Central University of Technology, Free State

Faculty of Management Sciences

Department of Business Support Studies

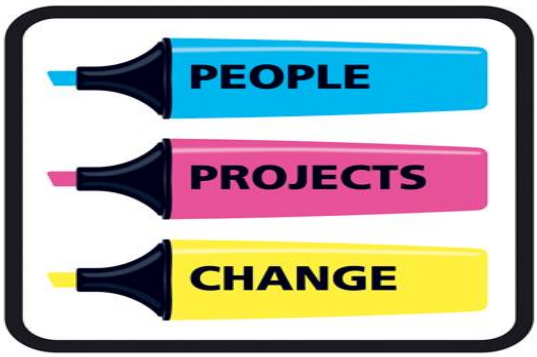
Project management

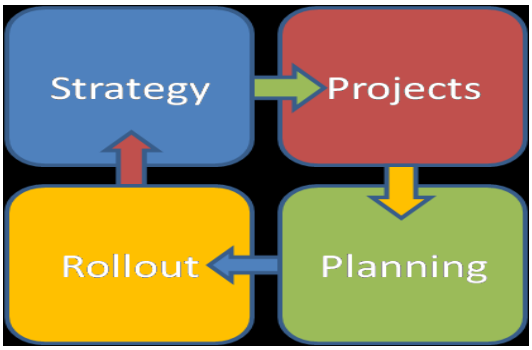
Business administration

Office management & technology

Learning Guide for 2016

Project Research IV: PKN41AB





PROJECT MANAGEMENT PROGRAM

BTECH: PROJECT MANAGEMENT

Programme Code	NQF level	Credits
BEBBTPJ	**7**	**12**

Compiled by:

Professor DY Dzansi (PhD)

E-mail: ddzansi@cut.ac.za

Tel: 051 507 3219

Room 001

Dr EK Agbobli

email: eagbobli@cut.ac.za

Tel: 0515073955

Room B106

Faculty of Management Sciences Building

****

WELCOME TO

PROJECT MANAGEMENT

@

CUT

NB: The material in this study guide is largely based on: (1) Leedy, PD & Ormorod, JE. 2010. *Practical Research, planning and design*. Pearson; and (2) Salkind, JN. 2009. *Exploring Research.* Pearson.

TABLE OF CONTENTS

COURSE DETAILS

COURSE NAME	**Project Research IV**
COURSE CODE	**PKN41AB**
NQF LEVEL	**7**
PROGRAMME	**B-Tech. Project Management**
CREDITS	**12**
NOTIONAL HOURS	**120**
CONTACT HOURS	**2 hours per week**

BRIEF DESCRIPTION OF CONTENT

Research is an indispensable tool in understanding and solving various forms of management problems. Thus, it is crucial that students enrolled in the BTech project management program understand the fundamentals of research in preparing for careers in project management. For this reasons, this course (Project Research IV) provides learners with introduction to **methods and techniques typically employed in social research**. Specifically, the purpose of this course is to equip learners with knowledge of basic research methods and strategies used in management/business research. It covers the scientific process including problem identification, problem statement, research question, and hypothesis formulation, measurement and instrument construction, data collection, research design and in particular, survey research methods, statistical analysis, qualitative research and ethical considerations. It is also intended to develop learners in exemplary academic writing skills and critical analysis. You will also be introduced to the philosophical issue of research paradigms.

NB: The material in this study guide is largely based on: (1) Leedy, PD & Ormorod, JE. 2010. *Practical Research, planning and design*. Pearson; and (2) Salkind, JN. 2009. *Exploring Research.* Pearson.

COURSE NORMS AND RULES

Every orderly human undertaking has ground rules. The following are some of the most important but not all of the ground rules for the course.

Late submission

Be warned that there is absolutely no room for late submission. As a rule, **late submission = no submission = ZERO (0).** Due dates are provided well in advance sometimes a year in advance so no excuse will be justifiable. Make sure you familiarise yourself with due dates.

STUDENTS TO NOTE CAREFULLY!

