drawn;
Hypothesis should be very specific and limited to the piece of research at hand because it has to be tested. As it has to be related to the body of knowledge, survey of literature is of immense importance in formulating hypotheses.

Usually it is derived form the deductive logic of the objectives under investigation. A one-to-one correlation might exist between the objectives and their corresponding hypothesis. Thus there could be as many research hypotheses as there are objectives. There are also objectives that do not need formulation of hypothesis. It is usually important when a researcher is interested about causal relations among variables.

The main approach to develop a working hypothesis includes:

i) Discussions with colleagues and experts about the problem, its origin and objective in seeking a solution
ii) Examination of data and records if available concerning the problem for possible trends, peculiarities and clues.
iii) Review of similar studies in the area of the study – literature survey.
iv) Exploratory personal investigation which involves original field work, interviews on a limited scale, with interested parties and individuals with a view to secure greater insight into the practical aspects of the problems.
v) Main problems faced commonly in formulating working hypothesis:
vi) Lack of clear theoretical framework
vii) Lack of ability to utilize the theoretical framework logically

viii) Failure to be acquainted with available research techniques so as to be able to phrase the hypothesis properly.

Note that a hypothesis can appear in your report on either:

i) Deductive form: makes positive statement about the outcome of the study. It can come in the form of directional or non-directional way:
Directional: stipulates the direction of the expected result.

Ex: In AMU, the academic status of those who studied new curriculum is significantly higher than those who studied old curriculum

Non-directional: Does not specify the direction of expected difference or relationship.

Ex: In AMU, there is a difference in academic performance of students who studied old curriculum and new curriculum.

ii) Null Form: Makes a statement that states no relationship.

Ex: In AMU, there is no significant difference in academic status of students who studied new curriculum and old curriculum.

iii) Question form: Put the hypothesis in question form

Example: Does the change in curriculum at high schools affect the academic status of students in Arba Minch University?

iv) Alternative hypothesis: This signifies a statement written opposite to the null form, i.e., when the final decision is made at a given significance level if the null hypothesis is to rejected alternative hypothesis gets accepted, the alternative hypothesis too on equal footing along with the null hypothesis provides direction to the research.

Ex: In AMU, there is significant difference in academic status of students who studied new curriculum and old curriculum.

Check Your Progress

2.9 What is a hypothesis? What does it help a researcher to formulate relevant hypothesis for her researcher?

Lesson Summary

A hypothesis is a tentative scholarly prediction of the outcomes of a research.

A hypothesis helps make a research more focused and clearly defines the data requirements.

Review Exercises

1. If the null hypothesis is 'Tourist inflow in Ethiopia is positively affected by urbanization', then the alternative hypothesis may be

   _____

   _____

2. Which part of the research proposal dominates in terms of importance in signifying why the research has to be conducted?
3. What exercise usually helps identify a gap in knowledge and hence potential topic(s) in research?
4. What measures whether or not objectives of a give research are good?
5. What types of studies employ hypothesis formulation and testing?
6. When do we say that a topic is overdone?
7. What factors need your research proposal to consider?

UNIT THREE:

3. THE RESEARCH DESIGN

Well done learners! We are done with the initial phase of the research process and this unit will take us to the planning phase. This is generally known to deal with the *research design* or the *research proposal*. In this unit, we will try to see what a research design (proposal) is and does, its kinds, why we need it and its basic components. Why do you think one needs to plan for an activity? More specifically, why do researchers plan about their research before the actual implementation? Please write your answers on the space provided below:

_____

_____

_____

This unit is organized as:

Objectives of Unit Three

After studying this unit, learners will be able to:

define and explain research design;

justify the need to design a research;

elucidate the components of a research proposal;

and elaborate the types of research design.

Lesson 1: What is Research Design

Description of Lesson 1

In this lesson, we try to see what a research design is about and what it does. We also try to appreciate why one needs to plan about a research.

A research design may be defined as the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure (Kothari, 1986:44). Krishnaswami (2000:132) defined it as a logical and systematic plan prepared for directing a research study and hence the overall scheme or program of research. Note that research design is the planned sequence of the entire process involved in conducting a research study which depends mainly on the research purpose. It is an overall plan, scheme, structure and strategy designed to obtain answers to the research questions or problems that constitute the research.

It should outline the various tasks you plan to undertake to fulfill your research objectives, test hypotheses and obtain answers to your research questions. A narrative of the project in the sequence in which the research segments or activities will be performed is expected with a justification for each method or approach. It involves selecting the most appropriate methods and techniques to solve the problem under investigation and reasons for undertaking the study. This is why some scholars put research design as a comprehensive *plan of the sequence of operations* that a researcher intends to carry out to achieve the research objectives.

Learners shall note why one first needs to design a research – because it facilitates the smooth sailing of the various research operations, thereby making research as much efficient as possible (yielding maximal information with minimal expenditure of effort, time and money). It has to be done very carefully as it has a great bearing on the reliability of the results arrived at and as such constitutes the firm foundation of the entire edifice/building of the research work. In this context, designing a research project can be compared to that of an architect's design for a building. Nevertheless, no design is a precise and specific representation of actual implementation like an architect's design for a building to be followed without deviations. Rather, a research design is a series of guideposts to keep one going in the right direction. It is a tentative plan which may undergo modifications.

On the other hand, without a proposal, a research attempt becomes unfocused and aimless empirical wandering and researcher gets lost in a welter/confusion of irrelevancies. Even though most actual research may not materialize as fully as proposed in the design, it still helps researchers to organize their ideas in a form whereby it will be possible to look for flaws/defects & inadequacies and make the necessary corrections accordingly. Generally, a research proposal should communicate the purpose and plan of the research in such a way that it:

enables the researcher to return to the proposal for her own guidance in decision-making at different stages of the research process; and

convinces supervisors/reviewers that your proposed methodology is commendable, valid, appropriate and feasible in terms of obtaining answers to your research questions or objectives.

i Check Your Progress

3.1 What is a research design? Why one needs to design her research before embarking on actual implementation of research project?

ii Activity 3.1

Guess what kind of [specific] questions your proposal should attempt to answer? Please write them on the space provided below and compare it with the ones given below:

_____

_____

_____

More specifically, a research proposal/design, in order to achieve the above functions, must tell a reader, reviewer, supervisor or donor what it is proposing to do, how to proceed about it, and why the proposed strategy is needed (justification). Some of the most common questions a researcher has to answer in developing a proposal include:

What is the proposed study about?

Why is the study being conducted? [very important] Where will the study be carried out?

What is the objective of the proposed study?

What type of data is required to achieve these objectives and where & how is it collected?

What period of time will the study cover?

Is the study a sample or census survey? If sample, what will be the sampling design? How will be the data analyzed?

In what style will the report be written?

Who are the potential beneficiaries of the proposed research?

What is the resource requirement of the research?

When and how the research output is to be disseminated? And, so on!

The *answer* to these and other questions shall aggregate to achieve the research objectives. Answering these and other questions is addressed in the various components of the research proposal as discussed below.

Lesson Summary

A research design may be defined as the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure

Some scholars put research design as a comprehensive *plan of the sequence of operations* that a researcher intends to carry out to achieve the research objectives.

One first needs to design a research – because it facilitates the smooth sailing of the various research operations.

A well designed research is more effective and efficient than otherwise.

Lesson 2: Components of Research Proposal

This lesson deals with what a typical academic research proposal contains and discusses the nature of each component in brief. We also try to elaborate why these components are required for the planning phase of the research process.

A research proposal may answer more questions than what is stated in **activity 3.1** above. The order in which these and other questions appear in your proposal may, nevertheless, differ based on the requirements of relevant institution or purpose of the research. The discussion that follows on the specific components can, therefore, serve only as a framework and not universal guideline. This format is generally followed in the Department of Economics of Arba Minch University and hence illustrations would follow accordingly.

Dear learners, while studying this unit, if you try to develop your proposal on the working title you identified in the previous unit, you will have a fairly commendable draft design of 8-12 pages for your term-paper. However, you shall develop it further using the knowledge you will have gained from subsequent units.

Let's see these components briefly [*Forget not, developing your design while studying what follows*].

i) Title Page[1]

The title is the most widely read part of your proposal. It should contain the key words – the important words that indicate the subject. This begins with concise representation of the working title followed by the purpose of the study [in your case may be stated as: A Research Proposal for the Term-Paper in Partial Fulfillment of the Requirements for the Course Introduction to Research Methods for Economists, Econ-341]. Then, the name of the author/researcher and advisor/course instructor followed by institutional affiliation stated as Department of Economics, Faculty of Business and Economics, Arba Minch University. Then, the date (month and year) and place (name of town/city) of proposal development at the right bottom corner shall be presented.

[1] See appendix B for model title page for your proposal.

ii) Operational Definition/Glossary of Terms

This is a list of uncommon, professional or strange words or phrases (terms or concepts) used in the proposal. It is important to define all unusual terms and concepts that may otherwise be misinterpreted (or misunderstood). This is needed because the technical terms, words, or phrases having special meaning should be defined operationally/the way they are used in the proposal.

iii) Abstracts

Often, the abstract and the title are the only parts of the proposal that both the technical and the non-technical staff of a reviewers, panel of supervisors or funding agency will read. An effective abstract should present highlights of the main aspects of the proposal concisely and clearly. A good informative abstract starts by stating the problem to be solved through the work proposed, specifying the significance of the problem at the national, regional or international level. It covers the purpose, expected outputs, beneficiaries, expected impact of the work being proposed and the methods to be used. A statement specifying the budget should also be given. Thus, the abstract presents all parts of the proposal, each part condensed according to the significance of information in it but the total shall not exceed a page (usually half a page).

A good technique is to write a draft abstract before writing the main proposal. This gives you an outline for your proposal of the key component and thus guides the writing of the complete proposal. Rework the abstract after completing the full proposal. Edit your abstract for length, adequacy, clarity and conciseness. When your proposal is meant for competition for research grant, you are advised to strictly follow the guideline forwarded by the relevant funding agency.

Check Your Progress

3.2 What does operational definition of your proposal include? Why does one need to have this section in the proposal?

3.3 What is the role of abstract in a given research proposal?

Introduction

The introduction should create reviewers' interest and lay a broad foundation of the problem. It shall give a background to the proposed study, brief review of literature (if any), define the problem, and discuss your motivations for the proposed study.

It is important in the proposal that the problem standout that the reviewer(s) can easily understand it. A problem statement should be presented within a context, and that context should be provided and briefly explained, including a discussion of the conceptual or theoretical framework in which it is rooted. You are expected to clearly & concisely identify & explain the problem within the framework of the theory or line of inquiry that undergirds the study. Sometimes, obscure and poorly formulated problems are masked in an extended discussion. In such cases, reviewers and/or supervisors will have difficulty in recognizing the problem. This should place the research problem in its historical perspective; state the need for studying it and researcher's precise interest in the study of the problem.

Statement of the problem constitutes the *scientific justification* for the study; i.e., the basis of the need for research to generate further knowledge that will contribute to existing knowledge. The statement must be written in a way that gives empirical references to describe the situation and also clearly specify the gaps in existing knowledge of the problem and/or the existing controversy and the non-conclusive evidence. Moreover, there may be very conclusive evidence for knowledge considered to be established, but the investigator questions the accumulated knowledge because of certain events that he or she intends to subject to verification. It is at this point where the investigator defines the object of study and conveys the questions or broader issues motivating the research.

The problem statement should make a *convincing argument* that there exists not sufficient knowledge to explain the problem and its possible alternative solutions, or for the need to test what is known and taken as fact, if it is called into question by new findings or conditions. The significant content of the justification is that no work has been carried out to try to solve the stated problem, or that little work has been carried out to try to solve the stated problem, or that some work has been done but gaps still exit; hence the need to conduct the study. If, as far as you know, no work has been conducted in the area being proposed for study, then this part becomes a simple, one- or two-sentence statement. But, be cautioned that most *student-topics* are likely to have some literature, attempts have usually been made to solve the problem, or work on a related

topic may have been conducted. In this case, the work being proposed may be intended to fill gaps, use existing knowledge to provide solutions, or carry further an effort that has already been started or is even advanced.

For this purpose, your survey of the literature must show that the work is firmly premised in previous relevant works and thereby building the theoretical framework. This, derived from the statement of the problem, is the argumentation and demonstration that the question has a basis/ground for probable answer(s) and/or working hypotheses from theory. It involves the establishment of relationships – identification of the relationships between the independent variable and the response variable(s). What is known, and how has it been explained? Are the results conclusive? What are the bases of the question? How are the possible answers to the question explained and defended? What are the assumptions? What are the relationships? What are the working hypotheses?

In this section you need to explain how your research problem intersects with the main theoretical debates in the field. One way to organize your review is to explain the historical transformation of your field and where your work fits in with this transformation. Another way is to organize your review around major debates in the field.

It is advised to start broad, become increasingly specific, and end with a review to broaden out and discuss potential applications (importance) of the proposed work. Remember that your goal in reviewing the literature is, in a few words as possible, to sum up all reviewers need to know to have a perspective on what is coming in your analysis. It is not enough just to describe what authors X, Y, and Z did but should help in defining the problem, formulating hypotheses and developing the methodology.

The results of your review are expected to be interpreted in the context of the overall review and study objectives. In particular, it is wise to make a discussion on contrasting evidence, possible sources for discrepancies (experimental design, lack of controls, sensitivity of measures, etc.), and the importance of resolving the differences.

Besides, it must show topics to be addressed and uncover what has been done and what has not; what is needed and why; and indicate your part or contribution on the proposed research. The placement of literature review and motivation of the proposed study, however, is flexible and

hence these could assume separate sections. It is important to revisit the requirements of the relevant institution you are submitting your proposal be it for funding or degree award or completion of a course.

iv) Objectives of the Study

This section states the objectives of the proposed study usually in two categories: general and specific objectives. Recall that objectives have to be SMART (specific, measurable, achievable, realistic and time-bounded) and that the specific objectives derive from the general objective. It is typically very brief, usually half a page at most. This is because the rationale for each objective will already have been established in the previous section, while the ways of achieving the objectives should be explained in the methodology section.

The general objective provides a short statement of the scientific goal being pursued by the research. The specific objectives are operational in nature. They may indicate specific types of knowledge to be produced, certain audiences to be reached, etc. These are the objectives against which the success of the research will be judged. Not only is it important to distinguish the specific objectives from the means of achieving them, such as pursuing fieldwork, organizing a workshop, or publishing a book, but also that the objectives flow logically and clearly from the research purpose, problem statement and justification already stated.

v) Methodology of the Study

The methodology generally addresses issues related to data collection, sample size and sampling techniques, data analysis, hypothesis testing and the likes. Hence, now, you are actually talking about *what and how* you are proposing to do your research. Again, be sure that this follows naturally from the introduction in which you should have set up and highlighted some critical issues that needs to be resolved. What you should do now is to tell the reader how you would resolve this issue. It should describe the proposed method among alternatives for the experiment (including whom the participants will be, what equipment will be used, and the procedure that will be followed). You should specify the variables (independent and dependent) that will be used in the experiment.

I want to underline that a clearly structured methodology is very important and, therefore, this section is really *the heart of the research proposal*. The activities should be described, with as much detail as possible, and the continuity between them should be apparent. You should indicate the methodological steps you will take to answer every question or to test every hypothesis illustrated in your *hypotheses* [sub] section.

It is best to organize the methodology to explain how each specific objective will be achieved. The proposal should provide enough detail to enable an independent scientific assessment of the proposal. Assuming that the research questions and research hypothesis to be addressed have been clearly identified, the purpose of the methodology section is to show how these questions will be answered in the most rigorous way possible. The methodology section deserves greater emphasis than students typically give to it. The proposal needs to be clear about what activities are envisaged in the pursuit of each objective, and this must be done before your proposal (and funding) is approved. Indeed, it is impossible to define the budgetary needs of the research project in the absence of a solid methodology section.

You should outline the general plan for collecting the required data. This may include survey administration procedures, interview or observation procedures – indicate what approaches and methods would be used to collect primary and secondary data. You also need to provide details on available sources of secondary data. The investigator should describe the quality of the secondary data. Besides, the instruments to be used for the collection of primary data, such as questionnaires and interview protocols should be developed and attached as an annex to the proposal. You should outline the procedures for the development, pre-testing, and administering of any data collection instrument. If survey work is involved, give detailed information on the study area. If the research is related to human populations, information on the study population should also be provided. Include a description of the procedures for selecting the population sample and the sample size with the necessary justification.

Another issue to be addressed in your proposal is the *sampling design*. The key reason for being concerned with sampling is that of *validity* – the extent to which the interpretations of the results of the study follow from the study itself and the extent to which results may be generalized to other situations with other people. Sampling is critical to *external validity* – the extent to which findings of a study can be generalized to people or situations other than those observed in the

study. To generalize validly the findings from a sample to some defined population requires that the sample has been drawn from that population according to one of several sampling techniques. Another reason for being concerned with sampling is that of *internal validity* – the extent to which the outcomes of a study result from the variables that were manipulated, measured, or selected rather than from other variables not systematically treated. Perhaps the key word in sampling is representativeness. You must ask yourself, *how representative is the sample of the survey population* (the group from which the sample is selected) and *how representative is the survey population of the target population* (the larger group to which we wish to generalize)? Generally, probability samples are believed to have more internal and eternal validity than samples drawn out of non-probability techniques. When the latter is the case, the rationale and limitations must be clearly furnished.

Besides, a detail of procedures to follow to obtain informed consent and ensure anonymity and/or confidentiality of respondents/informants should be provided. Then, you outline the instruments you propose to use (surveys, scales, interview protocols, observation grids, etc). If instruments have previously been used, identify previous studies and findings related to reliability and validity. If instruments have not previously been used, outline procedures you will follow to develop and test their reliability and validity. In the latter case, a pilot study, if possible, is nearly essential and hence needs to be incorporated in the proposal. Because selection of instruments in most cases provides the operational definition of constructs, this is again a crucial step in the proposal. Strictly speaking, results of your study will be directly relevant only to the instrumental or operational statements and hence operational definitions have to be made very carefully.

A careful researcher tries to take advantage of accounting the comments of reviewers by including an appendix with a copy of the instruments to be used. For a mailed survey (most widely used by student researchers), in particular, you need to identify steps to be taken in administering and following up the survey to obtain a high response rate.