Special exam

Except for the mid-year (June) exam where a student may with the permission of the Assessment and Graduation Unit write sick (when you were really sick) or special exam (e.g. work commitment or some mishap), **there is no end of year special exam for PKN41AB.**

Language Medium

The official medium of communication at CUT is English and in compliance, the language medium for this course is English. Therefore, classes, tests, and exams will be conducted only in English.

Group work

Successful project management requires teamwork and in preparing you for that, a lot of your work will be group based. Every learner is expected to belong to a group. **Groups must strictly consist of five members.** In rare cases and only with the permission of your lecturer will any deviation from this norm be allowed. Each group member's participation must be indicated on the cover page of assignments as a percentage. Failure to do so gives the lecturer liberty to award marks as (s) he deems fit and **NO appeals will be tolerated.**

Class attendance and lectures

The BTech project management is a contact learning program. Therefore, class attendance is compulsory and there are penalties for non-attendance. All students must sign the attendance register.

Your lecturers will **not refer exclusively to textbook material since knowledge acquisition in research is not confined to textbooks**. Rather, lectures **WILL** cover similar concepts and approaches to research using examples from the broader research literature. Beyond merely attending class, students are expected to contribute to class discussions as a result of careful preparation of course materials. Overall participation will involve individual, as well as group-based participation**. It is assumed that students have read the assigned readings before class.** Some exam questions will be drawn from books that will not be discussed in lecture. Students are expected to attend each lecture and will be responsible for the material.

Other expectations from learners

The lecturer expects every learner to**:**

- **Search** for and get other sources that relate to the course in addition to what is provided. The facilitator will only provide guidelines and assistance in this regard.
- **Maintain** academic integrity (not to engage in any academic dishonesty). Some examples of academic dishonesty are: signing the attendance register on behalf of someone else; cheating in exams, assignments, and tests; plagiarism in any form; fabricating information or citations.
- **Accept responsibility** and take ownership of own learning.
- **Arrange** consultation with the lecturer through the secretary (0515073219 or 0515073955) as there cannot just be walk-in consultations due to time constraints on the part of lecturers.
- **Get exposed** to and become competent in the use of **Blackboard**. This requirement is very important because it is the main medium of communication between the lecturer and you, other than in-class communication.

LEARNING ASSUMED TO BE IN PLACE

(A) A formal qualification on NQF level 6 (old) or level 7 (new);

(B) Computer literacy - ability to: use Microsoft word, PowerPoint, and excel and search for information on the internet; and

(C) Basic numeric literacy - ability to: construct frequency tables and calculate mean, mode, and median.

LEARNING OUTCOMES FOR THE COURSE

Upon completion of this course, the candidate will in the context **of project management**, provide evidence that (s) he knows, understands, and can apply the principles of quantitative and qualitative research in an own or group research project; can compile a research proposal; and carry out a small scale empirical research.

CRITICAL OUTCOMES

After successfully undergoing this course, learners will:

1. Develop and can demonstrate scientific writing skills through writing assignments, tests, exams and research report;
2. Develop and be able to apply critical thinking skills through individual and group work;
3. Develop understanding of and will be able to examine basic concepts in project management through reading and class discussion;
4. Develop research skills and be able to examine and apply scientific and research concepts and methods related to project management in class, writing assignments, and conducting empirical research on a project management topic;
5. Develop collaborative skills and will be able to collaborate in group interactions through class and group projects;

6. Develop communication skills and will be able to communicate understanding of concepts through in-class presentation, assignments, tests, exams, and research report.

ASSESSMENT CRITERIA

Candidates should display knowledge and understanding of the principles of social research and apply these correctly in the compilation of a **research proposal**.

ASSESSMENT METHODS

1. Course mark counts 40% of final mark

Formal group assignments and/or tests addressing the critical outcome areas identified above.

- Test 1 – multiple choice (individual) = **50% of course mark = 20% of final mark**
- Test 2 – multiple choice and/or long questions (individual) = **50% of course mark** = **20% of final mark.**

Note Carefully

You must obtain a course mark of at least 40% in order to qualify to write the semester exam

2. Exam mark (60% of final mark)

(a) A formal 3 hour examination that counts 60% of the final mark addressing the critical outcome areas identified above.

3. Final mark = course mark (40%) + exam mark (60) = 100%

IMPORTANT NOTICE

The BTech project management like all fourth year courses is pitched at honours degree level and all students are expected to produce work that reflects academic maturity. Therefore, the following should be carefully noted:

1. Appropriate language, logical academic arguments, insight, and appropriate technical outlay of work all count towards the mark you get.
2. It is the student's responsibility to make assignments, tests and exams readable. Hand written assignments will not be permitted.
3. **Group work means just that.** A group must consist of **five (5) members**. No deviations will be accepted.
4. Distance, work commitment or any other flimsy reason is unacceptable.
5. **Individual work means just that.**
6. The BTech. Project management programme at the CUT is **a part-time contact** program and **not distance learning (or correspondence)** program. Attendance is therefore compulsory. There is a clear institutional regulation regarding absence from class, exams, and tests. No deviations from established regulations will be granted.
7. **Late submission means no submission = ZERO (0). You will be provided with assignment and test schedules. You must make your employers adhere to them and not the University adhering to your work schedule!**
8. Main evaluations will have up to but not more than 40% multiple choice questions.

PLEASE NOTE WELL!!!!!

- **Under no circumstance will late submission of any part of the research (proposal, literature review, and report) be permitted!**
- **Late submission = no submission = 0 (zero)**

STATEMENT ON ACADEMIC INTEGRITY

Academic dishonesty is a serious offence at CUT. The following examples are not exhaustive but describe CUT's policies for what constitutes academic dishonesty and penalties thereof (**It is your responsibility to know what constitutes a violation of academic integrity).** Academic dishonesty includes, but is not limited to: signing an attendance sheet using somebody else's name, cheating on exams, plagiarizing, handing in papers that were downloaded from the web, fabricating information or citations, facilitating acts of academic dishonesty by others, etc. The usual punishment for academic dishonesty ranges from deduction of marks to total expulsion from CUT.

PRESCRIBED TEXT BOOKS

Ranjit Kumar. Research Methodology – a step-by-step for beginners. 3rd Edition. Sage.

ADDITIONAL READING

1. Practical Research – Planning and Design 10th Edition: Leady & Ormond – Pearson
2. Exploring research 7th Edition: Salkind – Pearson
3. Research Methods in Business studies: Ghauri and Gronhaug – Prentice Hall

IMPORTANT DATES AND ACTIVITIES

Tests and assignment submission dates will be provided by respective lecturers as there are different student groups

COURSE OUTLINE

UNIT 1: RESEARCH FUNDAMENTALS **WEEKS 1 & 2**	
Objectives (expected outcome): To understand-	**Assessment criteria: The ability to-**
The nature of research	Define research;
Research paradigms or philosophies	Explain and differentiate the key paradigms in research;
The characteristics of good research	State the most important characteristics of a good research;
The research process	Describe the basic steps in the research process; Describe the activities carried out at each stage of the research process;
The key terminologies used in research	Define and explain key terms such as concepts, definitions, theories, methods, etc.
Common ethical dilemmas in social research	Describe typical dilemmas social researchers encounter; Explain steps that can be taken to deal with ethical dilemmas in social research; Conduct research in an ethically acceptable manner;
Study: Kumar Chapter 2; Ghauri Chapters 2,3; Leedy Chapters 1,2,5; Salkind Chapters 1, 2, 3 and any other similar source	

UNIT 2: FOCUSING THE RESEARCH WEEK 3	
Objectives (expected outcome): To understand-	**Assessment criteria: The ability to-**
1. Problem formulation	Identify and describe research problem; Divide a research problem into sub-problems; State: hypotheses, research questions and objectives;
2. Delimitation of research	State what will not be done in the research;
3. Operational definitions	State operational definition of key terms used in the research;
4. Assumptions	State underlying assumptions in a research;
5. Ethics in	List, explain, discuss, and address common ethical dilemmas in social research.
Study: Kumar Chapter 4; Leedy Chapters 3 & 5; Salkind Chapters 2, 3a and any other similar source	