Your discussion on the *method of analysis* may explain how to process the data like the type of statistical tools and inference models, econometric models, mathematical programming models, simulation models, etc to be used. Fortunately, economics is rich in methods of analysis to

choose from. Your review of literature may help in your selection of analysis techniques. Include an explicit statement covering the field controls to be employed. Specify the procedures you will use in your data analyses and label them. If coding procedures are to be used, describe in reasonable detail. Communicate your precise reasons for these intentions to the reader. This helps you and the reader evaluate the choices you made and procedures you followed. Indicate briefly any analytic tools/softwares you will have available and expect to use [like STATA, EVIEWS, and SPSS – all available in the Department of Economics].

Learners have to note that all research is plagued by the presence of confounding/difficult variables (the noise that covers up the information you would like to have). Confounding variables should be minimized by various kinds of controls or be estimated and taken into account by randomization processes. One has to indicate the variables you propose to control and how you propose to control them, experimentally or statistically, and the variables you propose to randomize and the nature of the randomizing unit. Be aware of possible sources of error to which your design exposes you. Nevertheless, you will not produce a perfect, error-free design (no one can). Yet, you should anticipate possible sources of error and attempt to overcome them or take them into account. In addition, you should disclose to the reader the sources you have identified and what efforts you have made to account for them.

In connection, your methodology is expected to show how you are planning to test your hypotheses. A hypothesis, as you might recall, is a proposition that is assumed, perhaps without belief, to draw out its logical consequences and by this method to test that it is in accord with facts that are known or may be determined. Hypotheses can be developed from a body of theory. You can also construct hypotheses by logical deduction from theory. Theory is the most fertile seedbed for hypotheses. A theory presents what is known. Logical deductions from this constitute the hypotheses that must be true if the theory is true. There are three different approaches in formulating hypotheses:

Deductive approach – some social scientists argue that all research should derive hypotheses from theories about behavior

Inductive approach – place emphasis on discovering relationships in the data and generalizing from the observed relationships

Middle ground – many social scientists see the importance of

- Relating their research to pre-existing theory
- But allow for the discovery of the relationships in the data – not limited to testing of pre-established hypotheses

In a nutshell, the methodology section must provide a well thought-out rationale for your decision to use the design, methodology, and analyses you have selected. From the above discussion, learners may understand that the methodology is the part that guides decisions as to:

when and how often to collect data
construction of data collection measures
identification of the sample or test population
choice of strategy for contacting subjects
what data to gather and from whom/where
selection of statistical tools
how to analyze the data,
presentation of the findings, etc

Activity 3.2

Answer the following questions based on the progress of your term-paper proposal:

1. Your proposed study is
   A. Sample Survey
   B. Census Survey
2. What kind of data you planned to use and Why?
   A. Primary data
   B. Secondary data
   C. Both
3. If your answer to question 1 is A, what is
   the a. sample size? Why?

   _____

   _____

b. sampling technique used to draw you sample? Why?

_____

_____

4. How do you plan to analyze your data? And why?

_____

_____

Very good, if you are in a position to answer such questions, then you are good enough to design the methodology of your proposed research.

The diagram below may help capture the major decisions to make and activities to include in your methodological design. The graph, though not comprehensive, may help guide how to go about in subsequent decisions one has to make in designing the methodology of a given research.

A researcher has to first decide on the type of study (pure or applied), which is usually applied for a student research like your as it is based on real world problems. Then, the specific type of study based on relevant variables of consideration shall be determined. You also need to decide on whether the target population is to be treated comprehensively (census survey) or partly (sample survey). Next, you need to decide on the type and source of data as well as its techniques of analysis. Moreover, the methods and instruments of data collection need to be identified if you are using primary data.

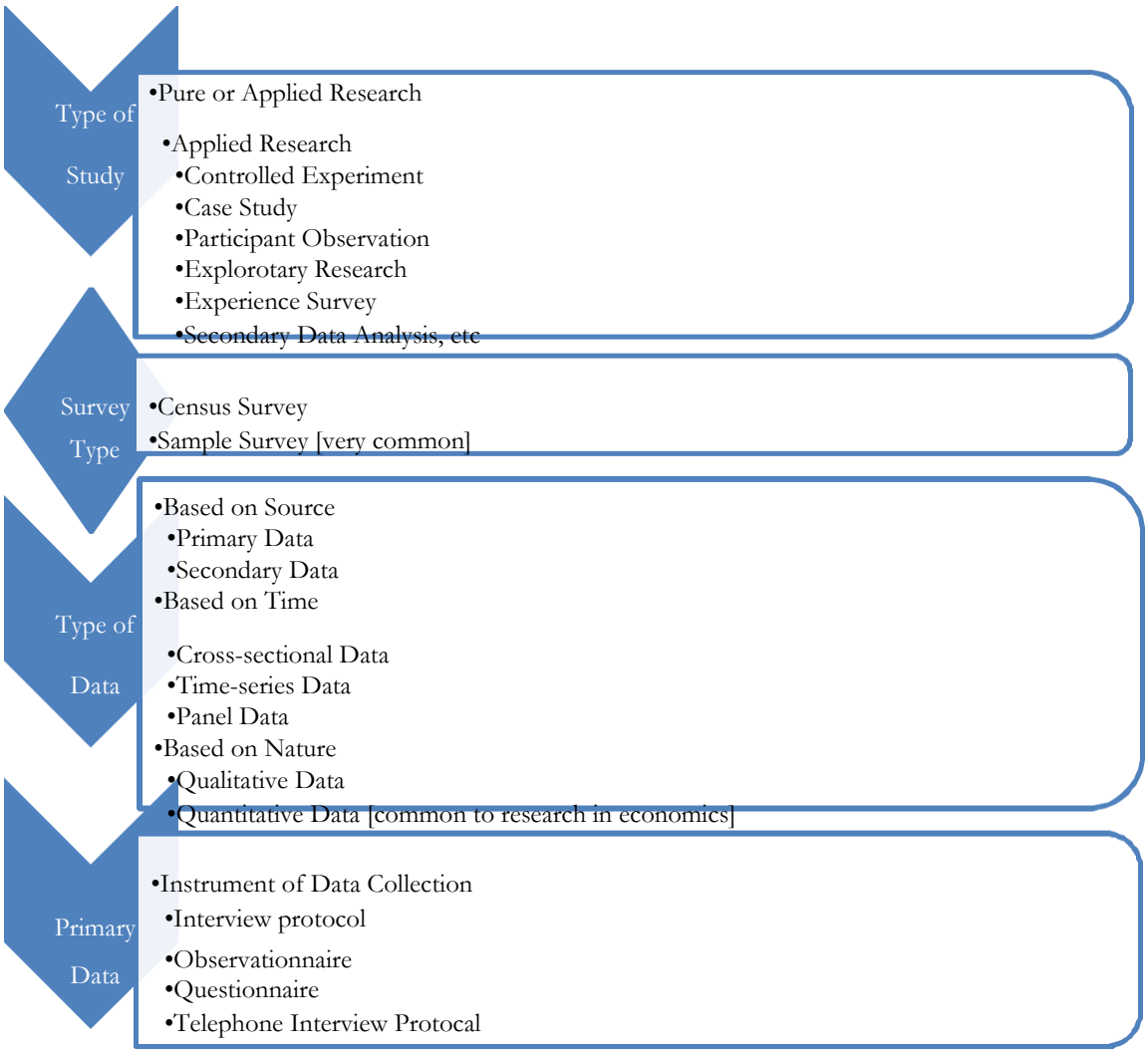


Figure 3.1 Scheme of Research Design

vi) Scope of the Study

This is the place to explain the things that you are not doing and why you have chosen not to do them since a complete study of any problem is almost non-manageable. Such an attempt, if any, would entail such an over-whelming mountain of data that would require more than the researcher's life time to comprehend and complete the study. In many research projects, researchers cannot treat all aspect of the research problem. The researcher, therefore, should carefully state the boundary of the problem to be investigated. Delimitation of a proposed study addresses how a study will be narrowed in scope – how it is bounded. Limit your scope to the things that a reader might reasonably expect you to do but that you, for clearly explained reasons,

have decided not to do. Alternatively, this will be undertaken by specifying areas of all variables which will be considered in the study or by eliminating all aspects and factors which will not be considered in the study.

vii) Significance of the Study

It is the rationale to conduct the research and may be stated, for example, as:

There is no detail local investigation before about your proposed study;

The study may be conducted on the assumption that there is a strong need for local investigation;

To alert decision-makers/policy-makers;

To fill the gap in the existing studies;

To resolve some inconsistency in previous research;

The finding may lead to new problem for further

study; To provide basic data about the problem; etc

Your statement of significance may indicate how your research will refine, revise, or extend existing knowledge in the area under investigation. Note that such refinements, revisions, or extensions may have substantive, theoretical, or methodological significance. You are expected to think pragmatically. Most studies have two potential audiences, viz. practitioners and professionals. Statements relating the research to both groups are always in order. You try to think about *implications* – how results of the proposed study may affect scholarly research, theory, practice, educational interventions, curricula, counseling, policy and other issues of interest. For beginners, however, this can be a difficult section to write. Soon, I will try to make your life simple – when thinking about *significance*, ask yourself the following questions:

What will results mean to the theoretical framework that framed the

study? What suggestions for subsequent research arise from the findings?

What will the results mean to the practicing educator and researcher?

Will results influence programs, methods, and/or interventions of concerned institutions?

Will results contribute to the solution of educational, social or economic problems?

What will be improved or changed as a result of the proposed research?

How will results of the study be implemented, and what innovations will come about?

Put briefly, the significance section of the research proposal describes the type of knowledge expected to be obtained and the intended purpose of its application. It should indicate the strategy for disseminating and using the research findings according to the potential users of the knowledge generated. When your proposal is competing for a research grant, this section should answer the following:

How does the proposed research relate to the priorities of the region and the country? How is the proposed research of interest to the funding agency?

What knowledge and information that is of interest to the agency or practitioners will be obtained?

What is the ultimate purpose that the knowledge obtained from the study will serve? How will the results be used, and who will be the beneficiaries?

Your answers to these and related questions should make a convincing argument that the knowledge generated will be useful and generally applicable within the regional context or of interest to the potential funding agency. This section is so important it determines whether or not you will win the research grant you are competing for.

Either way you need to show where your research will contribute to the current theoretical literature or what practical solution will result from this particular inquiry.

viii) Budget Schedule (Time, Finance and Other Resources)

It is interesting to see that it is time to show the constraints of the proposed study, mainly time and finance. In this section, your are expected to provide a general outline of the time schedule you expect to follow, and cost breakdown in terms of financial, human labor and other resource requirements. Indicate not only the time needed to carry out each phase of the research and its total duration, but also possible constraints in adhering to the timetable.

I would advise you to allow for inflation (or more generally as contingency) and indicate the level of inflation used in the estimate. All budget items must be quoted in relevant currencies (birr or otherwise).

ix) Reference[2]

It is important that you correctly cite all consulted published and unpublished documents that you refer to in the proposal development. This allows the reader to know the sources of your information. Every reference you cite must appear in the bibliography at the end of the proposal.

x) Appendix

This component of your proposal contains information less relevant to include in the body of your proposal yet may be of interest to some of the readers, reviewers, supervisors or funding agency. It usually comprises of details of activities/strategies, time schedule, and cost breakdown.

Check Your Progress

3.4 What does one show in writing the significance of a proposed study?

3.5 Which section of your design is most read by technical or non-technical unit of a potential funding agency? Why?

3.6 What is the role of defining the scope of a proposed study in the research design?

Lesson Summary

Research proposal mainly contains the title page, operational definitions (if any), abstract, introduction, objectives, methodology, scope, significance, budget schedule, appendix and reference.

The format for a research design may not be identical between different institutions.

[2] More on writing of bibliography and reference in chapter 6

Lesson 3: Kinds of Research Design

Description of Lesson 3

Lesson 3 touches briefly on the kinds of research design. The details of this discussion is beyond the scope of this module and is of less interest to distance learner of Economics.

There are numerous research designs, which fall into a limited number of main kinds. Each has advantages and corresponding disadvantages. The most common trade-off is between precision and reality. Usually the more precisely controlled your design is, the more unrealistic the experimental task and setting are. Wilkinson and Bhandarkar (2007) have listed the following kinds/ major parts of research design (but this classification is not binding for all types of research):

i) The sampling design: technique(s) of selecting the items to be studied;
ii) Observational design: conditions under which observation/data collection is to be made;
iii) Statistical design: how many items are to be observed, how the information or data generated is to be analyzed; and
iv) Operational design: techniques the procedures specified in sampling observational and statistical design can be carried out.

Note, however, that these are not the only kinds of designs/parts of a design. More detailed discussions are available in modern research method books. [3]

[3] If you are interested in this subject please refer to other books

Lesson 4: Features of Good Design

Lesson 4 briefly touches on the characteristic features of a good research design. It is a preliminary introduction of the most common desirable features for a proposal to be judged as good.

A good research proposal must posses the important characteristics of *objectivity*, *reliability* (the degree to which a measure is consistent or dependable and the degree to which it would give you the same result over and over again, assuming the underlying phenomenon is not changing), *validity* (the best available approximation of the truth of a given proposition, inference, or conclusion) and *generalization*.

We need to strive to make our proposal qualify the above desirable characteristics because if the design is good, all may go well. A well conceived research design will pay rich dividends. In spite of all these features, a research design is *not a rigid plan* to be followed without change, but like a polestar which help the researcher to find out the right direction. As research progresses, a change in original design may become imperative due to new facts, ideas and conditions. In addition to the regular characteristics of a good research design, the following are the additional features to support a good research design: *flexibility*, *appropriateness*, *statistical and economic efficiency minimize bias* and *maximize the reliability* of data, *smallest experimental error*, *effectiveness* (yields maximum information and considers many different aspects of the problem).

Finally, a good research design is organized, has a logical flow, concise (but also complete), uses a good grammar. To achieve all these, it's usually a good idea to have a colleague/friend read it before submitting it to the advisor/supervisor/funding agency. Often, little errors or small changes will be identified and addressed in this process of proof-reading.

Check Your Progress

3.7 What is a good research design? What makes it different from a bad design?

Lesson Summary

A good research design is objective, reliable, valid, general, flexible, efficient, flexible, and effective, among other qualities.

Review Exercises

1. What is a research proposal used for?
2. What are the basic questions a research proposal is intended to answer?
3. What is the methodology of a proposed research concerned about?
4. Why does a researcher need to define the scope of the proposed research?
5. How do you measure the quality of a certain proposal?
6. State some of the components of a research proposal for academic purpose.

UNIT FOUR

CENSUS AND SAMPLE INVESTIGTION

INTRODUCTION

A process of gathering information for the purpose of analysis and resolving a problem is generally known as *survey*. The entire collection of people or things – you are interested about and may consist of all elements, individuals, institutes, regions, kebeles, items, or objects whose characteristics are being studied is called *population/universe*. A survey that makes a complete enumeration of all items in the population, it is known as *census survey*. The technique of collecting information from a portion of the populations is called *sample survey*. A *sample*, therefore, is a (proper) subset of the population. Accordingly, *sampling* is the procedure of selecting a sample from a population.

In sample surveys, a researcher draws a sample from a larger pool of cases, or elements. We said that the large pool that comprises all the sampling elements is called the *universe* or *population*. The population that is being studied is also called *target population*. Alternatively, population refers to any collection of specified group of cases to be studied, also known as *sampling elements*. Hence, the *sampling element* is the unit of analysis or case in population from which information on variables[4] of interest is collected and provides basis for analysis – it is the subject on which measurement is being taken. Whereas, s*ampling unit* refers to the element or set of elements considered for selection in some stage of sampling. This might involve several stages based on the nature of the population and the problem under investigation – for example, if population is 'rural people in SNNP', Zones, and Woredas may serve as sampling units. [5]

A population with countable and known number of cases is known as *finite population* (students in Arba Minch University, Households in Arba Minch town, Textile firms in Ethiopia, the nations of the world, civil wars in Africa since 1900, etc). On the other hand, a population with an infinite number of units is *infinite population* or *hypothetical population* (number of stars in the sky, soil in Arba Minch Zuria Woreda, etc).

The actual list of sampling units from which the sample or sample stage of a sample is selected is known as *sampling frame* – for example, students roaster, tax records, payroll, telephone directory, municipality records of households, etc. Finally, the proportion of the sample drawn from a population may be measured by s*ampling ratio* (s) – the ratio of sample size (n) to the size of the population (N) [s=n/N].

---

[4] A variable is a characteristic under study that assumes different values for different elements.

5 As there exist several types of sampling techniques, some require a list of sampling units for sampling.

Note, however, that the ultimate purpose of conducting a sample survey is to make decisions about the corresponding relevant population. It is important that the results obtained from a sample survey closely represent the results that we would have obtained by conducting a census. For example, to find the average income of families living in Arba Minch Town by conducting a sample survey, the sample must contain a family who belongs to different groups in almost the same proportion as they exist in the population. Such a sample is called a *representativ*e sample.[6]

As indicated above, when one scrutinizes a sample, she is interested about the population. That is, a sample *statistic* (a number that results from measuring all the units in the sample) is useful only as far as it can be inferred about the population parameter (a number that results from measuring all the units in the population – hypothetical in the case of sample survey).

Note also that certain techniques of sampling require the list of members of the target population. Such a list used in the sampling process is known as *sampling frame* (e.g., a telephone book). Given the target population, the *sampling unit* is unit of analysis that refers to the type of specific object of interest (for instance, students, firms, households, etc). For example, to find out the average income of households in Arba Minch in 2001:

The target population is all households in Arba Minch in 2001;

The list in the municipality of the town or respective kebeles in that period may serve as a sampling frame;

The sampling unit is a household;

The design of sampling may be either probability or non-probability as shall be seen soon;

Sample size may be, say, 50 or 100;

A variable, for example, may be monthly income;

A statistic may be the arithmetic average of income from selected households in the sample;

A parameter may be the conclusion about the monthly average income you inferred from the statistic found in the sample survey.

Activity 4.1

1. A researcher needs your help in identifying her target population for a proposed study entitled ‘The Role of Microfinance in Poverty Reduction in SNNPR’.

   _____

   _____

[6] More on representativeness later

2. A learner is doing her term-paper on ‘Family Background and Academic Performance of Students in Arba Minch University’. What is the target population, a relevant variable?

   _____

   _____

3. A student-researcher is conducting a research on ‘The Sources of Financing College Fee for Continuing and Distance Program Students: the Case of Arba Minch University.’

   i) Identify the population _____

   ii) What is the sampling element? _____

   iii) Indicate potential sources of sampling frame _____

   iv) If the sample size is 200, what is the sampling ratio (assume the population is 5,000)?