UNIT 3: LITERATURE REVIEW WEEK 4	
Objectives (expected outcome): To understand-	**Assessment criteria: The ability to-**
1. The importance of literature review	Discuss the purpose of literature reviews; Explain the difference between primary and secondary data;
2. Sources and strategies for conducting thorough and efficient literature review	State the typical sources of credible secondary data; Conduct literature review.
Study: Kumar Chapter 3; Leedy Chapter 4 and any other similar source	

UNIT 4: KEY ISSUES RELATED TO QUANTITATIVE DATA GATHERING WEEKS 5 & 6	
Objectives (expected outcome): To understand-	**Assessment criteria: The ability to-**
1. General criteria for a research project	Describe the features common to all research
2. The nature and role of data in research	Differentiate between primary and secondary data
3. Sampling, sampling error, and generalizability	Define generalizability Define and differentiate between a sample and a population Determine when sampling is more preferable to census Describe various sampling techniques and when to choose a particular one Explain why good sampling is important Explain what sampling error is Identify sources of sampling error Explain how to reduce sampling error Explain how to estimate (calculate) minimum (required) sample size
4. Measurement	State and explain the different levels of measurement
5. Reliability and validity	Define validity and reliability Describe the different types of reliability and validity Explain the role of measuring instrument in reliability and validity Explain how reliability and validity can be increased and measured
Study: Kumar Chapters 12, 11, 10, Salkind Ch4; 5. Leedy chapter 5 and any other similar source	

UNIT 5: CLASSIFYING RESEARCH - *QUALITATIVE RESEARCH* WEEK 7	
Objectives (expected outcome): To understand-	**Assessment criteria: The ability to-**
Qualitative Research	Define and explain the term *research design* Identify, describe, and distinguish between various qualitative research designs Know when to choose the qualitative approach Describe the various methods for collecting data in qualitative research
Study: Kumar Chapter 7, 8; Leedy Chapter 7 Salkind Chapter 10 and any other similar source	

UNIT 6: CLASSIFYING RESEARCH - *QUANTITATIVE RESEARCH* WEEKS 8 & 9	
Objectives (expected outcome): To understand-	**Assessment criteria: The ability to-**
1. Quantitative Research	Describe the various quantitative research designs (excluding experimental designs) at least as listed in ***Leedy .*** ***NB. Although experimental research will be briefly discussed in class knowledge of it is not required for assessment purpose. The student will however do well to mater it for own future use.***
2. How to conduct interviews in quantitative research	Explain the use of face-to-face and telephone interviews in quantitative research
3. How to construct questionnaire for data collection in quantitative research	Describe various types of questionnaire/scale formats used to gather quantitative data Construct and administer questionnaire
Study: Kumar Chapter 7, 8; 9, 10; Leedy Chapter 7; Salkind Chapter 4, 5, 7 and any other similar source	

UNIT 7: STATISTICAL TECHNIQUES FOR DATA ANALYSIS WEEKS 10 & 11	
Objectives (expected outcome): To understand-	Assessment criteria: The ability to-
The use of statistical techniques in analysing quantitative data	Explain what is meant by "inferential statistics" & "descriptive statistics" Classify/identify a statistic as either descriptive or inferential Explain the use of "inferential statistics" in: Estimating population parameters Hypotheses testing Classify/identify a statistic as a measure of central tendency Identify/classify a statistic as a measure of variability Explain the concept of correlation and correlation coefficient
Study: Kumar Chapter 16; Leedy chapter 11 & Salkind chapter 7, 8, 9 and any other similar source	

DETAILED SYLLABUS

UNIT 1: RESEARCH FUNDAMENTALS

WEEKS 1 & 2

1.1 DEFINING RESEARCH

Research can be defined as "a systematic process of collecting, analysing, and interpreting data to increase understanding of a phenomenon about which we are interested". It is a process through which new knowledge is discovered.

1.2 RESEARCH PHILOSOPHIES/PARADIGMS

A scientific (research) paradigm, in the most basic sense is a framework containing all of the commonly accepted views about what direction a research endeavour should take and how it should be performed. You can read more on this subject at: http://www.experiment-resources.com/what-is-a-paradigm.html#ixzz1hqniEsGd

Thomas Kuhn is known to have popularised the term 'paradigm', which he described as essentially a collection of beliefs shared by scientists (a set of agreements about how problems are to be understood). According to Kuhn, paradigms are essential to research because as he saw it, "no natural history can be interpreted in the absence of at least some theoretical and methodological belief that permits selection, evaluation, and criticism." Therefore, paradigms guide the research efforts of scientific communities, and it is this criterion that most clearly identifies a field as a science.

There are two main research paradigms namely positivism and anti-positivism (or interpretivism or naturalistic inquiry).

Positivism

The positivist paradigm of exploring social reality is based on the idea that *observation* and *reason* are the means of understanding social phenomena. According to this philosophy, **true**

knowledge should be based on *experience of senses* which can only be obtained by *observation* and *experiment*. Positivists adopt this so called "scientific method" as a means of knowledge generation. Positivist paradigm is based on certain principles and assumptions of science.

According to Conen et al. (2000), the assumptions that underpin positivism are: *determinism*, *empiricism*, *parsimony*, and *generality*.

- ***'Determinism'*** means that events are caused by other circumstances; and hence, understanding such casual links are necessary for prediction and control.
- ***'Empiricism'*** means collection of verifiable empirical evidences in support of theories or hypotheses.
- ***'Parsimony'*** refers to the explanation of the phenomena in the most economical way possible.
- ***'Generality'*** is the process of generalizing the observation of the particular phenomenon to the world at large.

With these assumptions, the ultimate goal of any research endeavour should be to integrate and systematize findings into a meaningful pattern or theory bearing in mind that these are only tentative and not the ultimate truth. Theory is subject to revision or modification as new evidence emerges. The guiding principle of the positivist paradigm is the ***belief in objective reality*** meaning:

- Knowledge is only gained from sense data that can be directly experienced and verified between independent observers.
- Phenomena are subject to natural laws that humans discover in a logical manner through empirical testing, using inductive and deductive hypotheses derived from a body of scientific theory.
- Reliance of methods on quantitative measures, with relationships among variables commonly shown by mathematical means.
- Stable external reality that should be studied in a law-like manner.

- The nature of the relationship between the researcher and what is researched is objective and done in a detached manner.
- The methodology usually followed is experimental, quantitative, and hypothesis testing.

Positivists therefore aim at uncovering general laws of relationships and/or causality that apply to all people and at all times. The positivist researcher attempts to develop and test theories and models, for example, entrepreneurship can be explained and predicted. **Positivistic paradigm** thus systematizes the knowledge generation process with the help of **quantification,** which is essentially to enhance precision in the description of parameters and the identification of relationships among them.

Although positivistic paradigm continued to influence research for a long time, it was criticized due to its lack of regard for the subjective states of individuals. A main criticism is that it regards human behaviour as passive, controlled and determined by external environment. Hence human beings are dehumanized without their intention, individualism and freedom taken into account in viewing and interpreting social reality. According to the critics of this paradigm, objectivity needs to be replaced by subjectivity in the process of scientific inquiry. This gave rise to Interpretivism or anti-positivism or naturalistic inquiry.

Interpretivism/Anti-positivism

Anti-positivism emphasizes that social reality must be viewed and interpreted by the individual according to the ideological positions that person possesses. Therefore, knowledge is personally experienced rather than acquired from or imposed from outside. The belief is that reality is multi-faceted and complex and any single phenomenon has multiple interpretations. They emphasize that the verification of a phenomenon is adopted when the level of understanding of a phenomenon is such that the concern is to probe into the various unexplored dimensions of a phenomenon rather than establishing specific relationship among the components, as it happens in the case of positivism.

Anti-positivism is marked by three schools of thought in the social science research. These are *phenomenology*, *ethnomethodology* and *symbolic interactionism*. All the three schools of thought emphasize human interaction with phenomena in their daily lives, and suggest qualitative rather than quantitative approach to social inquiry.