Check Your Progress

4.1 What is population? How does a researcher identify a target population?

4.2 What is sample?

4.3 Why does one engage in sampling while the interest remains with the target population?

4.4 What is a variable? What is the difference between statistic and parameter?

iii Why Sampling?

As compared to census, sample surveys dominate in real world research. And, most student-researches, if they use primary data, are sample surveys. Why do you think is the case? Please your answers on the space provided below and compare it with the discussion that follows:

_____

_____

First, sometimes 'measuring' or 'testing' something destroys it. Suppose the Ethiopian Quality and Standards Assurance Authority wanted processed coffee exporters demonstrate the flavor of their products. Obviously, the companies can't be expected to make coffee out of all their produce to show its quality. So, the companies make a sample of few cups of coffee.

Second, though seldom, not all units in the population can be identified, such as all the soil in Arba Minch. So to measure, say, fertility of soil, one must take a sample of soil from areas under investigation. Otherwise, an attempt to identify all the soil in the study area is extremely costly, or infeasible and useless.

Third, sampling not only saves money, time and effort but also is simple to undertake. Fourth, it can be conducted under [serious] resource limitation.

Fifth, it may result in superior quality – since census requires large staff, long interviewing period, difficult to specify the time to which the data refers, managerial problems in supervision & data administration, record keeping, training, etc that may all have a huge implication on the data quality and hence the quality of expected results.

Sixth, some problems are so urgent one may not have the luxury of waiting for census results as its results may turn out irrelevant when the project matures. In such cases, it dictates the bad need of quick and timely results which is possible only via sample survey.

Finally, if a population is infinite, the only feasible survey would be using a sample as infinite population with its diversity & composition carries it own problems.

Therefore, these justify the dominance of sample survey in real world research than census survey. Even though sample survey considers only part of the population, the basic idea of sampling as indicated above is that some of the elements in a population provide useful information on the entire population. That is, sampling aims at obtaining *consistent* and *unbiased* estimates of the population. A question of interest, therefore, may be how we draw a sample which enables a consistent and unbiased estimate of the population behavior? How do we draw a sample from a population?

A census survey is fortunate enough to be totally reliably correct if errors of administration, interview, respondents, etc are fully avoided since there is no sampling error involved. Besides, the parameters computed are more precise. For instance, policy decisions that involve population size require census survey as, for example, the Central Statistical Agency does every 10 years. However, census is usually not only costly but also involves [huge] errors of administration, interview, respondents, data management, etc. Hence, the aggregate implication in most real world research is that census survey may involve large non-sampling errors. This dictates, therefore, sampling is most often (and census less often) employed. The question, however, remains how we can arrive at a representative sample – think about the question!

Activity 4.2

Learners were grouped into two teams and allowed to debate on the relative importance of sample and census surveys. Put yourself in each team in turns and state your points briefly:

[for sample survey] _____

[for census survey] _____

iv Sampling Techniques

One thing from which you are unlikely to suffer when it comes to data collection is lack of choice of sampling techniques. I presume, you already decided in designing on not only which method to use but also why you chose that specific method. I believe you have revisited your materials on the various sampling techniques in your statistics. Yet, it is indispensable to discuss them in the context of their respective merits and demerits.

The selection process in sampling is called *sampling technique*. There are two categories of sampling techniques, namely *random* (probabilistic) and *non-random* (non-probabilistic). In a random sampling technique each element of the population has a known (equal) chance of being included in the sample. However, in a non-random sampling technique elements have no known chance of being selected in the sample. A sample drawn in such a way that each element of the population has a known (equal) chance of being selected is called *a random sample*; otherwise a *non-random* sample.

The logic of the theory of sampling is *induction*; that is, we proceed from particular (sample) to the general (population) and all results are expressed in terms of probability. If the population was identical or perfectly homogeneous, we do not need a careful sampling procedure, any sample would suffice. The real world social problems, nevertheless, often involve imperfectly homogenous or heterogeneous populations. So, when faced with variation or heterogeneity, in the population, more controlled sampling procedures are required. How homogenous is the population you identified for your term-paper?

4.2.1. Probability Sampling Techniques

We said that probability sampling techniques give every member of the population a known, non-zero or equal probability of selection. It is selection done haphazardly/without purpose or aim on specific elements of the universe. It is widely used technique of sampling. There are four types of probability sampling techniques relevant for our discussion: simple random sampling, systematic sampling, stratified sampling and cluster sampling.

i) Simple Random Sampling

A *simple random*/*unrestricted* sampling is a technique that draws a sample in such a way that every element in the population has the same known/equal chance of being chosen in the sample. In this technique, every unit of the population of interest must be identified, and all units must have a known, equal non-zero chance of being selected in the sample. In addition, the selection of each unit is independent of the selection of every other unit – selection of one unit does not affect the chances of other units.

One way to select a random sample is by lottery method; i.e. giving a distinct number (code) for each and every element of the population; form lots of equal size, shape and color; and drawing as much of the lots as the size of the sample desired. It is the same as the practice of picking lottery winners; hence also sometimes known as the *lottery method*.

Often, researchers use *random tables* that are attached in most statistics books' annexes for drawing simple random samples. For example, to select a sample of 75 households in a given woreda where 500 registered households live, make a list of all the 500 households and assign each household a code, say, between 001 – 500. Then, refer to the *table of random numbers*. Starting at any point of the table read across or down (which ever direction); draw every number whose last three digits fall between 001 and 500. Use the numbers/codes you have found to pull the names from the list that correspond to the selected households. These 75 households constitute your sample.

Another common way to draw simple random sample is to take 500 ping-pong balls and number them from 001 to 500; to put them into a large barrel and mix them up; and then grab 75 balls; read-off the numbers. Those are the households in your sample.

This technique is widely applicable and results in more representative sample. Make sure that you cannot apply simple random sampling techniques without a reliable sampling frame. However, when the target population has a pattern of distribution that assumes sub-grouping, it may not guarantee adequate representation of each group of the population.

Advantages of Simple Random Sampling:

Of all random sampling techniques, this is the easiest technique to apply; It may result in more representative sample if conducted properly;

It does not require prior knowledge of the true composition of the population; The amount of sampling error can easily be computed.

Disadvantages of Simple Random Sampling:

It is often impractical because of lack of (up to date) sampling frame or making one is very costly or impossible;

Its use may be wasteful as we fail to use all the known information about the population;

It does not ensure proportionate/appropriate representation to various groups constituting the population – it is possible that all sample elements may be drawn from a certain group of the population;

The sampling error is usually greater than other random sampling techniques;

It requires larger sample to ensure representativeness and hence costly, especially if the survey covers a large geographic area – why? Because your sample may be scattered evenly across large geographic area and hence involves costly administration;

This technique of sampling is suitable only for a small homogenous population and available sampling frame (or easy to make one).

ii) Systematic Sampling

A systematic random sample is a sample which contains every k^{th} element of the population. The first element is chosen randomly between 1 and k (k=N/n and it is an integer; where N and n are the population size and desired sample size, respectively), but the rest elements systematically – hence seldom known as *quasi-random* sampling technique. As the interval between sample units is fixed (k), it is also known as *fixed interval method*. Strictly speaking, it is not a probability sampling; it possesses characteristics of randomness and some non-probability traits. Hence, it is sometimes known as *pseudo-random* sampling technique. The first element may be drawn using the lottery method or any other relevant random sampling technique. It is one of the most widely used probability sampling technique. For example, from the population give above (500 households) we may draw a sample of 75 households by a random start between 1 and 7 (the integer approximating 500/75) and then every 7^{th} element sequentially. Suppose the 3^{rd} element is randomly selected. Then, the 10^{th}, 17^{th}, 24^{th}, and so on elements are selected to constitute your sample.

Though this method may be applicable to a wide range of populations, it brings a better result if a population is frame re-arranged in a systematic order on the basis of important population characteristics, especially for populations with greater variability.

Advantages of Systematic Sampling:

It is much simpler than simple random sampling – it is easy to use;

It is easy to instruct field investigators and data collectors to use this technique;

It requires less time – a researcher operating in limited schedule will prefer this technique; It is cheaper than simple random sampling;

It is easier to check whether every k^{th} item has been included in the sample; Usually the sample is evenly spread over the population;

It is statistically more efficient than simple random sampling when population is (can be) ordered chronologically;

It is flexible, less expensiveness and easy to carry out.

This method is useful for selecting large samples, say 100 or more. It is less cumbersome than a simple random sampling. For example, you might have to sample files in a large filing cabinet, then it is easier to select every 17^{th} file than to pull out all the files and code them for simple random sampling.

Disadvantages of Systematic Sampling:

It ignores all the elements between two k^{th} elements selected;

Except the first element, other selected elements are not chosen randomly; hence it is not a probability sampling technique in the strict sense of the term;

It is not advised to apply for studies that aim at estimation and generalization because of non-random nature involved in selecting k^{th} elements except the first.

It is dangerous when the population has a *trending nature* or *systematically arranged*. For instance, a research on market demand for a commodity, if placed at every market day, doesn't reflect the true picture of the market condition – this is *biased*.

Activity 4.3

v Suppose a researcher has data on the average monthly pocket-income distribution of graduating students of Economics in Arba Minch University as given below:

vi	**vii**	**viii**	**ix**	**x**	**xi**	**xii**	**xiii**	**xiv**	**xv**	**xvi**	**xvii**	**xviii**	**xix**	**xx**	**xxi**	**xxii**	**xxii**	**xxi**	**xxv**
xxv	**xxv**	**xxvii**	**xxi**	**xxx**	**xxx**	**xxx**	**xxx**	**xxx**	**xxx**	**xxx**	**xxx**	**xxxvi**	**xxx**	**xl**	**xli**	**xlii**	**xliii**	**xliv**	**xlv**
xlvi	**xlvi**	**xlviii**	**xlix**	**l**	**li**	**lii**	**liii**	**liv**	**lv**	**lvi**	**lvii**	**lviii**	**lix**	**lx**	**lxi**	**lxii**	**lxiii**	**lxiv**	**lxv**
lxvi	**lxvi**	**lxviii**	**lxix**	**lxx**	**lxxi**	**lxxi**	**lxxi**	**lxxi**	**lxx**	**lxxv**	**lxx**	**lxxvi**	**lxxi**	**lxx**	**lxx**	**lxx**	**lxx**	**lxx**	**lxx**
lxx	**lxx**	**lxxxv**	**lxx**	**xc**	**xci**	**xcii**	**xcii**	**xciv**	**xcv**	**xcvi**	**xcv**	**xcviii**	**xci**	**c**	**ci**	**cii**	**ciii**	**civ**	**cv**
cvi	**cvii**	**cviii**	**cix**	**cx**	**cxi**	**cxii**	**cxii**	**cxiv**	**cxv**	**cxvi**	**cxv**	**cxviii**	**cxi**	**cxx**	**cxx**	**cxx**	**cxx**	**cxx**	**cxx**
cxx	**cxx**	**cxxvi**	**cxx**	**cxx**	**cxx**	**cxx**	**cxx**	**cxx**	**cxx**	**cxx**	**cxx**	**cxxx**	**cxx**	**cxl**	**cxli**	**cxli**	**cxli**	**cxli**	**cxlv**
cxlv	**cxlv**	**cxlvii**	**cxli**	**cl**	**cli**	**clii**	**cliii**	**cliv**	**clv**	**clvi**	**clvi**	**clviii**	**clix**	**clx**	**clxi**	**clxi**	**clxi**	**clxi**	**clxv**

1. If the researcher wants to draw a sample of 40 students to analyze the source and outlay of their pocket income, what is k (interval with which elements of the sample have to be drawn)?
2. Draw 3 samples of the same size (40) using:
    a. simple random sampling,
    b. systematic sampling with the population as given, and
    c. systematic sampling with the population chronologically arranged

Compare the representativeness of the samples (you may compare the population mean and the mean of each sample). Which sample yields the best sample?

iii) Stratified Random Samples

If a population does not constitute a homogenous group, stratified sampling technique is generally applied in order to obtain a representative sample. Virtually most populations can be segregated into a number of mutually exclusive groups. Dividing a population into non-overlapping groups based on certain criteria is called *stratification*. Stratification helps increase the sample statistical efficiency, provide adequate data for analyzing the various groups and also make possible to apply different techniques to different groups. When stratifying a population, make sure stratum is homogenous internally and heterogeneous compared with other strata – intra stratum homogeneity and inter strata heterogeneity in order to ensure representation of each group. This is why this technique is particularly useful when we have a large *heterogeneous* population.

When one chooses to use this technique of sampling, she needs to decide on:

- the stratification base or bases to be used; the ideal variable being the variable under study. For example, for a study on determinants of investment in Ethiopia, one may stratify investors into agriculture, manufacturing, service, etc.
- the number of strata; though there is no precise answer to this question. Note that the larger is the number of strata, the more is the degree of representativeness of the sample. This decision is made based on the number of groups to be studied and the cost of stratification.
- strata sample size; there are two alternatives: first, the strata sample sizes may be proportionate to strata's share in the total population – *proportionate stratified sampling*; and second, they may be disproportionate to strata's shares – *disproportionate stratified sampling*.

For example, in the above research on the determinants of investment in Ethiopia, you may end up with investors in only one sector (say agriculture) or inadequate representation of other sectors with simple random sampling. With stratified sampling, on the other hand, you may come up with adequate representation of investors in each sector. It should, therefore, be underlined that when the population is heterogeneous, it is preferable to stratify the population.

The example given below (using hypothetical data) may illustrate how a stratified sampling technique may come up with a better/adequate representation of members from each group. The data given in the table below shows the number of tourists arrived in Ethiopia from various parts of the world both in the *derg* regime *and EPRDF* regime. Note also that the stratification bases are two, viz. the origin of tourists and year/regime of visit.

	Derg Regime				EPRDF Regime			
Origin of	**1976-**	**1981-**	**1986-**	**Total**	**1991-**	**1996-**	**2001-**	**Total**
USA	325	210	150	**685**	350	800	1200	**2350**
Latin	35	45	50	**130**	85	75	90	**250**
Asia	540	670	1260	**2470**	400	500	800	**1700**
Europe	850	1225	1110	**3185**	1250	2400	3650	**7300**
Africa	35	65	50	**150**	45	55	80	**180**
Total	**1785**	**2215**	**2620**	**6620**	**2130**	**3820**	**5820**	**11780**

If you select a simple random sample of 460 tourists out of a total of 18,400 tourists during the two regimes, you might not get any from Africa or the period during the *derg* regime. To make sure that you get some tourists from each group, you can divide them into these groups (time wise and origin wise), and then select the *same percentage* (460/18,400 = 2.5% from each group)

of tourists from each group using a simple random sampling method. This is *proportional stratified random sampling*.

Generally speaking, if the number of units selected from each stratum is proportional to the total number of units of the strata, we have *proportionate sampling* and hence the allocation among proportionate sampling becomes $n_1/N_1 = n_2/N_2 = \ldots = n_s/N_s$ [Where N=strata size and n=sample from each stratum]. Accordingly, if N_1 is not equal to N_2, then n_1 is not equal to n_2. But when the number of units drawn from each stratum is equal regardless of the size of the strata or not proportional to the size of the strata, it is called *disproportionate stratified sampling*. This applied to surveys that give special importance to certain sub-groups of the target population or when one needs equal number of units from each sub-group.

Can you do the corresponding table that dictates on the sample size from each group yourself please? Very good! Now, compare your results with the one given below. However, you may still have too few of some tourists of a certain origin or period, such as those from Latin America and during periods like 1986-1990.

Hence, you may, instead, divide them into groups and then select the same number of students from each group using a simple random sampling method regardless of differences in their number in categories of different origin and period of visit. This is *disproportionate stratified random sampling*. This allows you to have adequate number of tourists in each group and sub-group so that you can perform some meaningful statistical analyses. The desirability of such sampling is usually determined by the size of the strata, internal variances among strata and sampling costs – a researcher may take a larger sample from a given stratum if the stratum is large, or has more internal variability or cheaper to draw.

Origin of Tourists	Derg Regime			EPRDF Regime		
	1976-1980	**1981-1985**	**1986-1990**	**1991-1995**	**1996-2000**	**2001-2005**
USA	8	5	4	9	20	30
Latin America	1	1	1	2	1	2
Asia	14	17	32	10	13	20
Europe	21	31	28	31	60	91
Africa	1	2	1	1	1	2
Total	**45**	**56**	**66**	**53**	**95**	**145**

Suppose 12 business enterprises in Arba Minch have total monthly labor cost of 1, 2, 3, 10, 11, 12, 50, 51, 52, 100, 101, and 102 (in thousands of Birr). Let's compare the mean with and without stratification: Without stratification, if our sample size is 11, we have maximum and minimum sample means 44.9 and 35.7, respectively and the range of error is 9.2 (44.9 – 35.7). But, if our sample size is 4, the maximum and minimum sample mean are 88.75 and 4,

respectively and the range of error is 84.75. With stratification, we may form four groups based on the relative size of their wage expenditure as given below:

Relative Size of wage expenditure	Very small	Small	Medium	Large
Wage Expenditure (absolute terms)	1, 2, and 3	10, 12, and 13	50, 51, and 52	100, 101 and 102

Then, if our sample size is 4, we will draw one from each stratum (proportional or non-proportional?). Accordingly, for this sample size the maximum and minimum sample mean are 42.52 and 40.25, respectively and the range of error for this data is only 2.27. Hence, the minimum error is attained when the stratified sampling system is adopted. Can you try the mean and error yourselves?

Advantages of Stratified Sampling:

enhances the representation of different groups of the population especially when it is heterogeneous;

easy to carry out; and

facilitates giving weight to different groups;

Disadvantages of Stratified Sampling:

requires a prior knowledge of the composition of the population and its distribution;

researcher may commit classification error and hence may vitiate the interpretation of survey results; and

practical feasibility is questioned especially for disproportionate sampling that aims at optimal classification based on variability and when one does not know relative variability of strata.

iv) Cluster Sampling

Where population elements are scattered over a wider area and/or a list of population elements is not readily available, the use of other random sampling would be too expensive or even infeasible. So, if the total area can be divided into groups or *clusters* of elements and some of the groups or clusters are selected randomly, we may have cluster sampling. It views the units in a population as not only being members of the total population but as members also of naturally-occurring in clusters within the population. For example, city residents are also residents of neighborhoods, blocks, and *sefers*. The cluster, thus, may be an institution or geographical area or any other appropriate group depending on the nature of the study. Each group is considered as

a *cluster* and its number of elements referred to as *cluster size* – in most cases the cluster sizes among clusters differ.