'Phenomenology' is a theoretical view point which believes that individual behaviour is determined by the experience gained out of one's direct interaction with the phenomena. It rules out any kind of objective external reality. During interaction with various phenomena, human beings interpret and attach meanings to different actions and thereby construct new experiences. Therefore, the researcher has to develop empathic understanding: (1) to know the process of interpretation by individuals so that (s)he can reproduce in her/his mind feelings, motives and thoughts that are behind the action of others.

'Ethnomethodology', deals with the world of everyday life. According to enthomethodologists, theoretical concerns centre on the process by which common sense reality is constructed in everyday face-to-face interaction. This approach studies the process by which people invoke certain 'take-for-granted' rules about behaviour which they interpret in an interactive situation and make it meaningful. They are mainly interested in the interpretation people use to make sense of social settings.

'Symbolic interactionism'. The peculiarity of this approach is that human beings *interpret and define* each other's actions instead of merely *reacting* to each other's actions. Human interaction in the social world is mediated by the use of symbols like language, which help human beings to give meaning to objects. Symbolic interactionists, claim that by only concentrating attention on individuals' capacity to create symbolically meaningful objects in the world, human interaction and resulting patterns of social organizations can be understood. As a result, not only do human beings change themselves through interaction, but also bring in change in societies.

The two paradigms presented above are concerned with two concepts of social reality. While positivism stands for *objectivity, measurability, predictability, controllability and constructs*

NB: The material in this study guide is largely based on: (1) Leedy, PD & Ormorod, JE. 2010. *Practical Research, planning and design*. Pearson; and (2) Salkind, JN. 2009. *Exploring Research.* Pearson.

laws and rules of human behaviour, interpretivism essentially emphasizes *understanding* and *interpretation* of phenomena and *making meaning* out of this process.

Alongside the presence of these two major paradigms, a third paradigm of research called *Critical Theory* emerged during the post-sixties. Thereafter, a plethora of paradigms emerged but we shall not get into any more discussion on these.

Research paradigms and choice of research methods

Each of the paradigms discussed above implies associated research methods which can be used in carrying out scientific investigation.

Positivism which emphasizes objectivist approach to studying social phenomena gives importance to research methods focusing on quantitative analysis, surveys, experiments and the like.

Anti-positivism which stresses on subjectivist approach to studying social phenomena focuses on qualitative analysis, e.g. personal interviews, participant observations, account of individuals, personal constructs etc.

The question arises: how does a researcher select a research paradigm and corresponding methodology and methods? Answering the following questions can help you:

1. What is the nature or essence of the social phenomena being investigated?
2. Is social phenomenon objective in nature or created by the human mind?
3. What are the bases of knowledge corresponding to the social reality, and how knowledge can be acquired and disseminated?
4. What is the relationship of an individual with her environment? Is the researcher conditioned by the environment or is the environment created by the researcher?

1.3 CHARACTERISTICS OF A GOOD RESEARCH

High quality research:

i. Is able to be replicated
ii. Is able to be generalized
iii. Is based on logical rationale and theory
iv. Originates with a question or problem
v. Generates new questions
vi. Is an incremental process
vii. Is undertaken to benefit society
viii. Requires clear articulation of a goal
ix. Requires a specific plan for proceeding
x. Usually divides the principal problem into more manageable sub problems
xi. Is guided by the specific research problem, question, or hypothesis
xii. Accepts certain critical assumptions
xiii. Requires the collection and interpretation of data
xiv. Is, by its nature, cyclical or helical

YOU MUST BE ABLE TO ELABORATE ON EACH OF THE ABOVE. THAT IS, YOU MUST BE ABLE TO EXPLAIN EACH OF THE ABOVE

1.4 BASIC STEPS IN THE RESEARCH PROCESS

The following can be regarded as some of the key steps in the research process.