A convenient way in which a sample can be taken is to divide the area into a number of smaller non-overlapping areas and then to randomly select a number of these areas; and hence the name *area sampling*. When a survey is large scale, however, it may be conducted at several stages called *multistage sampling* – in the first stage, a sample of a set of geographical regions or clusters is taken randomly; and next, a subset of them are sampled within each of those regions, and so on. Finally a sample of elements is drawn from the last stages/sub-clusters until the required group of clusters/sub-clusters is arrived at (it must use a random sampling method at each stage).

This technique reduces cost by concentrating the survey in selected areas. In that respect, it may be less precise than other random samples. While statistical efficiency for cluster sampling is usually lower than simple random sampling chiefly because clusters tend to be homogenous, economic efficiency is often great enough to overcome this weakness and hence this technique of sampling is justified even at the cost of loss in statistical efficiency.

Broadly, it is used when the population is infinite, sampling frame for all the units of the population is practically unavailable or difficult to prepare but boundaries can be well-defined, the geographical distribution of the population is highly scattered and/or sampling individual units is not convenient for several administrative reasons. For example, to obtain information about the drug habits of all high school students in Addis Ababa, you could obtain a list of all the school districts in the city and select a simple random sample of school districts. Then, within in each selected school district, list all the high schools and select a simple random sample of high schools. Within each selected high school, list all high school classes, and select a simple random sample of classes. Then use the high school students in those classes as your sample.

Therefore, if you use this sampling technique, you ought to identify clusters depending on the nature of the study, examine the nature of clusters, determine the number of stages, and select the units accordingly. This may result in a somewhat larger sample than using a simple random sampling method, but it saves time & money and also cheaper to administer.

Advantages of Cluster Sampling:

much easier and more convenient to apply when large populations are studied or large geographical areas are covered;

less costly;

promotes convenience of field work;

units of study can be readily substituted for other units with in the same random section;

flexible as it allows to employ different types of random sampling in successive stages; and

gives the best result if population is relatively evenly distributed over a given geographic area of the study.

Disadvantages of Cluster Sampling:

when cluster sizes vary significantly, bias of the resulting sample may be worrisome; greater sampling error; and

if the population follows a certain pattern in its distribution, cluster sampling may result in unrepresentative sample and hence misleading results – this is because adjacent units of study tend to have more similar characteristics than do units of distantly apart.

Activity 4.4

1. Compare and contrast simple random sampling and systematic sampling in 50 words.

   _____
   _____
   _____

2. Compare and contrast stratified sampling and cluster sampling in 50 words.

   _____
   _____
   _____

3. A student-researcher is conducting a sample survey on *The Determinants of Adoption of Maize Seed Variety in Gamo Gofa Zone* using primary data. What sampling technique(s) do you suggest she should employ? Why?

   _____
   _____
   _____

4.2.2. Non-Probability Sampling Techniques

As noted earlier, a non-probability sampling does not adopt the theory of probability and hence no known chance of elements to be part of the sample – generally considered inferior to probability sampling techniques. Despite accepted superiority of the latter sampling techniques, non-probability sampling techniques may be used when probability sampling is prohibitively expensive and when precise representation is not necessary.

These techniques focus on volunteers, easily available units, or those that just happen to be present when the research is done. They are useful for quick and cheap studies, for case studies, for qualitative research, for pilot studies, and for developing hypotheses for future research. Usually such techniques of sampling are used because of cost and time requirements, if there is no desire to generalize on the population parameters, and the total population may not be available for the study in certain cases.

We shall consider four different techniques of non-random sampling: haphazard sampling, quota sampling, judgment sampling and snow-ball sampling. Other non-probability sampling techniques include library research, participant observation, marketing research, consulting with experts, and comparing organizations, nations, or governments which are less commonly used.

i) Haphazard or Convenience Sampling

This technique involves selecting anyone who/which is convenient, close at hand, easy to reach, etc; like interviewing people whom you happen to meet in a just *hit and miss* fashion and hence also known as *accidental sampling* or *man-on-the-street sampling*. This also means selecting whatever sampling units are conveniently available. It may be used for simple purposes such as testing ideas or rough impression about an issue of interest. At times when one cannot define the target population or sampling frame is not available, this technique may be used.

Advantages of Haphazard Sampling:

cheapest and simplest to use;

does not require any statistical expertise;

can be applied in the absence of sampling frame;

very common in media and other behavioral sciences. For example, journalists may interview people who are convenient for their views on the impact of inflation on their livelihood;

Disadvantages of Haphazard Sampling:

can produce ineffective, highly unrepresentative sample – not recommended in economic research for its biased results and systematic errors;

may involve bias from the researcher's subjectivity;

least reliable as there is no way of estimating representativeness of the sample;

and not easy to generalize the findings

ii) Purposive or Judgment Sampling

When one draws a non-probability sample that confirms to a certain criteria, it is called *purposive sampling*. It uses the judgment of researcher in deliberate selection of cases with specific purpose in mind; for example, students who live in dorms on campus, or farmers using fertilizer, banana producers selling their product in central markets.

This sampling technique is crucial when one wants to make sure the inclusion of *relevant* elements in the sample – it matters to the judgment of the researcher (research team whether or not to include an element in the sample).

Advantages of Purposive Sampling:

less costly and more convenient;

handy method when a researcher is in critical shortage of time;

and guarantees inclusion of relevant elements in the sample

Disadvantages of Purposive Sampling:

does not ensure representativeness;

less efficient especially compared with probability samples;

requires more prior information about the population; and

does not allow the use of statistical inference and hence results can't be generalized.

iii) Quota Sampling

This involves assigning quota (restricted number of units to be drawn) to different groups. For example, to interview 28 shoppers at a mall/market, half of whom are male and half of whom are female (14 each). Although it is an improvement over haphazard sampling, it is also a weak type of sampling technique, at least as compared to random sampling techniques. In this technique, a researcher first identifies categories of units in the population and then decides how many to select in each category. Thus, the number of units to draw in various categories is fixed but gives no assurance that the sample is a representative on the variable being studied. It is usually used in studies like marketing surveys and opinion polls which do not aim at precision. Nothing prevents the researcher from selecting units that act friendly or convenient or want to be part as there is no

random element involved in this sampling technique. As there is no element of randomization, the extent of sampling error cannot be estimated.

This technique of sampling guarantees inclusion of elements from all groups of the target population. Moreover, it saves cost as it does not involve randomization and hence allows the researcher to draw elements in a way that saves resources; hence is less costly. The degree of representation of each group depends on the judgment of the researcher (research team). This may be done either proportionate (when all groups are *equally* important for the research objectives) or disproportionate (when some groups are more important to the research objectives than the others) to the size of the group.

Advantages of Quota Sampling:

less costly and takes less time;

does not require sampling frame; and

easy for field work as it does not require supervisions.

Disadvantages of Quota Sampling:

may not yield a precise result as the sample is unrepresentative;

strict control of field work, if need be, is difficult; and

difficult to apply when dimensions and variable for grouping are large as the degree of classification error would be very large

iv) Snowball sampling or network/chain referral or computational sampling

This involves building up a list or a sample of a special population by using an initial set of its members as informants. It is a method for identifying and sampling or selecting the case in a network. It begins with one or a few people or cases and spreads out on the basis of links to the initial case. For example, when one needs to study the socio-economic problems of Ethiopian females working in Arab States; she needs to find some females as initial informants who shall later point out for more units for sampling or census of the list developed.

Advantages of Snowball Sampling:

very useful in studying social groups, informal groups in a formal organization and diffusion of information among professionals of various kinds; and

useful for smaller populations for which no frames are readily available.

Disadvantages of Snowball Sampling:

does not allow use of random sampling;

difficult to apply for large population; and

does not guarantee inclusion of all elements in the list.

Finally, it should be noted that not only one technique of random sampling may be used at a time for a given research. Rather, a combination of more than one technique may be used especially when there is no single way to sample a particular population. Such a technique is known as *hybrid sampling*. For instance, a sample of farmers selected using a combination of cluster sampling from their woreda and heterogeneous sampling based on their crop type may be considered as a hybrid sampling.

Check Your Progress

4.5 Why do sample surveys dominate research in Economics? What is the rationale of sampling? How does a researcher justify studying only part of the population while her aim is to know about the population? What are the major considerations in deciding to sample?

4.6 Distinguish between random and non-random sampling.

4.7 How and when does one draw a simple random sample? Give illustration please!

4.8 Why is systematic sampling considered as semi-random design?

4.9 What is stratification? How and when does one draw a stratified sample?

4.10 When do we recommend a cluster sampling to be used? Why?

4.11 Why it is not advised to use accidental sampling for economic research?

4.12 'A quota sampling brings better representativeness than convenience sampling' – argue for or against!

4.13 What is snow-ball sampling? When do we use it and how?

Activity 4.5

1 Briefly compare and contrast probability and non-probability sampling techniques:

_____
_____
_____

2 Compare and contrast quota sampling and stratified sampling:

_____
_____
_____

3 Distinguish between stratified and cluster sampling:

_____
_____
_____

4 A researcher realized from literature that farmers in Ethiopia are heterogeneous at a national level and homogenous in the woreda level with regard to their major crop type. If she is

interested to conduct an economic research concerning their major crop, which (combination of) sampling technique(s) do you suggest? Why?

_____
_____
_____

5 A student-researcher is afraid to employ a non-probability sampling for an economic research for her term paper. Why mainly is she curious about using such techniques?

_____
_____
_____

6 A student-research is supposed to cover a wide geographic area. However, she is constrained by time and financial resources.

_____
_____
_____

7 What considerations you suggest a researcher has to make in selecting a (combination of) sampling technique(s)? Compare your answer with the discussion below.

_____
_____
_____

clxvi Dear learners, note that for a researcher to select the relevant sampling technique, she needs to consider the following:

Purpose of the Study: If the interest is generalization, then random sampling technique must be employed as opposed to non-random sampling. If the purpose is to generate rough information on a certain issue of interest, then a non-probability sampling may be used.

Measurability: If research objectives dictate statistical inference which requires computation of sampling error, simple random or stratified sampling must be employed based on the homogeneity or heterogeneity of the population.

Degree of Precision: Where a high degree of precision of results is required, probability sampling should be used. Where even crude results would serve the purpose, any convenient non-probability sampling may be utilized.

Information about the Population: In the absence of information about the population (at least sampling frame), it is usually difficult to apply random sampling technique.

Nature of Population: In terms of the variables under study, a heterogeneous population demands stratified sampling. Where there is a sampling frame or easy to make one and the population has no trending nature, systematic sampling may result in equally good samples.

Geographic Area and Size of Population: When a sample survey has to cover a large geographic area or involves a big population size, (multi-stage) cluster sampling is often used.

Resources: This refers to time, human and financial resource allotted for the survey. These constrain your decision on the selection of the sampling technique(s).

clxvii The task of the researcher, therefore, is to make a *right mix* of these and other variables to optimally choose the sampling technique(s) that bring a *good sample*. What is a good sample? How large is a good sample? Compare your answers with the discussion below:

_____

_____

clxviii Sample Size

There is a widespread but erroneous touching belief of students that the answer to sample size should be *as many as possible* or that the bigger the sample, the better the survey. The right answer is *as many as necessary*. If your survey is badly designed and therefore producing only garbage, then a bigger sample size will only lead to a bigger pile of garbage. Even if your survey

is beautifully designed, then bigger is not always better; there is a point of *diminishing returns* where you will not be learning anything very new from gathering any more data, and where you are starting to waste everyone's time if you continue to gather data.

How do you know if you have reached that point? This is where statistics appear into view – we know from probability theory that if we take a very large number of simple random samples of n students from our student population, and find the average monthly pocket income for each sample, that those averages would tend to distribute themselves in the pattern of a *bell-shaped* curve, also called *the normal curve* which has well-established properties – for example, approximately 68% of the sample averages would fall within plus or minus one standard deviation of the true population average; approximately 95% of the sample averages would fall within plus or minus two standard deviations of the true population average and approximately 99% of the sample averages would fall within plus or minus three standard deviations of the true population average.

Using these established principles, we do not have to take repeated simple random samples (fortunately!). Instead, we can use these principles to estimate how well our sample statistic estimates the population parameter and sample size determination.

Another issue is that sometimes you'll want to leave some people unsampled, so that you can use them for a further piece of research, without affecting their behaviour by surveying them in your first study.

Generally speaking, the size of the sample depends on the type of research design being used; the desired level of confidence in the results; the amount of accuracy wanted; and the characteristics of the population of interest but has little to do with the size of the population As a rule of thumb, the lower the degree of precision required, the lesser the number of groups studied, the lower the budget for the study, etc the lower is the sample size, and vice versa.

To calculate sample size, we must make four decisions:

a. Are we doing a true experimental design (e.g., control-group, pretest-posttest design) or a non-experimental design (e.g., a cross-sectional survey)? The former can use smaller sample sizes, while the latter require larger sample sizes.
b. How sure do we want to be that we could get the same results if we did the study multiple times? Do we want to be 50%, 90%, 95%, or 99% sure? This is called the *confidence level*. The more sure we want to be; the larger the sample size needs to be.
c. How accurate do we want to be at estimating the population parameter? Will a margin of error of (plus or minus) 5% be acceptable or 4%, 3%, 2%, or 1%? This is also called the *confidence interval*. This means that if we find that 66% of the students have their source of

pocket income from either of their parents, we really mean 66% plus or minus 5% (or any other confidence interval).

d. How is the population distributed on the variable of interest? That is, in a yes/no situation, how many do we think will say yes? How many will say no? The most conservative way to approach this is to guess that the population is split 50/50 on the question.

Then, the formula suggested for *sample size* is: $= ()(1 -)$

For example, if a researcher faces a population evenly distributed in the variable of interest, and wants to have 95% confidence level and allowed a 5% margin of error, what is the sample size?

$\sqrt{s} = \frac{1.96}{0.05}$

$\sqrt{s} = \frac{(0.5)(1 - 0.5)}{0.25)}$

$\sqrt{s} = 39.2$

$s = 384.16$ 385 students

Activity 4.6

If the researcher wanted only a 90% confidence level, what is the recommended sample size? How about allowing for 1% margin of error? Hint: t-value for 90% confidence level is 1.645 [Answer: 270.6 271 and 9604, respectively].

Note also that if you have a fixed sample size, you can increase the confidence level and decrease the accuracy (or vice versa) but you cannot do both. Besides, as the variability in the population on the variable of interest increases, the sample size increases. A probability of 50/50 demonstrates the greatest variability in the population. If the variability decreases to 60/40, or 70/30, then a smaller sample size will result. If the researcher wants to study sub-populations as well as the whole population, then larger sample sizes will be needed. In addition, if more than one variable is being studied at the same time, then the *rule of thumb* is to have a total of at least 10 cases per variable.

If the research is to be a controlled experiment, then smaller sample sizes can be used. However, it is recommended to use samples of no smaller than 30 for each group in the experiment (like experimental and control groups). Many common statistics are based on sample sizes of a minimum of 30; for sample sizes of less than 30, other special statistics must be used.

Check Your Progress

4.14 What information do you need to compute the desirable sample size for your research?

4.15 A learner argued that larger samples are always better. How do you criticize this argument?

Characteristics of Good Sample

A good sample must be as *representative* (of the entire population) as possible; and ideally it must provide the *whole* of the information about the population from which the sample has been drawn. It must be unbiased/accurate and efficient. So, the ultimate test of a sample design is how well it represents the characteristics of the population it purports to represent. In measurement terms, the sample estimates must be valid. Moreover, it must be viable in the context of funds available for the research.

Validity depends on two factors: accuracy and precision. Accuracy refers to the degree of closeness to the correct population parameter (bias is absent from the sample). An accurate/unbiased sample is one with the underestimations (negative deviations) and overestimations (positive deviations) balanced among the members of the sample. There is no systematic variance with an accurate sample.

Precision measures the degree to which the standard errors are minimized. No sample will fully represent its population in all aspects. A sample statistic may be expected to differ from its parameters as a result of random fluctuations inherent in the sampling process referred to as the

error of variances or *sampling error*. Precision is measured by the standard errors of the estimates; the smaller the standard error of estimates the higher is the precision of the sample.[7] Sampling error arises from two principal sources: *random error*, and *non-random error*:

> Random error results from taking a sample from a population, instead of measuring the entire population. It is predictable, using probability theory. It is the reason that sample statistics only provide estimates of population parameters, but the amount of random error is known. Non-random error, on the other hand, results from bias being introduced into the sample from some flaw in the design or implementation of the sample. For example, using a telephone book as the sampling frame for all the residents of a city will result in some bias, because some people are not listed in the directory or do not have telephone. People who refuse to take part in a study (which is their right) also may introduce bias into the sample. Some people may provide erroneous information, which also biases the results. Finally, mistakes in computing the required sample size, in identifying the actual units to be included in the sample, or other errors can introduce bias into the sample.

[7] Revisit you Econ-242 materials for details of standard errors

To assess whether an adequate sample was used in your research, ask the following questions:

Size – was the size adequate for the purpose of the study, especially if there were many sub-groups included in the analysis, or many variables used simultaneously?

Representativeness – was the sample selected randomly from the population, using probability theory? Was the sampling frame adequate?

Implementation – was the sampling design carried out carefully, was it adequately supervised, was there some quality control plan, and did it result in a good response rate?

Check Your Progress

4.16 What are the desirable properties of a good sample?

4.17 How do we measure the adequacy of a sample?

4.18 Distinguish between random error and non-random error.

4.19 Distinguish between validity and viability of sample.

Activity 4.7

1. For each of the following, suggest an appropriate sampling design and sample size.

a. A public health official wants to estimate the number of babies who are being born infected with HIV in Ethiopia in a year.

b. The Arba Minch town Meir wants to know, by next fasting season, the extent of depletion of fish resource in Abaya and Chamo Lakes.

2. Briefly discuss qualities of good sampling design and good sample.

_____

_____

3. A researcher wants to increase the degree of precision of the sample estimates. What do you suggest she should do? Why?

_____

_____

Steps in Sampling Design

Learners, I hope, can imagine how difficult a task is to sample a given population. The following guideline may help you ease this activity. Sampling broadly involves:

Identifying the relevant population – this is a very critical step in any research since a study, however fabulous it might be, sampling the wrong population is a mere wandering around the bush;

Determine the method of sampling – this comprises decisions not only on probability or non-probability techniques but also the specific (combination of) technique(s) with all the justifications;

Secure a sampling frame if so required;

Identify variables of interest – what population attributes/behaviors are of interest to the study? This helps determine the appropriate sample size to attain the objective of the study and also for designing instruments of data collection (to be discussed later).

Determine the sample size – this does not depend on homogeneity/heterogeneity of the population, intended degree of confidence (how much precision is required), number of sub-groups to be studied, cost and time factors.

Overall, the task of the researcher is to come up with a representative sample. Hence, the techniques and activities in sampling may differ. The above steps, however, may serve as a guide to major activities in any sampling exercise. Then, a researcher shall evaluate how good the drawn sample is against certain criteria.

Sampling and Non-Sampling Errors and Biases

Recall that a survey aims at estimating or inferring selected population characteristics by studying either the entire population (census) or part of it (sample). Nevertheless, the research results may differ from the true values of the parameters under study. Such differences are known as *errors* and biases. These may be classified into sampling and non-sampling errors (and biases).

1. Sampling Errors and Biases

The errors which may arise because of studying only part of the population are called *sampling errors*. These arise from *non-representativeness* of the samples and the *inadequacy* of sample size. When several samples are drawn from a population, their results would not be identical; and their degree of variation among themselves is measured by standard deviation and is known as *standard error* which is indirectly related to sample size – the larger the sample, the lower the sampling error and vice versa. Learners need note that *error* is not synonymous with mistake.

The average of the estimates of a population parameter derived from an infinite number of samples is called the *expected value* of the estimator. The difference between this value and the 'true value' of the parameter is the *sampling bias* which may arise if the sample is non-random, sampling frame is inaccurate or incomplete, and/or response rate is low. This bias cannot be compensated by increasing sample size.

2. Non-Sampling Errors and Biases

These arise from sources other than sampling. *Non-sampling errors* include errors of observation, measurement, and responses. The implementation of data collection and processing activities are subject to imperfection which cause errors (non-sampling).

Non-sampling biases, on the other hand, pose problems for scientific measurement as they affect both the population and sample values and account for the difference between the population value and the 'true value'(practically usually unknown). This consist of biases of observation (caused by obtaining and recording observations incorrectly) and non-observation (arise from failure to obtain observation from some segments of the population due either to non-coverage or non-response), response biases (arise from imperfections in field observation or interviewing) and process biases (occur during coding, tabulation and computing).

Generally speaking, the measurement of sampling error does not pose much problem, but the measurement of non-sampling errors require special procedures and is costly. Besides, you may see that sample survey involves both kinds of risk, but census has only the latter. But the risk of non-sampling error may be high in census survey. Thus, there are many occasions a sample survey could produce less overall errors than census survey.

 Check Your Progress

4.20 What is the difference between error and bias? What constitutes sampling error and bias and their sources?

4.21 What is non-sampling error and bias? How do these occur in a survey?

4.22 What can a researcher do to minimize non-sampling error? How about sampling error?

4.23 A researcher argues to avoid both sampling and non-sampling errors in a sample survey. Is this possible? Why/Why not?

4.24 How is it possible to have less total errors in a sample survey than that of census of the same target population?

Lesson Summary

Data may be obtained from either or both primary and secondary sources. Data collected afresh for the study at hand by the researcher/research team is known as primary; whereas, secondary data is one that is collected by others for their own purpose.

A survey that treats all the elements of the target population is known as census. If a researcher treats only part of the population, it refers to sample survey.

A survey that treats all the elements of the target population is known as census. If a researcher treats only part of the population, it refers to sample survey.

Data can be acquired from various sources which can generally be categorized in two major groups based on the *originality of data* – primary and secondary. Secondary sources of data include data which has been collected by someone else and which has already been passed through the statistical processes for her own sake; where as primary sources of data are afresh collected by the researcher or her team for the first time (original).

In order to use secondary data for your research, you need to locate, evaluate and verify it. It refers to those that were written or come into being by the people directly involved in the research. These are statistical materials which the investigator or her team produces for the purpose of the research at hand and hence are collected afresh (for the first time)

The difference between secondary and primary sources of data is one of degree – primary data in the hands of one person may be secondary in the hands of another.

A process of gathering information for the purpose of analysis and resolving a problem is generally known as *survey*. A survey that makes a complete enumeration of all items in the population, it is known as *census survey*. The technique of collecting information from a portion of the populations is called *sample survey*.

The logic of the theory of sampling is *induction*; that is, we proceed from particular (sample) to the general (population) and all results are expressed in terms of probability.

Sampling element is the unit of analysis or case in population from which information on variables of interest is collected and provides basis for analysis – it is the subject on which measurement is being taken.

The selection process in sampling is called *sampling technique*. There are two categories of sampling techniques, namely *random* (probabilistic) and *non-random* (non-probabilistic).

Random sampling techniques include simple random sampling/lottery method, stratified sampling, cluster sampling and systematic sampling.

Non-random sampling techniques include convenience sampling, judgment sampling, quota sampling, and snowball sampling.

The choice of sampling technique depends mainly on purpose of the study, measurability of variables, degree of precision desired, available information about the population, nature of population, geographic area and size of population, available resources (time, human and financial resource).

A widespread but erroneous touching belief of students is that sample size should be *as many as possible* or that the bigger the sample, the better the survey. The right answer is *as many as necessary*. If your survey is badly designed and therefore producing only garbage, then a bigger sample size will only lead to a bigger pile of garbage.

If the desired confidence level and margin of error is known, the formula suggested for *sample size* determination is: = ( )(1 −)

A good sample must be as *representative* (of the entire population) as possible; and ideally it must provide the *whole* of the information about the population from which the sample has been drawn. It must be unbiased/accurate and efficient

In measurement terms, the sample estimates must be valid. Validity depends on two factors: accuracy and precision. Moreover, it must be viable in the context of funds available for the research.

Research results may differ from the true values of the parameters under study. Such differences are known as *errors* and biases. These may be classified into sampling and non-sampling errors (and biases).

Sample may be drawn using various techniques broadly divided into probability (simple random sampling, systematic sampling, stratified sampling and cluster sampling) and non-

probability sampling (convenience sampling, judgment sampling, quota sampling and snow-ball sampling).

A researcher has to choose carefully among the various alternative methods of sampling to arrive at a reasonably representative sample for her study.

UNIT FIVE:

DATA COLLECTION

Introduction

I guess you realized that we are now embarking on the *implementation* of the research project – Great! If you recall, we said that research should be empirical. The search for answers to research questions calls of collection of data; and hence every research requires *data*. This necessitates decision-making on the source, type, method of collection, relevance and reliability of data for a given research. This unit basically deals with issues related to data collection.

This unit is organized as

Objectives of Unit Four

After the completion of studying this unit, learners will be able to:

distinguish between primary and secondary data;

explain different types of methods of data collection;

clarify the difference between rationale of different sampling techniques; draw sample from a given population;

elucidate how to design a tool of data collection; and

evaluate quality of data for research.

Lesson One: Data and Its Importance

Description of Lesson One

This unit defines what data is and illustrates briefly on what may constitute data, it shall also see importance of data in research.

Data are facts, figures and other relevant materials, past and present serving as bases for study and analysis. Alternatively, data refers to measurements or observations examined and used to find out things or to make decisions. For example, age of distance learners of Arba Minch University, the opinion of people on HIV prevention methods, general price of goods and services in Ethiopia, etc may be considered as data.

Data serves as a basis of analysis. Without analysis of data, no inference can be drawn on the questions under study. Otherwise, it would be an *arbitrary* guess or imagination of the issue under scrutiny and hence unreliable. Besides, having data doesn't guarantee a valid inference – the relevance, adequacy and reliability of data determine the quality of the findings of a given study. Not also that all data is not important for your analysis and I advise you to be as much precise as possible in your data collection. If you plan seriously and design your data collection carefully, this should not be a problem.

Check Your Progress

4.25 What is data? Give examples of data you planned to collect for your term-paper? In what way does this specific data help you answer your research questions?

Lesson Summary

Data are facts, figures and other relevant materials, past and present serving as bases for study and analysis and its relevance, adequacy & reliability determines the quality of the findings of a given study

Lesson Two: Types and Sources of Data[8]

Description of Lesson 2

Lesson two tries to see the nature and relative merits as well as demerits of primary and secondary sources of data.

Data can be acquired from various sources which can generally be categorized in two major groups based on the *originality of data* – primary and secondary. Secondary sources of data

[8] Revisit your 'Econ-242, Introduction to Business Statistics' materials for better understanding of this unit.

include data which has been collected by someone else and which has already been passed through the statistical processes for her own sake; whereas primary sources of data are afresh collected by the researcher or her team for the first time (original). Thus, studies made by others for another purpose represent secondary data – this may include published and sometimes unpublished materials. Accordingly, data obtained from primary (secondary) sources is referred to as primary (secondary) data.

Activity 4.8

1. A learner doing her term-paper is collecting data from unpublished works of MSc students of the Department of Economics. Is this data primary or secondary? Why?

clxx _____
_____
_____

2. A learner argues that primary sources of data and hence primary data are invariably reliable than secondary ones. Do you agree? Why or why not?

clxxi _____
_____
_____

3. A learner-researcher competing for a research grant has designed her methodology in such a way that she collects data from regular students of the Faculty of Business and Economics by distributing questionnaire and having a discussion with selected students from each department. What kind of data is expected – Primary, Secondary or Both?

_____
_____
_____

clxxii Secondary Data

We said that secondary data refers to data that were collected for other studies. Data collected for the first time by the first researcher is primary; but for the other researchers, it shall be considered as secondary. There are enormous amounts of data that are collected everyday by government agencies, universities, private organizations, profit and non-profit organizations, public opinion polls, and students. Some examples include the labor force survey of Central Statistical Agency, annual budget reports of the Ministry of Finance and Economic

Development, quarterly financial sector reports of the National Bank of Ethiopia, Ethiopian economy databases of the Ethiopian Economic Association, the Water Journal of the Arba Minch University, and others.

In order to use secondary data for your research, you need to locate, evaluate and verify it.

i) **Locating the Data**: Secondary data can be located by using printed or electronic indices, such as the bibliographies of books and journals or internet search engines (like Google, Yahoo, hi5, etc). There are also on-line databases of secondary data, for example, in the CSA, IMF and WB home pages.

ii) **Evaluate the Data:** Evaluating secondary data is similar to doing a critique of a published research report. Everything about the original project that produced the data should be scrutinized to ensure that the project had high validity and reliability, such as:

theoretical or conceptual model used,
variables and hypotheses posited,
operational definitions of variables and measures employed,
the population, sample frame, sampling design, and sample obtained,
the data collection method and response rate obtained,
quality control measures employed,
data coding, data entry, and data analysis procedures utilized, and
factors which could have affected the study, such as current events.

iii) **Verify the data**: If the data seem valid and reliable, you need to make sure that you have an accurate copy of the data, especially if you obtained it through an electronic medium. This includes verifying that you:

have proper documentation
have the correct number of observations or
cases have the correct number of variables
have the correct coding scheme
can reproduce the original summary statistics

Check Your Progress

4.26 Please answer the following questions and get the comment of other learners:

4.2.1 What is the danger of using secondary sources of data for your research?

4.2.2 How do you minimize potential dangers of using secondary data?

4.2.3 What should one do if has to use secondary data?

4.2.4 What do we mean by ‘classification of data into primary and secondary is that of degree’?

Why Use Secondary Data?

One may gain the following advantages/merits by using secondary data for her research:

It is found more quickly and cheaply;

Helps much on past events and distant places, where the researcher cannot afford to collect the data afresh;

It can be less expensive than gathering the data all over again;

It may allow the researcher to cover a wider geographic or temporal range; It can allow for larger scale studies on a small budget; and

It does not exhaust people's goodwill by re-collecting readily available data.

However, if you decide to use any secondary data source for your term-paper/research, you need to check whether:

the nature of the data is consistent with the need of the ongoing research - data may have been intended for consumption by particular groups, which differ from the present project;

the data is adequate for the study;

the data satisfy the demand of the study being undertaken;

the units and definitions are in line with the interest of the study or can easily be converted; the time and geographic coverage address the needs of the study;

the data is reliable;

the institution is dependable – skilled and trained staff, acceptable data collection instruments and reliable department processing the data, efficient reporting system, etc;

the members of the team of data collectors well-trained;

the instrument of data collection up to the standard;

the data is timely and relevant and so on;

the conditions that led to their production are unbiased – it may not be free from personal bias and prejudices;

data may have been originally gathered to persuade, justify, or otherwise convey a particular point of view;

data may have decayed over time, been censored or purged

If the research team belongs to an institution, secondary data sources can be classified as either *internal* or *external*. Internal data come from internal sources related with the functioning of an

organization or firm where records regarding staff profile, purchase, production, sales, profits, etc are kept on a regular basis. However, internal data is usually either insufficient or inappropriate (or both) for the enquiry into a phenomenon under scrutiny. In such cases, we [additionally] need external data.

External data sources include all sources outside the institute, such as books, periodicals, government documents, reference books, encyclopedia, journals, magazines, reports of other institutions, etc. For example, a *team of researchers* in Arba Minch University doing a study, say, on 'Determinants of academic performance of students' may use part of the data from the Office of Registrar of various higher institutions in Ethiopia. Data obtained from the Arba Minch University Registrar may be considered internal where as data obtained from other Universities' Registrar Offices is external. Recall that the team of researchers (or the research) belongs to the Arba Minch University; hence is *internal* to this specific research (this is why the Arba Minch University Registrar is considered as internal source of data).

Potential Drawbacks of Secondary Data

Secondary data are only as good as the research that produced them. When you use it, you must assume what the author(s) meant by the terms they used – adopt the terms and definitions used in the author. Unless handled carefully, the sub-culture references, jargons, or idiomatic expressions may be misleading. Worse is, some data may be neither valid nor reliable for instruments or data collection methods may have changed over time, data may have been modified by the researcher already (e.g., weighted averages), the data is ill documented, electronic format incompatibilities, etc

Other factors that affect the use of secondary data include limited access to the data [for example, on-site access only], confidentiality considerations that lessen its usefulness and substantial purchase or loan cost. As stated above, the information may not meet ones specific needs since it is collected by others for their own purpose – definitions, units of measurements would differ, and different time periods may be involved. Besides, it is difficult to assess the accuracy of the information because one knows little about the research design or the conditions under which the research or its data collection took place. Finally, it a serious problem when secondary data used in a research is outdated. These all imply the need to check secondary data for potential merits and demerits before we use them.

Yet, if data from secondary sources do not satisfy the interest of the study, it necessitates collecting required information afresh from primary sources. Thence, primary and secondary data may be used simultaneously jointly or individually depending on the demands of the research and judgment of the researcher.

clxxiii Primary Sources

It refers to those that were written or come into being by the people directly involved in the research. These are statistical materials which the investigator or her team produces for the

purpose of the research at hand and hence are collected afresh (for the first time); hence these happen to be original in character. If the same data is used by other (team of) researchers for their own sake, nonetheless, the data would be called secondary while it remains to be primary for the firs producers/collectors. This, therefore, implies that the difference between secondary and primary sources of data is one of degree – primary data in the hands of one person may be secondary in the hands of another.

Activity 4.9

1. From where and how do you think one collects primary data? Please write your answer on the space provided below and compare it with the discussion that follows.

_____

_____

2. Suppose a researcher collected data for a research and used the same data for another related research. Is the data considered as primary or secondary for the second research? Why?

_____

_____

Primary data may be collected through a variety of *methods* including personal observation, interview, questionnaire, etc. Since these methods are relatively costly many research involve sampling the population under study. However, as shall be discussed soon, cost is not the only reason for sample survey as opposed to census – what do we mean by these?

Lesson Three: Methods of Data Collection

Description of Lesson Three

Lesson Three discusses the methods of data collection, their relative merits and demerits, their application, and introduces the corresponding tools of data collection. It also compares and contrasts between the alternative methods of data collection including activities and exercises.

Dear learners, did you manage to decide on your sampling design? I hope you did! Let me suppose you are using primary data for your research (term-paper) and try scrutinize the various methods of [primary] data collection to help you discriminate between them.

Using the precept that no data is better than bad data, collecting bad data cannot yield good results. Likewise, there is no statistical test that will overcome bad research design and no statistical test that will overcome bad data. It is, therefore, important from the beginning to

design the method(s) of data collection carefully. Method refers to the way or mode of gathering data while a tool is the instrument used for the method. While the former is the concern of this lesson, the latter is the concern of the next lesson.

One thing from which you are unlikely to suffer when it comes to data collection is lack of choice of methods. The most common methods include observation, interviewing, mail survey, experimentation, simulation and projective techniques; the choice of which depends, among others, on nature of the study, unit of enquiry, size and spread of the sample, scale of the survey, educational level of respondents, type & depth of information to be collected, availability of skilled labor power, and desired accuracy of data. Learners shall note that we do not have the luxury of discussing all the methods in this module but only more relevant methods in economic research.

4.2.1. Observation

Observation means viewing or seeing – a systematic viewing of a specific phenomenon in its proper setting for the specific purpose of gathering data for a particular study. This, as a method, may include both seeing and hearing and not asking questions. It is widely used in behavioral sciences. In the context of studies in social sciences, we often talk about participant and non-participant forms of observation. The former refers to the case where the observer makes herself, more or less, a member of the group she is observing and experiences the phenomenon with the group be it in the knowledge of other about her presence (open participant observation) or otherwise (disguised participant observation). The latter, on the other hand, applies to cases where the researcher/observer acts as a detached emissary without any attempt to her part to experience through participation what others feel.

Advantages of Observation:

Enables to record the natural behavior of the group;

Helps verify the truth of statements by informants in other methods of data collection – helps in data triangulation;

May provide context for the observer to give meaning to observed behavior and heard statements;

Has virtue of *directness* – makes it possible to study behavior as it occurs;

More suited to studies whose subjects are unable to articulate meaningfully (for example, studies on children, tribal groups, animals, etc);

Less demanding and less biasing effect; and

Allows the use of mechanical devices for recording; like camera, tape recorder, etc

Disadvantages of Observation:

It cannot be used indiscriminately for all purposes; for instance, it is not applicable to past events;

Not suitable to study opinions and attitudes;

It poses difficulties in obtaining a representative sample;

It cannot be used as and when the researcher finds it convenient to use it but has to wait for the event to occur;

A major limitation is that the observer/researcher must be at the scene of the event when it takes place, especially when it is difficult to predict when the event will occur; and

It is a slow and expensive process, requiring human observers and/or costly surveillance of equipments.