1 **A problem arises** In the mind of the researcher, a problem arises that has no known solution.	5 **Data generation** The researcher collects data that potentially relate to the problem.
2 **Putting it into words** The researcher converts the question to a clearly stated research problem.	6 **Arranging the data** The researcher arranges the data into a logical structure
3 **Guessing the solution** The researcher poses a temporary hypothesis or series of hypotheses or alternatively research question(s).	7 **Data analysis** The researcher analyses and interprets the data to determine their meaning
4 **The search for a solution** The researcher searches the literature for ideas that shed light on the problem and for strategies that may help to address the problem	8 **Drawing conclusion** The data either seemingly resolves the research problem/question(s) or do not. That is, they either support the hypothesis/hypotheses or do not

Source: Leedy & Ormond, 2010 with adaptations

RESEARCH METHODOLOGY – A TEACHING/ LEARNING FRAMEWORK

<table>
<tr><td colspan="9">Research Philosophy</td></tr>
<tr><td colspan="6">Philosophical issue</td><td colspan="3">Research paradigm</td></tr>
<tr><td colspan="6">Ontology - Nature of knowledge</td><td colspan="2">Objectivism</td><td>Subjectivism</td></tr>
<tr><td colspan="6">Epistemology - How to acquire “true” or “authentic” knowledge</td><td colspan="2">Positivism</td><td>Interpretivism</td></tr>
<tr><td colspan="6">Approach – Broad classification of the research – mainly concerned with data type</td><td colspan="2">Quantitative</td><td>Qualitative</td></tr>
<tr><td colspan="9">The research process</td></tr>
<tr><td>Phase</td><td>Phase I</td><td colspan="4">Phase II</td><td colspan="3">Phase III</td></tr>
<tr><td rowspan="3">Tasks</td><td>Deciding</td><td colspan="4">Planning</td><td colspan="3">Undertaking/executing</td></tr>
<tr><td>WHAT</td><td colspan="4">HOW</td><td colspan="3">COLLECTING EVIDENCE</td></tr>
<tr><td>Research questions to answer</td><td colspan="4">To gather evidence to answer research questions</td><td colspan="3">The required information</td></tr>
<tr><td>Steps</td><td>Step 1</td><td>Step 2</td><td>Step 3</td><td>Step 4</td><td>Step 5</td><td>Step 6</td><td>Step 7</td><td>Step 8</td></tr>
<tr><td></td><td>Formulate research problem</td><td>Identify research design</td><td>Data collection instrument</td><td>Sampling</td><td>Write proposal</td><td>Collect data</td><td>Process data</td><td>Write report</td></tr>
<tr><td rowspan="5">Required theoretical knowledge</td><td>Steps in formulating research problem</td><td>Design functions</td><td>Methods & tools of data collection</td><td rowspan="5">Sampling theory and design</td><td rowspan="5">Essential parts of proposal</td><td rowspan="5">Methods, tools of data collection</td><td rowspan="5">Methods of data analysis</td><td rowspan="5">Principles of scientific writing</td></tr>
<tr><td>Variables</td><td rowspan="4">Design types</td><td rowspan="2">Validity</td></tr>
<tr><td>Questions</td></tr>
<tr><td>Hypotheses</td><td>Reliability</td></tr>
<tr><td>Objectives</td><td></td></tr>
<tr><td>Required activity</td><td>Literature review</td><td>Literature review</td><td>Literature review;</td><td>Literature review</td><td>Literature review</td><td>Test instrument</td><td>Coding</td><td>Literature review</td></tr>
<tr><td></td><td></td><td></td><td>Test instrument</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td></td><td>Coding</td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Refer to Kumar pages 18-27

1.5 SOME KEY CONCEPTS IN RESEARCH

Theory

- A theory is the explanation of the relationship, between elements in adequate detail to explain an event or phenomenon, enabling the making of predictions arising from the theory (Cooper & Schindler, 2006:47). Theories tend to be complex, abstract, and involve multiple variables.
- An organized body of concepts and principles intended to explain a particular phenomenon
- Tentative explanations that new data either support or do not support

Research methodology

- The general approach the researcher takes in carrying out the research project.
- The techniques one uses to collect and analyse data; may be specific to a particular academic discipline.

Hypothesis

- A logical supposition, a reasonable guess, an educated conjecture;
- Provides a tentative/proposed explanation for a phenomenon/problem under investigation;
- An assumption about a population parameter which may or may not be true;
- Restates the general problem in a form that is precise enough to allow testing;
- The best way to determine whether a statistical hypothesis is true would be to examine the entire population. Since that is often impractical, researchers typically examine a random sample from the population. If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected. There are two types of hypothesis namely **alternative hypothesis** and **null hypothesis.** They are the two rival hypotheses which are compared by a **statistical hypothesis test**.