4.2.2. Interview

It is a two-way systematic conversation between an investigator and an informant, initiated to obtaining information relevant to a specific study. This, therefore, involves presentation of oral-verbal stimuli and reply in terms of the same either face-to-face or via telephone. This calls for interviewing skill and is done by using a structured schedule or unstructured guide. This method is the only method of collecting primary data from illiterate or less educated respondents and applies to surveys that cover a compact area or having adequate number of qualified interviewers. It is usually superior to other methods as people are usually more willing to talk than to write. With probing and efforts of interviewers, it may yield even confidential information. It also allows the investigator not only to grasp the behavioral context of data furnished by respondents but also seek clarifications when necessary.

4.2.3. Advantages of Interview:

The depth and detail of information that can be secured is larger than that of other methods;

Interviewer(s) can do more things to improve the response rate and quality of information received;

It gives greater flexibility, esp. in the case of unstructured

interview; It may enable to obtain personal information easily;

The language of the respondents cab be adapted to the ability or educational level; and

Interviewers can collect supplementary information about the respondent's personal characteristics and environment which is often important in interpreting results.

Disadvantages of Interview:

Very expensive for sample scattered geographically;

It may yield biased information – interviewers or respondents bias. For example, either or both party may be over-stimulated and end up with imaginary data;

It may be difficult to reach some potential respondents; like important officials, business people, etc; and

When it is telephone interview, it requires the existence of the facility for both interviewer and interviewee.

4.2.4. Mail Survey

This is also known as self-administered method of data collection and is very common in economic research. In this method a *questionnaire* is sent (usually by mail) to individuals who can read and write and selected in your sample with a request to answer the questions.[9] However, if respondents cannot read and understand it, others may read for them and write down their answers – such a method is known as *schedule* method.

Advantages of Mail Survey:

Lower cost than other methods even for a large universe and evenly distributed geographically;

It is free from interviewer's bias and respondents may have adequate time to give well-thought answers;

Persons who might otherwise be inaccessible can be contacted; It

provides more anonymity than interview method; and

Respondents may complete the questionnaires at their convenience.

Disadvantages of Mail Survey:

The scope of application is limited especially in areas with high illiteracy rate;

A major challenge in this method is low response rate;

Non-response error (incomplete answers) is expected and is serious when big;

Additional information on the personal characteristics of respondents cannot be secured; It takes longer time and may be slow; and

May result in biased outcomes if the questionnaire is not well-designed.

[9] A questionnaire is a form or set of forms that consists of list of questions printed/typed in a definite order to respondents with spaces to answer the questions.

Quite often, this method is considered as the heart of a survey operation especially in economics. Besides, for most [student] research using primary data the method is widely used. Hence, lesson four shall deal with how to develop a tool of data collection with emphasis on questionnaire.

Lesson Summary

Data may be collected using different methods including observation, interviewing and mailed questionnaire.

Mailed questionnaire method is popular in economic research for its cost advantages.

Mailed questionnaire method may involve low response rate and non-response errors which may, however, be improved if efforts are committed.

Check Your Progress

4.27 Compare and contrast observation, interviewing and mailed questionnaire method of data collection.

4.28 What conditions must be satisfied for one to conduct:

Observation

Telephone interview

Face-to-face interview

Mailed questionnaire

Lesson Four: Tools of Data Collection

Lesson four discusses how a tool of data collection in general and questionnaire in particular is designed. It tries to see how to ask questions to draw relevant information for your research and illustrates the issue using questionnaire.

Dear learners, fortunately there is another challenge in this sub-topic to address the issue of how to ask good questions to collect appropriate data. This shall be addressed using questionnaire as one of the most commonly used tools of data collection. A questionnaire is defined in Webster's New Collegiate dictionary as *a set of questions for obtaining statistically useful or personal information from individuals*. It translates the research objectives into specific questions. The answers to those questions provide the data for testing the research hypothesis. Questions must also interest the respondents enough that they will provide the information.

Obviously, an effective questionnaire is much more than that since poor questionnaire design contributes to non-sampling errors; specifically response errors, the questionnaire should be well designed. The reliability and validity of survey results are dependent on the way the specific questions are planned, constructed, and executed. Hence, the questions should minimize the possibility that respondents will give inaccurate answers. To this end, we try to the main aspects of questionnaire design: general form, sequence, type and question wording. Let me forward a piece of advice on the major steps in questionnaire design first:

Determine data need through a deeper analysis of the research objectives, hypotheses, and your operational definitions (if any). You need to decide on what data is required to answer the research questions, to test the hypotheses

Prepare a *dummy* table that displays the data to be gathered to ensure adequate coverage of the data needs of the study. This helps avoid duplications and omissions, make appropriate additions, corrections and deletions while designing.

Determine respondents' level of knowledge, experience, language ability, etc in order to choose appropriate words and concepts in the questionnaire.

Choose the type of tool to design: questionnaire, observationnaire, interview guide (structured or unstructured), schedule, etc.

Draft the tool of data collection such that you:

lay a broad outline of listing broad categories of data (for example, demographic characteristics, consumption, income, etc).

sequence these groupings in a logical order – usually from simple and less sensitive to complex and more sensitive questions.

design the specific questions under each group making use of the dummy table prepared.

Evaluate the draft tool and get the comment of others preferably qualified people. If you cannot find one who can comment on your draft, you may do it yourselves based on these criteria:

the relevance of each question to the research objectives

appropriateness of the type of question (mainly, open-ended or close-ended)

clarity and unambiguity of the question

practicability of the question (it should not demand knowledge and information respondents do not possess or remember)

validity of the question – it should be neither leading nor loaded

check the logical and psychological order of the questions

pre-test the revised questionnaire to identify the weaknesses and make the necessary further adjustment

specify procedures and instruction on its administration – mode of answering, ensuring anonymity, covering letter, etc

design the format in a way it suits the needs of the study – divide the questionnaire into different sections relating to the different aspects of the problem with each section header

printed in **bold**, allowing adequate space for open-ended questions, and reasonable line spacing between questions

Then, we construct questions based on the *dummy table* considering the relevance and content of questions, the wording, response form, and the order/sequence.

Relevance of questions is measured by whether it yields significant information for answering the research questions. Relevant questions shall be of proper scope in providing information needed to interpret the response fully. Question construction ought to consider the respondents' information level and recall ability especially for event in the past. Balance between generality and specificity is also important.

Question wording is a critical element of your construction. We should choose a shared vocabulary – common to you and your respondents. The words shall also have a uniform meaning to all respondents. Make sure the wording is neutral – free from the distorting influence of fear, prestige, bias or emotion. When you ask, do not assume respondents know key words, phrases or concepts. In some [sensitive or embarrassing] areas, you may need to use indirect questioning; like 'what would people say about…' instead of 'what would you say about…' as in the former the respondent will likely reply in terms of her own attitude and experiences freely than the later. Unless you have information on your respondents do not presume anything; try having a filtering question rather. For instance, before asking 'are you satisfied with ETV transmissions?' you got to ask/filter whether the respondent watches ETV. I would like to advise you not to ask hypothetical questions, like 'what would you do if…' as such questions are of very little value for your study. Finally, I would also like to advise you to make your questions brief and simple and put in affirmative/positive than negative.

The response form of questions refers to the type of questions classified into *open-ended* and *close-ended*. Close/structured questions may be dichotomous (two choices) or multichotomous (many choices). The choice between these depends inter alia on your objective, respondents' information level, motivation of respondents to express themselves (how sensitive the issue is), and your awareness of these factors. If your objective is classification of responses, closed questions are appropriate; but if you are interested in suggestion of (or more information from) respondents, then opened questions are appropriate. Besides, opened questions are good in measuring the information level of respondents than are closed questions. It is advised to have some opened questions around the beginning of

your questionnaire as these may arouse interest of respondents. In closed question forms, you have to either make sure the list of alternatives provided is exhaustive or include the category ‘others, please specify’ in order to accommodate unforeseen responses. On top of this, we have to make sure overlapping choices do not exist, fair balance between choices and all alternatives are reasonable.

Table 4.1 Comparison of Response Forms: A Summary

Criteria	**Closed-ended**	**Open-ended**
Researcher's objectives	Confirm or deny the researcher's hypothesis	learn about how respondents come to possess a particular point of view
Level of information needed	Never know how the answer was arrived at	Obvious shortcomings in the respondent's knowledge become apparent
Influence on responses	Response choices may influence respondent's memory; may produce less unaware	No fixed response alternatives can jog respondents who were previously information but attitude formation is not affected by the response options
Amount of effort required	Easy to reply with few communication problems	More frequent refusals to answers or "don't know" responses are possible? Also a chance to educate the respondent

Generally, in designing your questionnaire, you must **try avoiding** the following type of questions:

Leading questions: questions worded in such a way as to influence the respondent to give a certain answer. For example, ‘Don’t you think the President’s speech is interesting?’, ‘You would like to own a new *habasha* dress for Easter, wouldn’t you?’

Loaded questions: questions that contain emotionally colored words and suggest an automatic approval or disapproval. If you ask students ‘Have you ever tried to get special favor from your instructor by pressurizing them one way or another?’, then no student can answer *yes*.

Ambiguous questions: questions with unclear meaning or mean different things to different people. Usually questions having subjective words invite different understanding by different respondents.

Double-barreled questions: questions with two or more different ideas with possibility of different answers shall not be barreled as one. If you are asked 'Whether the quality of modules and tutorial for the courses in this semester are up to your expectations', then you may have different answers for the quality of module and tutorial sessions. Therefore, a single question may not address respondents' true feelings.

Long questions: when questions are long, they often become ambiguous and confusing. Try keeping questions simple and short.

Double negative: usually, questions with double negative are confusing; hence, you are advised not to include such questions.

Check Your Progress

4.29 How do you check the relevance of questions to include in your questionnaire?

4.30 If one is not sure on the exhaustiveness of the list of alternatives in closed questions, what do you suggest she should do about it?

4.31 After drafting a questionnaire the designer wants your advise on final screening of the questions. What kind of questions would you screen out or amend?

Very good, you are now having a draft tool of data collection. The next step would be to check the logical and psychological sequence of the questions. Broadly speaking, it is advised to start your questionnaire with motivating questions, followed by non-sensitive, less sensitive and finally sensitive questions. This helps you win the attention and interest of the respondents and build confidence on the part of the respondents. Besides, you shall include a statement stating the anonymity of any information provided and guarantee her/his personality shall not be revealed in public. Moreover, the answers to the questions shall not be influenced by the content of the preceding questions.

The final step of your design exercise is pre-testing your tool of data collection. If this is done on the actual (or similar) population at a small scale, it is known as pilot survey. This helps in any one or more of the following ways before we embark on the large scale actual survey:

Provides a better knowledge of the problem under study and its

dimensions. Provides guidance on conceptualization

Helps in discovering the nature of relationship between variables and hypothesis formulation.

Shows the nature of the target population and may help in more efficient sampling design.

To check whether the available sampling frame is adequate, complete, accurate, up to date and convenient.

To check the adequacy of the tool and make necessary adjustments.

Guides efforts to develop appropriate plan of analysis.

May give information on the probable cost and hence financial implications as well as the potential problems we may face in the field so that we make the necessary preparations.

Finally, quality of the tool and its delivery are crucial in your research because:

response quality can be improved by design (to decrease respondent fatigue): limit the length;

clear, concise, & unambiguous items;

items well-informed and not biased;

response categories are appropriate;

items arranged by topic;

good quality paper, attractive layout,

numbered items with ample spacing;

offer incentives, if needed.

Response rate can be improved by delivery:

use a cover letter indicating the importance of the project and the value of the respondent's participation; offer incentives and/or feedback;

include a postage-paid return envelope

use follow-up contacts by letter or telephone to increase the response rate. However, if there is a guarantee of anonymity, then follow up must be done with the entire sample, since it cannot be known which persons returned the survey and which did not. In this regard, there may be a trade-off between anonymity and response rate.

Activity 4.10

1. A student-researcher needs you advice on what to consider in designing a questionnaire. Write your advice on a page and read it again yourself.
2. A researcher is worried about the response rate of her questionnaire. What should she do not to suffer this problem? State the pre and post design activities she might do to mitigate the problem.

Lesson Summary

Questionnaire design may involve the following steps:

Determine the specific information needed to achieve the research

objectives. Identify the sources of the required information

Choose the method of administration that suits the information required and sources of information

Determine the types of questions to be used and form of

response Develop the specific questions to be asked

Determine the sequence of the questions and the length of the

questionnaire Predetermine coding

Pretest the questionnaire

Review and revise the questionnaire

Review Exercises

1 Distinguish between primary and secondary data

2 A researcher drew a sample of 100 households using multi-stage sampling technique from SNNP region for a certain study. If she selected 10 households from randomly selected 50 woredas drawn from 5 zones. What is:

   a. The sampling element?
   b. The first stage sampling unit?

3 A student researcher went to a microfinance institution and interviewed people on the spot about the behaviour of people towards the use of the services provided by the institution. How do you comment on the representativeness of this sample?

4 A researcher is worried about the vast area her researcher topic covers. How do you suggest should she draw the sample if hers is a sample survey with the aim of saving time and money?

5 A researcher hides herself in the target population (fixed income earners) in order to collect data on the consumption behaviour of people before and after salary day. What is the method of data collection?

6 If one includes a question in a questionnaire like *'Are the transmissions of ETV entertaining and educative? Yes; No'*, what is the likely weakness of this question? What is the solution for such a problem?

7 A learner has collected information on the following variables of interest from a sample drawn out of her target population. Group the data in the categories given in brackets:

   a. Age of respondents [qualitative; quantitative]
   b. Religion of household head [qualitative; quantitative]
   c. Monthly household expenditure on food [cross-section; time series]

8 What is the use of data?

UNIT SIX

DATA PROCESSING AND ANALYSIS

Introduction

Hello learners! I suppose you have a good understanding of data collection techniques and tools. I also presume you are reading this unit coming back from your data collection field work. What are you going to do with your data? The row data you collected from primary and secondary sources using the appropriate methods and tools is of no use for your study unless processed to convey the messages in a way understandable to the readers.

This requires editing, classifying, and presenting the data, analyzing it and communicating the result to others via reporting. This unit, therefore, shall deal with the research activities after data collection.

The rest of the unit is organized as follows:

Objectives of Unit Five

After studying this unit, learners will be able to:

explain the different steps involved after data collection

to edit row data and decide how to present classified data in different ways

elucidate the techniques of writing a research report and explain the different elements of a research report

write a reasonably acceptable report for their term paper

evaluate the contents of a technical report for academic purposes

Lesson One: Data Processing and Interpretation

Description of Lesson One

This lesson show how collected rata is processed and interpreted to achieve research objectives. Row data, unless processed to answer the research questions, is useless. Hence, in this lesson, we try to see how we process row data, basic steps in data processing and techniques of interpretation.

The data you just collected from primary (and secondary) sources is known as *raw data* as it cannot right away give answers to the research questions and defy comprehension; rather it has to be processed to meet the research needs. Processing eliminates the irrelevant details and their significant features standout prominently. It is only after a proper processing of the data that the analysis is carried out efficiently which involves editing, coding, summarizing and presenting/tabulating.

clxxiv Editing

The first step of processing your data is to ensure that it is *clean* – examining whether it is free from incompleteness, inconsistencies, outliers, contradictions, errors, omissions, misclassifications, etc and to correct these when possible. This may be done at two stages: in the field (field editing) and in the office (central editing). Dear learners, when you do editing, I advise you not to erase original entries given by respondents but just crossed out with a single line and your corrected entries in a different color.

Field Editing is very common for interview and schedule method of data collection. The enumerators/data collectors shall edit while in the field and make the necessary rectification. Because it is done in the place of data collection/field, we call it field editing. Nonetheless, enumerators should be trained enough to do field editing.

Office Editing is done almost in every survey. It is editing at the center of the research either by the researcher(s) or by a team of editors depending on the scale of the survey. In large-scale surveys, editing may be done in different rounds by different groups of editors. It may be done either by examining answers to one question or variable at a time or/and examining answers to all questions by a respondent at the same time. The latter provides a total picture of the responses and helps assess internal consistency.

Having cleaned the data, the next step is a pre-requisite of summarizing and tabulation, namely coding.

clxxv Coding

It refers to the process of assigning numerals or other symbols to answers so that responses can be put into a limited number of categories or classes that may be appropriate to the research problem at hand. It is necessary for efficient analysis. This calls for pre-coding the questionnaire when designing so that this step is simplified. Otherwise, it may involve developing a code book, pre-testing the code book, coding the raw data and verifying the coded data. A code book provides a set of rules for assigning numerals to answers obtained from respondents. If one is using softwares for data processing s/he must have an understanding of the working of the software before coding. Besides, the way you wanted to communicate your findings to readers also matters to your coding. For instance, a closed question on age group classified, say, as 0-15, 16-30, 31-50, and above 50 may have its coding as 1, 2, 3, and 4 respectively. For less

cumbersome data, however, coding may be less important and may result in mere waste of resources.

clxxvi Classification

Classification of data is comparable to the sorting operation; i.e. it is a function very similar to that of sorting letters in a post office. It is the process of arranging things (either actual or notational) in groups or classes according to their resemblance and affinities/closeness. It is grouping of related facts into different classes. Facts in one class differ form those of other class with respect to some characteristics called bases of classification that depends mostly on the nature of the data and the purposes the data have to serve. If the interest of the researcher in a certain study is income level, then classification with respect to this variable may be done as 0 500 birr, 500 1500 birr, 1500 5000 birr and above 5000 birr based on the nature of the research. In this case, the classification base is income level of the units in the study. Such type of classification based on the magnitude of the variable is known as quantitative classification and is most widely used. Broadly, data may be classified on the following four classification bases: spatial, chronological, qualitative and quantitative.

Dear learners, we may classify a hypothetical data on agricultural subsidy in Ethiopia as given in table 5.1 to illustrate other classification bases: spatial/geographical, chronological and qualitative.

In spatial classification [see 2nd row of Table 5.1] data is classified on the basis of geographical or location differences between the various items; that is, area wise (such as cities, kebeles, woredas, regions, etc).

Table 5.1 Government subsidy for agricultural inputs in selected regions (‘000 birr): 2006-2008

Year	Type of Subsidy/ Recipients’ Major Activity*	Region											
		Tigray			Amhara			Oromia			SNNP		
		tion	ck	g	tion	ck	g	tion	ck	g	tion	ck	g
2006	Fertilizer	21	0	10	21	0	10	10	21	0	10	0	21
	Cattle Hybrid	0.4	8	3	0.4	8	3	3	0.	8	3	8	0.4
	Pesticide	3	0.	6	3	0.	6	6	3	0.5	6	0.	3

	Others	1	0.	2	1	0.	2	2	1	0.5	2	0.	1
2007	Fertilizer	10	10	21	0	10	21	10	10	21	0	10	0
	Cattle Hybrid	3	3	0.4	8	3	0.4	3	3	0.4	8	3	8
	Pesticide	6	6	3	0.5	6	3	6	6	3	0.5	6	0.5
	Others	2	2	1	0.5	2	1	2	2	1	0.5	2	0.5
2008	Fertilizer	21	0	10	21	0	10	10	21	0	10	10	10
	Cattle Hybrid	0.4	8	3	0.4	8	3	3	0.	8	3	3	3
	Pesticide	3	0.	6	3	0.	6	6	3	0.5	6	6	6
	Others	1	0.	2	1	0.	2	2	1	0.5	2	2	2

* This refers to the major and dominant economic activity in the production region and the type of subsidy for the specific purpose of agricultural input.

Source: CSA, 200x: pp

Chronological classification [see 1st column of Table 5.1] is made based on time and is common in time series analysis and panel studies. In cross-sectional data, however, there is no such classification as it is a snap shot outlook and does not involve time.

Finally, in qualitative classification data are classified on the basis of some attribute or quality such as sex, color, or hair, literacy, religion etc. The type of data used here are usually those who cannot be measured and one can only find out whether it is present or absent. When classification involves more than one variable, we call it cross-classification.

I hope, learners noted that identifying the classification base(s) is the most important decision a researcher has to make in data classification and that the decision depends on the needs of the research.

clxxvii Data Presentation[10]

Once the classification is done, learners have to choose the alternative data presentation techniques – table, pie chart, bar graph, line graph, pictogram, frequency polygon, ogive, etc. It conserves space and reduces the need to have explanatory and descriptive statements; facilitates comparison activities; expedite the summation of items and detection of errors and omissions; and provides bases for various statistical analyses.

[10] For details of data presentation techniques please refer back your materials on 'Introduction to Business Statistics, Econ-242' or any other book on statistics; like Gupta, 1994

You need to be careful in making this choice as they are different in terms of their efficacy in conveying the finding of your study. Besides, you need to make sure the elements are included properly in your data presentation. Because tables and graphs are very common techniques of data presentation, I dare to ask you what you think are the major elements of these. Please write your answers on the space provided below and compare your results with the discussion that follows.

_____

_____

_____

Dear learners, we may consider only the most common devices of data presentation, namely table and chart. One of the simplest and most revealing devices for summarizing and presenting data in meaningful fashion is *table*. A table is a systematic arrangement of statistical data in columns and rows. Rows are horizontal arrangements whereas columns are vertical ones. The purpose of a table is to simplify data presentation and to facilitate comparisons. The various parts of a table may vary from case to case depending upon the given data. But a good table contains the following parts: table number, title of the table, caption, stub, body, head note, and footnote.

In **the context of Table 5.1** above, these parts of a table may be discussed and illustrated. The discussion given in bold is an illustration from the above table to help capture the elements of a table. Table number [**Table 5.1**] helps easy identification from among other tables in your report and shall assume numbering in line with the chapter it belongs to. In chapter three, for instance, you may start numbering tables as Table 3.1, Table 3.2, and so on. Title of a table [**Government subsidy for agricultural inputs in selected regions of Ethiopia ('000 birr): 2006-2008**] must be suitable to describe its contents and shall answer the questions what, where and when in that sequence. It should be clear, brief and self explanatory. Caption [**Tigray, Amhara, Oromia and SNNP as well as Crop Production, Livestock and Mixed Farming**] of a table refers to column headings with their units. If different columns are expressed in different units, the respective units must be specified accordingly. Stubs of a table [**2006, 2007 and 2008 as well as Fertilizer, Cattle Hybrid, Pesticide and Others**] are the designation of the rows or row headings. The body of a table [**all the cells entered with data**] contains the numerical information. Data presented in the body arranged according to descriptions are classifications of the captions and

stubs. Head notes [**'000 birr**] are used to explain certain points relating to the whole table that have not been include in the title, nor in the captions or stubs. The units of measurement, if are common, are frequently written as the head note. Footnotes [**This refers to the major and dominant economic activity in the production region and the type of subsidy for the specific purpose of agricultural input**] are used to explain any thing in the table which the reader may find difficult to understand from the title, captions and stubs. If footnotes are needed, they are placed directly below the body of the table. Generally, they are used to clarify any thing in the table. And finally, source [**CSA, 200x: pp**] of the table indicates where the data is taken/computed from. When data is compiled from the primary data you collected, you may state the source as *own survey*.

Activity 5.1

1 In the context of table 5.1 rewrite the elements of a table and briefly describe each of them
2 In the process of classification of data, what decision does one have to make?

One of the most convincing and appealing ways in which data may be presented is through *chart* because pictorial presentations help in quick understanding of the data. Furthermore, charts have greater memorizing effect as the impressions created by them last much longer than those created by the figures. A chart can take the shape of either a diagram or a graph.

For representing data, *diagrams* are more commonly used than graphs. The following general *rules* should be observed while constructing diagrams:

Every diagram must be given a suitable *title* in as few and explanatory words as possible. A proper proportion between width and height of the diagram should be maintained.

Selection of appropriate scale that shows values, preferably in even numbers or in multiples of five or ten to make further analysis at ease.

Footnotes that clarify certain points about the diagram and may be given at the bottom of the diagram.

An index illustrating different types of lines or different shades, colors, should be given so that the reader can easily make out the meaning of the diagram.

Diagram should be absolutely neat and clean.

Diagrams should be as simple as possible so that the reader can understand their meanings clearly.

The independent variable should always be placed on the horizontal

axis. The vertical scale should always start at zero.

All axes should be clearly labeled and should clearly state both the variable and the units in both axes – what is measured on the x-axis and y-axis and what are the units?

Although many types of diagrams are used practically, the following are the most commonly used types of diagrams: bar chart, pie chart, pictograms, frequency polygon and cartograms. Nonetheless, presenting the data alone only makes the data amenable for analysis and does not complete the interest of the study per se. Hence, the data presented requires *analysis* and *interpretation* to answer the research questions.

Check Your Progress

5.1 Briefly discuss the major steps in data processing

5.2 Distinguish between field editing and central editing

Data analysis is the most skilled task of the research process and refers to the computation of certain measures along with searching for patterns of relationship that exist among data groups. It is a critical examination of the assembled and grouped data for studying the characteristics of the object under study and for determining the patterns of relationships among the variables relating to it. Particularly in the case of surveys, it involves estimating the values of unknown parameters of the population and testing of hypotheses for drawing inferences. Analysis, therefore, may be grouped as descriptive or inferential. Where as the former is the study of distributions of one variable, the latter is concerned with the process of generalization from the sample to the target population.

Analysis may serve several purposes:

It summarizes large mass of data into understandable and meaningful form (peculiar to descriptive statistics).

It makes exact descriptions possible. For example, when we say the educational level of people in Gamo Gofa Zone is high, the description is not specific; but when statistical measures like the percentage of literate, the description becomes exact.

It facilitates identification of the causal factors underlying complex phenomena, such as multivariate analysis.

It helps the drawing of reliable inferences from observational data – forecast based on past and present observations (peculiar to inferential statistics).

It also helps making estimations or generalizations from the results of sample surveys (inferential statistics).

It is useful for assessing the significance of specific sample results under assumed population conditions – *hypothesis testing*.

Activity 5.2

1 What purposes does analysis of data serve?

2 Why is a researcher usually interested in inferential statistical analysis in the case of sample survey?

Dear learners, you may follow different approaches to make your data analysis either manually or using computer softwares.[11] The first aspect is *descriptive analysis* which involves construction of statistical distribution, and calculation of simple measures like averages, percentages and measures of dispersion (such as mean, mode and median) for describing the features of the research. It is common to begin the analysis by showing the distribution of respondents by each of the pertinent variables (like age, sex, income level, etc). Then, some average value that represents the distribution is computed (mean, mode, or median).

The second aspect is to compare two or more distributions/sub-groups within a distribution usually using measures of central tendency, dispersion, ratio, proportion and percentages. The last three methods not only simplify comparison but also permit the comparison of groups of unequal size than the first two.

[11] EXCEL, STATA, EVIEWS, SPSS, LOTUS, PCGIVE, MATLAB, and WINRATS are some of the softwares used in economic analysis.

The third aspect is to study the nature of relationships among variables using methods like coefficient of correlation, partial and multiple correlations and regressions. Besides, other parametric (*t*-test and *F*-test) and non-parametric test (like chi-squared test, ^{2}-test and sign test) may be used in data analysis.

Tests for statistical significance are used to estimate the probability that a relationship observed in the data occurred only by chance; the probability that the variables are really unrelated in the population. They can be used to filter out unpromising hypotheses; and are used because they constitute a common yardstick that can be understood by a great many people, and they communicate essential information about a research project that can be compared to the findings of other projects.

However, they do not assure that the research has been carefully designed and executed. In fact, tests for statistical significance may be misleading, because they are precise numbers. But they have no relationship to the practical significance of the findings of the research.

Finally, one must always use measures of association along with tests for statistical significance. The latter estimate the probability that the relationship exists; while the former estimate the strength (and sometimes the direction) of the relationship. Each has its use, and they are best when used together.

Check Your Progress

5.3 Why is data analysis the most skilled task of the research process?

5.4 Distinguish between descriptive and inferential analysis using a hypothetical example.

Lesson Summary

The data you just collected from primary (and secondary) sources is known as *raw data*

Editing helps to ensure that it is *clean* – examining whether it is free from incompleteness, inconsistencies, outliers, contradictions, errors, omissions, misclassifications, etc and to correct these when possible.

In field editing the enumerators/data collectors shall edit while in the field and make the necessary rectification.

Central editing is editing at the center of the research either by the researcher(s) or by a team of editors depending on the scale of the survey. In large-scale surveys, editing may be done in different rounds by different groups of editors.

Coding refers to the process of assigning numerals or other symbols to answers so that responses can be put into a limited number of categories or classes that may be appropriate to the research problem at hand. It is necessary for efficient analysis.

Classification is the process of arranging things (either actual or notational) in groups or classes according to their resemblance and affinities/closeness.

There are plenty of alternative data presentation techniques – table, pie chart, bar graph, line graph, pictogram, frequency polygon, ogive, etc. Data presentation conserves space and reduces the need to have explanatory and descriptive statements; facilitates comparison activities; expedite the summation of items and detection of errors and omissions; and provides bases for various statistical analyses.

Data analysis is the most skilled task of the research process and refers to the computation of certain measures along with searching for patterns of relationship that exist among data groups.

CHAPTER SEVEN

WRITING THE RESEARCH REPORT

Lesson One

Description of Lesson Two

This lesson is about report writing – the final step of the research process. Research output, unless communicated to users, is waste of scarce resources. This step makes the research process complete. In this objective, this lesson introduces you with the techniques of writing a technical report and its main elements. We also finally make remarks on oral presentation and ethical issues in writing.

The final and most important step in implementation of a research project is writing a report. This need interpreting the data presented using different devices and analyzed using various techniques. Interpretation refers to the task of drawing inferences from the collected facts after an analysis. The task has two major aspects:

i The effort to establish continuity in research through linking the results of a given study with others/literature; and
ii Establishment of some explanatory concepts

Often, techniques of interpretation involve the following steps:

You must give reasonable explanations of the relations which you have found and you must interpret the lines of relationship;

Make use of extraneous information if collected during the study;

Consult professionals having insight on the issue before embarking on actual interpretation;

Learners shall note, however, that even if the data are properly collected and analyzed, wrong interpretation would lead to inaccurate conclusions. Hence, I advise you to pay attention to:

Whether the data invariably satisfies you that it is appropriate, trustworthy & adequate for drawing inferences; it reflects good homogeneity; and proper analysis is done.

Staying cautious about errors that may possibly arise in the process of your interpretation – false generalization, wrong interpretation, etc

The intertwined relation interpretation has with analysis and hence simultaneously done.

The purpose of *report writing* is to communicate to interested persons *what* was done, *why* it was done, *how* it was done, and *what* the outcomes are. It is, thus, a narrative but authoritative document on the outcome of a research effort. It presents highly specific information for a clearly designated audience. Learners must note that reports are the product of slow, painstaking, accurate inductive work that may involve:

the logical analysis of the subject matter – primarily concerned with the development of a subject (of interest to the readers/audience);

preparation of working and final outline – these serve as a framework for long written works construction, aid logical organization of the report and reminds points to be stressed in the report;

preparation of rough draft – follows the logical analysis of the subject matter and the preparation of final outline;

rewriting and polishing – involves checking for weaknesses in terms of logical analysis, coherence, grammar, spelling, etc; and

writing the final draft and preparing final bibliography – the final draft shall take into account possible remedies for weaknesses identified in the last step and try writing an exhaustive list of references used in writing the report.

Before embarking on the actual writing of your report, learners are advised to plan on the process. This starts with identifying your audience & their information/knowledge level, determine the intended purpose of your report (it is for a degree, reference material, implementation by firm, etc) as it may determine its contents and form of presentation, the type & scope of the intended report, and the format of the report. [12] Then, you shall start drafting your report by way of developing the likely content of the report and work accordingly. The outline may start with dividing the intended report into chapters (*chapterisation*).

[12] The standard guideline of the Department of Economics of the Arba Minch University for your term paper and other papers is attached in the appendix.

Then, you need to have a draft outline of contents for each of the proposed chapters which includes topic and sentence outlines. If you have a good outline, then your report *writes itself*. Of course, an outline should not be considered as a rigid scheme to be strictly followed. It is a tentative guide and is likely to undergo changes while the writing process proceeds. The elements of your (term) paper shall be illustrated using a comprehensive *technical report*, which is usually followed by academic papers and is broadly divided into: [13]

i the preliminary/prefatory pages
ii the body
iii the end matter

i Preliminary/prefatory pages:

The prefatory pages consist of, inter alia, the title page, declaration, certificate of supervisor/advisor and examiners, preface/acknowledgement, table of contents, list of tables and figures, abstract, if any, etc.

Any report shall start with the **title page** which consists of the institutional affiliation (if any) of the writer(s), the title/topic clearly identifying the study and may contain a description of the study design, purpose (like *in partial fulfillment of the requirements for the course Econ-341or Bachelor of Arts Degree in Economics, etc)*, writer's name and who it is submitted to (your course instructor/advisor), and date & place.[14]

The inner part of the report shall start with your **acknowledgment** furnishing credit to those individuals and institutions who contributed positively for the research (may include course instructor, your coordinator, advisor, colleagues, etc).

Then, you shall put the **table of contents –** an outline of the contents of the report. It contains the list of items like chapter and their sub-titles with respective page numbers. It facilitates ready location of items in the report. If your report consists of many tables and/or figures, you may need to have an outline of these as **list of tables and figures** (or each separately) in order to make reference of such data at ease.

[13] Standard guide of technical reports (term papers, practicum reports, senior essays and theses) is attached as appendix A(the title page) and appendix B (all other parts).
[14] A model title page is attached in appendix A.

These shall be followed by a succinct statement of the findings of the project – **abstract**. It should be as brief as possible to help readers get a quick overview of the report. It contains a brief and precise statement of the purpose and scope of the study and the method used for the investigation, and a bare summary of the findings or the results of the investigation. It should be written direct, concise and condensed.

Check Your Progress

5.5 What are the main elements of the prefatory pages of a research report and their purposes?

5.6 What are the necessary details of a title page?

ii The Main Text/Body

This is the major part of the report covering the formulation of the problem studied, methodology, review of literature, findings & discussion and a summary of the findings & recommendation. In a comprehensive report, the body of the report may have several chapters which shall start with the first chapter introducing the study.

This chapter may start by offering **background** to the problem under study in order to place it into a larger context to enable the readers know its significance in a proper perspective. This section must try creating reader interest by providing the broad foundation of the problem under study and place the study within the larger context of the scholarly literature.

Then, the **problem statement** describes the context for the study and it also identifies the general analysis approach. A *problem* might be defined as the issue that exists in the literature, theory, or practice that leads to a need for the study (motivation). Sometimes, obscure and poorly formulated problems are masked in an extended discussion. In such cases, readers, reviewers and/or committee members will have difficulty in recognizing the problem. This will clarify the writer's mind as to the purpose and it will inform the reader directly and explicitly. Effective problem statements answer the question *why does this research need to be conducted*. If you are unable to answer this question clearly and succinctly, then the statement of the problem will come off as ambiguous and vague.

Following the statement of the problem should appear the **objectives** stating the aim of the study very briefly. It may be organized in the structure of one general and other few specific objectives

derived from the general objective. One may state the objectives together with the investigative questions relating to each of the objectives. Try to incorporate a sentence that begins with "The general objective of this study is . . ." and state the specific objectives in bullets.

The dimensions of the study in terms of geographical coverage, designation of population under study and level of generality shall be delimited in the statement of **scope of the study**.

It is very common to include a discussion of the design of the study in this chapter usually stated as **data source and methodology**. This part is devoted for the presentation of all aspects of methodology and its implementation. The details of these should be so meticulously stated as to fully satisfy the criterion of *replicability*.[15] Though the technical details may be given in the appendix, one is expected to discuss the overall typology of research used, and the data collection methods. In addition, the data source, sampling design and other aspects of the design may be presented in separate (or combined) sections. Among others, the writer has to try to answer the following questions in this regard:

What is the overall typology of the research?

What is the source of data – primary or secondary?

If primary, what is the target population?

Is the study a sample survey or census survey?

If sample survey, what is the sampling technique employed and sample size? What is the instrument of data collection?

What is the response rate and limitations?

What are the techniques of data analysis?

What hypotheses are formulated and test techniques used?, etc

It is not uncommon to see (student) reports including not only **definition of terms and concepts** that help readers have a common understanding of technical terms used in the report but also **organization of the report**.

Most (student) researches go to **literature review** following the introduction which critically reviews the contemporary theory and wisdom on empirical findings related to the topic of study. One shall try to review not only what (economic) theory says about the issue but also critically

[15] See unit one lesson 7 on *criteria of good research*.

examine what other researchers did, how, and what they found with the aim of building theoretical and empirical framework for the study. This makes the work of analysis more interesting especially if the writer compares and contrasts against other findings. This assessment of literature is a way of uncovering the gap in knowledge the research is trying to fill. It provides the background and context for the research problem. It should establish the need for the research and indicate that the writer is knowledgeable about the area.

The core of the report where we see the efforts of the researcher is given in the **analysis and discussion** part of the report. This is basically the presentation of data, its analysis and interpretation. It may be divided into several parts based on the research objectives. The results should be reported as accurately and completely as possible, showing how they bear on the research questions/hypotheses. And, each section should be given an appropriate *heading*/sub-title. The discussion should be clear enough for the readers to be able to grasp the findings either by reading the text of looking at the data presentation device (table or chart).

The final chapter of the body of your report may include a **summary** of findings, **conclusion and recommendation** which is more extensive than the abstract given in the prefatory pages. It is a self-contained summary of the whole report – reproduction of the topical sentences of various findings and conclusion presented.

Summary is a brief presentation of the major findings; where as conclusion is a generalization of the findings to the population of interest. Recommendation, on the other hand, is a normative analysis of researcher as to what ought to be done about the problem based on the findings of the study. Make sure you neither over-generalize in your conclusion, nor be over-ambitious in your recommendations. The summary should be concise, the conclusion balance between specific and general, and the recommendation should not only be feasible but also attempt to show the roles of stake-holders in the suggested interventions. It is not uncommon to see reports combining these elements in one chapter.

iii The End Matter

The *end matter* includes the bibliography, index, and appendix (tools of data collection, extensive data, technical details, etc). Bibliography is the first item of the end matter presented at the end of the research report. It lists in alphabetical order all published and unpublished

references (including journals, magazines, books, pamphlets, newspapers, research reports, conference proceedings, etc) used by the writer in the report preparation. It gives a list of materials relating to the topic under study as a ready reference to the readers.

There may be several bibliographic entry formats, but the following points should be emphasized on the styles and formats in writing the bibliography:

All references noted in the text, including the sources for figures and tables must appear in the bibliographic list (note also that you should include only those references that are cited in the report).

Each item in the list should include all pertinent facts about the publication: author, title of the work in question and any journal or book in which it is included, place of publication, publishing company, date, page numbers, etc.

Do not number references

When there is more than one work by a singe author, place them in reverse chronological order i.e., put the most recent item first and work backwards without the need to repeat the name of the author.

Individual citation: list author's last name first (use initials for first names)

When there is more than one author of a single publication, the others are listed by their respective first names. However, when there are more than two writers, it is customary to write *et.al* (meaning, and others) following the name of the first (main) writer instead of writing the names of other writers (co-authors).

Alphabetize references according to the author's last names. Date of publication follows the names of the authors and is punctuated with a full stop.

Titles of articles are in double quotation marks with sentence style capitalization and punctuated with a full stop.

Books and journal titles are in italic's with volume and number.

Conference papers and presentations should be noted with title, sponsor, location and date. Publishing information is last with the publishing company's name.

A common way of writing bibliographic elements in Economics is the American Psychological Association (APA) style of writing. A brief summary of this bibliographic writing and the elements for different types of references is given next:[16]

For books:

- Name of the author, last name first (for Ethiopian names first name first)
- Date of publication
- Title of the book, *italics*
- Edition
- Place and publisher
- Number and volume

Periodicals (Magazines, Journal Articles and Newspapers):

- Name of the author
- Date of the issue/publication
- Title of the article
- Name of the periodical
- The volume or volume & number
- Pagination

In the body of the report, nevertheless, citations may be at the middle or end of a sentence or paragraph (author's name, year of publication; page numbers, if any (for example, Taddesse, 2009: pp 85).

With regard to index and other appended sections, the writer merely puts them next to the bibliography. If you make use of any appended section in your discussion in the body of the report, make sure you have made correct citation of the relevant element of the appended section.

Written reports should have a user-friendly layout, make a good use of headings and subheadings and must be organized around central points, has to leave technical details to the appendix. In writing, you are advised to use active verbs and short sentences, eliminate unnecessary phrases,

[16] Note also that there are various styles of writing bibliographic elements. If you are interested to read more refer to any of the references given in this module bibliographic section.

use spell check and grammar check, get proofread by an objective third party, double check all mathematical calculations and answers.

Check Your Progress

5.7 List the major and minor elements of the main text/body of a report and their roles.

5.8 What are the elements of the end matter and their roles in the report?

5.9 What are the bibliographic elements for books and periodicals in their order of appearance?

A final remark I would like to make is on **ethical issues in report writing**. Research requires many different kinds of writing: research proposals; progress reports; final reports; research instruments; instructions; informed consent forms; cover letters; etc. Students/researchers must conduct themselves ethically at all times.

Two major areas of concern are:

i Misconduct:

- making up data rather than gathering it;
- falsifying data that has been gathered, by altering it or selectively deleting it to obtain the desired results;
- plagiarizing--using someone else's ideas, words, or data without giving them credit;
- failure (intentional or unintentional) to properly cite used references;
- copy (in full or partial) of one's work and claiming as if it is your own work – plagiarism;
- failure to acknowledge the roles of other individuals, institutions and colleagues who took part either directly or indirectly in the research; etc

ii Handing Research Errors

- No research is error-free. The broad research community has adopted the practice of peer review to ensure that researchers practice ethical standards. Peer review means that both the methods and the results of one's research are reviewed by other researchers. In blind peer review, the reviewers do not know the identity of the author(s).
- Peer review is used extensively by agencies that fund research, to decide which proposals have scientific and technical merit. Peer review is used by Review Board to

decide if research is treating subjects ethically. And, peer review is used by all the major research journals, to decide which written reports merit publication.

- In peer review, there is full disclosure of research methodology and procedures. There should be clear and complete identification of the limitations of the study and their impact on the findings, conclusions, and recommendations.
- Copies of the research report should be made available to those who request them.
- Data should include instruments, code books, written instructions, field notes, informed consent forms, computer programs and files, printouts, diskettes, tapes, etc.

iii Follow carefully the content of the study program and try to clarify your own topic of the proposed project soon enough before you actually start to write it. The topic must reflect your interest in social research topic(s) and issues(s).

iv Follow carefully the lessons focused on empirical social research and its structure and think in advance of your own procedure in proposing your research project. There is no one absolute and definitive structure of research project but there are some very useful procedures and option of which you MUST learn and then you shall be able to select your own procedure.

v You shall be given the more precise instruction for writing a research project from the Department like the one attached in the appendix and your report should be organized as per the guideline.

Lesson Summary

The purpose of *report writing* is to communicate to interested persons *what* was done, *why* it was done, *how* it was done, and *what* the outcomes are. This helps one know/learn about the issue, use the information in the report for (optimal) decision making, identify a gap in knowledge for future research, etc.

The elements of a report that shall be followed by academic papers are broadly divided into:

i the preliminary/prefatory pages

ii the body

iii the end matter

The prefatory pages consist of, inter alia, the title page, declaration, certificate of supervisor/advisor and examiners, preface/acknowledgement, table of contents, list of tables and figures, abstract, if any, etc.

The *end matter* includes the bibliography, index, and appendix (tools of data collection, extensive data, technical details, etc).

Review Exercises

1. The row data on the variable **age** for a sample of 35 respondents (rural household heads) is distributed as given below:

18	Bekele Gutema	35	03	18	22	45
29	64	38	56	44	70	38
25	60	29	Gamo Gofa	48	44	52
43	260	56	60	49	44	43
27	49	51	58	47	46	No children

   a. Which of these are to be edited given the nature of the variable?
   b. How do you rectify some of these editorial problems?
2. Given a row data collected using questionnaire, how do you go about classifying the data?
3. A student researcher collected data using questionnaire and found out the source of financing college fee for week-end (non-regular) students in Soddo Station in different programs as given below:

Source of financing college fee	Own Source				Parents				Scholarship				others			
Department	A	E	G	C	A	E	G	C	A	E	G	C	A	E	G	C
Frequency	22	41	34	16	7	9	6	4	1	2	0	4	8	7	3	9

   a. How do you summarize the total number of students financing their college fee from different sources using a table?
   b. What is the other most convenient way of presenting this data with the objective of showing the relative share of different sources of financing college fee?
4. What is lacking in the above table for the researcher to include in the final research report? How do you suggest the student-researcher solves the problem?
5. What is the most read part of a research report?
6. List the most important elements the components of an academic research report.

REFERENCES

1. Ranjit Kumar. 2005. Research Methodology: A step-by-step guide for beginners. London. Sage Publications

2. C.R. Kothari. 2004. Research Methodology: Methods and Techniques. Second Revised Edition.

3. Guajarati, Damodar N. 2001. Basic Econometrics, : McGraw-Hill

4. Leedy, Paul. 1997. Practical Research: Planning and Design. Upper Saddle River: Prentice-Hall.

5. Mukherjee, C., H. White, and M. Wuyts. 1998. Econometrics and Data Analysis in Developing Countries. London: Routledge.

6. Newman, W. L. 1997. Social Science Research Methods: Qualitative and Quantitative Approaches. Boston: Ally and Bacon

7. Blaug, Mark. 1992. The Methodology of science: How Economists explain, second edition, 1992.

Answer Key for Review Exercises

Unit One

1 Research is a systematic and objective process of collecting, recording, organizing, and interpreting data for aid in decision-making.

2 Some of the criteria of research include:

Systematic and critical investigation of a

phenomenon Uses the scientific method

Objective and logical

Based on empirical evidence

Emphasis on the general than the specific

3 The scientific method refers to the ideas, rules, techniques and approaches characterized by the use of empirical evidence, logical reasoning and skepticism. It is widely used by the scientific community. Where as the arbitrary method seeks answers via imagination, belief, imagination, etc which all lack objectivity and suffer from subjectivity. People may gain knowledge via their opinion, belief, logical reasoning or observable experiences. The first two (and others that lack objectivity and use of empirical evidence) are arbitrary while the latter two are scientific.

4 Theories and facts are interlinked through the working of methods. Facts help modify, change or substitute theories and vice versa.

5 Generation of knowledge/information or filling the gap in knowledge.

6 The topic may be too broad, financial and time limitations, lack of expertise in appropriate research techniques, etc.

7 Where as applied research is based on real world problem(s), basic research aims at deriving general principles, theories and rules.

8 Based on the techniques employed, applied research may be classified in

to: controlled experiment

survey (case

study) observation

exploratory research

Unit Two

1 A1: Tourist inflow in Ethiopia is negatively affected by urbanization

A2: Tourist inflow in Ethiopia is not affected by urbanization

2 The problem statement

3 Survey of the literature

4 Whether objectives of a given research are good is measured by whether they are SMART

S=specific
M=measurable
A=achievable
R=relevant
T=time-bounded

5 Research aimed at identifying causal relations among variables of interest

6 When it reaches a point where the next research can hardly have a value added. When a topic is overdone, almost all aspects of the issue are already addressed by researches conducted so far with all available methods and hence additional research can only repeat them instead of contributing to knowledge.

7 Any research proposal shall consider, inter alia, the resource (material, finance, time, human – skilled and unskilled), the experience of the research team, the significance (cost compared to expected gain), the available tools, etc.

Unit Three

1 A research proposal is used for planning the research exercise in order to make it (more) effective and efficient. It is like a sailor's compass without which the exercise would be a mere wondering and wastes resources.

2 Some of the questions a research proposal is intended to answer include:

1 What is the proposed study about?

2 Why is the study being conducted? [very important]

3 Where will the study be carried out?

4 What is the objective of the proposed study?

5 What type of data is required to achieve these objectives and where & how is it collected? 6 What period of time will the study cover?

7 Is the study a sample or census survey? If sample, what will be the sampling design?

8 How will be the data analyzed?

9 In what style will the report be written?

10 Who are the potential beneficiaries of the proposed research?

11 What is the resource requirement of the research?

12 When and how the research output is to be disseminated?

3 Methodology generally addresses issues related to data collection, sample size and sampling techniques, data analysis, hypothesis testing and the likes. In this section of your proposal you are actually talking about *what and how* you are proposing to do your research. This determines the success of your proposed research.

4 Clearly defined scope of a proposed research avoids mere wandering and the subsequent wastage of scarce resources. If the scope of a proposed research is well-defined, it helps the researcher adequately address the research questions with lesser resources than otherwise.

5 A good research proposal must posses the important characteristics of *objectivity*, *reliability* (the degree to which a measure is consistent or dependable and the degree to which it would give you the same result over and over again, assuming the underlying phenomenon is not changing), *validity* (the best available approximation of the truth of a given proposition, inference, or conclusion) and *generalization*. A well conceived research design will pay rich dividends. In spite of all these features, a research design is *not a rigid plan* to be followed without change, but like a polestar which help the researcher to find out the right direction.

As research progresses, a change in original design may become imperative due to new facts, ideas and conditions. In addition to the regular characteristics of a good research design, the following are the additional features to support a good research design: *flexibility*, *appropriateness*, *statistical and economic efficiency minimize bias* and *maximize the reliability* of data, *smallest experimental error*, *effectiveness* (yields maximum information and considers many different aspects of the problem). A good research design is also organized, has a logical flow, concise (but also complete), uses a good grammar. To achieve all these, it's usually a good idea to have a colleague/friend read it before submitting it to the advisor/supervisor/funding agency. Often, little errors or small changes will be identified and addressed in this process of proof-reading.

6 The most common components of an academic research proposal include, among others, the title page, the introduction, the problem statement, the objectives (general and specific), the scope/delimitation, the methodology, and the time schedule.

Unit Four

1 Primary data is data collected for the first time for the research at hand either by the researcher of enumerators. Where as secondary data is one that is collected by others for their own purpose and used by a researcher again. Primary data is fresh and original, but secondary data is not.

2 Recall that this is cluster sampling that involves several stages

   a. The sampling element is the household in SNNP region
   b. The first stage sampling unit is the next administrative unit next to a region in the Ethiopian context; i.e. the *zone*.

3 This sample is unrepresentative as it treats only those who are already customers and came to use the services provided.

4 Generally, cluster sampling is known to reduce the cost of a research for studies that cover a vast geographic area.

5 Observation

6 This is double barreled question which might have two different answers. If one found it entertaining but not educative or the other way round, she can't answer the question in *yes* and *no*. The solution to such questions is to break it up so that it appears two different questions with the closed response form category for each and it appears:

   'Are the transmissions of ETV entertaining? Yes; No' 'Are the transmissions of ETV educative? Yes; No'

7 The data category is:

   a. Quantitative
   b. Qualitative
   c. Time-series

8 Data is used for aid in decision making

Unit Five

1 Given that the sampling unit is *rural household heads* and the variable of interest is age, one can see:

   a. the most obvious editorial problems include:

      i. Bekele Gutema [first row-second column] – there is no such age
      ii. 03 [first row-fourth column] – a household head cannot be this young
      iii. Gamo Gofa [third row-fourth column] – this is an administrative zone
      iv. 260 [fourth row-second column] – very large
      v. No children [last row-last column] – this cannot be age

   b. An editor may rectify some of these problems as

      i. Look in to additional information from the returned questionnaire
      ii. Instead of 03, it may be 30 if it is in line with other data given in that specific questionnaire
      iii. Look in to alternative ways to solve the problem
      iv. Instead of 260 it may be 26 or 20
      v. If the respondent has no children, an editor may opt for lower age groups

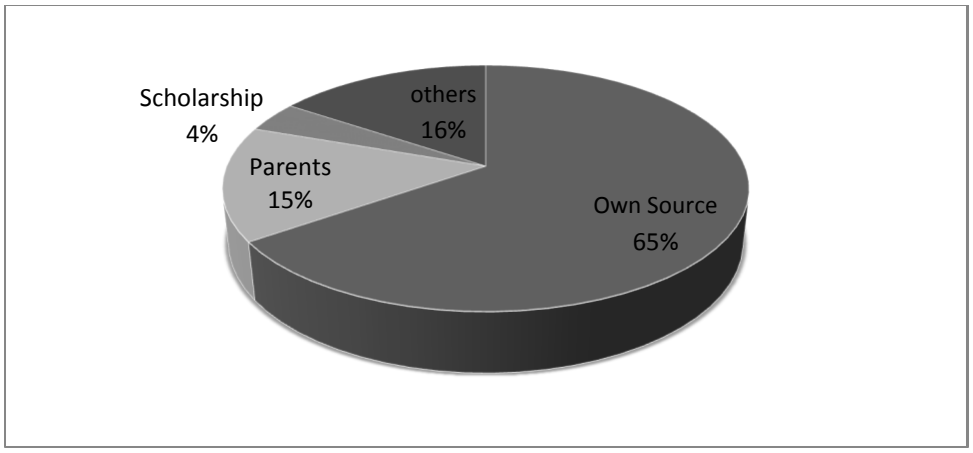
2 You need to have classification base and then a dummy table. Using tally (///) or other symbols you may count on the number of respondents in each category (frequency).

3 Pie chart

   a. The summary of students financing their college fee from different sources [this is the summation of the number of students financing their college fee from each source in each department]:

Source of financing college fee	Own Source	Parents	Scholarship	others
No of Students	113	26	7	27

   b. A pie chart (given next page)



4 The table lacks the table number, the title, key for abbreviations used, and source.

Table number shall be given in line with the chapter and section the table belongs to, like 3.1 or 4.2 and so on.

The title may be: Source of College Fee Financing: Soddo Week-end students of Arba Minch University, 2009

Key may be A=Department of Accounting and Finance; E=Department of Economics, G=Department of Geography; C=Department of Civil Engineering, etc

The source may be stated as own survey

5 The research topic/title

6 The major components include:

a. The prefatory pages
    i. Title page
    ii. The acknowledgement
    iii. Operational definition
    iv. Acronym
    v. The abstract
    vi. The table of contents
    vii. The list of figures and tables

b. The main text
    i. Introduction
        1. Background/overview
        2. Statement of the problem
        3. Objectives (general and specific)

4. Scope of the study
5. Methodology of the study
6. Limitation, if any
7. Significance, if any

ii. Literature survey (common but not a must for exploratory research)
    1. Empirical literature
    2. Theoretical literature

iii. Discussion and Analysis

iv. Summary/Conclusion and Recommendation

c. The end matter

i. Bibliography

ii. Appendix