Research/alternate hypothesis

The **alternative hypothesis** (or **maintained hypothesis** or **research hypothesis**) usually denoted by H_1 or H_a asserts:

- A particular relationship between two or more measured phenomena;
- A difference between groups (A statement of inequality) and it can take many forms, e.g. 'not the same' or 'greater than' or 'less than'.

  H_a: $\mu_1 \neq \mu_2$

  - H_a: Research hypothesis
  - μ_1: Theoretical average of say population 1
  - μ_2: Theoretical average of say population 2
- That a potential treatment has an effect.
- To compare against null hypothesis

Examples of alternative hypothesis:

- Female employees are more in favour of working flexible hours than males.
- People who exceed the speed limit when driving are significantly more likely to cause motor accidents than those who do not exceed the speed limit.
- Learners who have attended a study skills workshop are more likely to perform better than learners who did not attend the workshop.

Null hypothesis

Typically denoted by H_0 proposes a general or default position, such as: that

- There is no relationship between two or more measured phenomena;
- That a potential treatment has no effect; or
- That sample observations result purely from chance.
- Statement of equality (there are no differences between groups)

  H_o: $\mu_1 = \mu_2$

  - H_o: Null hypothesis
  - μ_1: Theoretical average of say population 1
  - μ_2: Theoretical average of say population 2

Examples of null hypotheses:

- Female employees are not more in favour of flexible working hours than male employees.
- There will be no difference between learner performance of learners who attend a study skills workshop and those learners who did not attend.
- There will be no difference between the performances of employees that engage in good health behaviour than those that do not.

Differences between null and research hypotheses

■ Null ❑ Equality between variables ❑ Refers to population ❑ Directly tested ❑ Stated using Greek symbols (μ_0) ❑ Explicit	■ Research/alternate ❑ Inequality between variables ❑ Refers to sample ❑ indirectly tested ❑ Stated using Greek symbols (μ_a) ❑ Implicit

A simple example:

Suppose we wanted to determine whether a coin was fair and balanced. A null hypothesis might be that half the flips would result in Heads and half, in Tails. The alternative hypothesis might be that the number of Heads and Tails would be very different. Symbolically, these hypotheses would be expressed as

H_0: P = 0.5

H_a: P ≠ 0.5

Suppose we flipped the coin 50 times, resulting in 40 Heads and 10 Tails. Given this result, we would be inclined to reject the null hypothesis. We would conclude, based on the evidence, that the coin was probably not fair and balanced. Source: (http://stattrek.com/)

Can one accept the null hypothesis?

Some researchers say that a hypothesis test can have one of two outcomes: you accept the null hypothesis or you reject the null hypothesis. Many statisticians, however, take issue with the notion of "accepting the null hypothesis." Instead, they say: you reject the null hypothesis or you fail to reject the null hypothesis.

Why the distinction between "acceptance" and "failure to reject?" Acceptance implies that the null hypothesis is true. Failure to reject implies that the data are not sufficiently persuasive for us to prefer the alternative hypothesis over the null hypothesis.

Source: (http://stattrek.com)

What makes a good hypothesis? A good hypothesis is:

- Stated in declarative form
- Posits a relationship between variables
- Reflects theory or literature
- Brief and to the point
- Testable

Samples and populations

- **SAMPLE** is a representative portion of a **POPULATION.**
- **POPULATION** is the entire group of interest.
- Results from a good **SAMPLE** should generalize to the **POPULATION** under study.

Logical reasoning

Good research is based on sound reasoning. Sound reasoning is finding the correct **premises**, testing the connections between facts and assumptions, and making claims based on adequate **evidence**. The **evidence**, called **premises**, must support the **claim**, known as

the **conclusion**, e.g.: *All professors are well read; Mpho is a professor; Therefore, Mpho is well read.* The first two are the premises and the third line is the conclusion. There are two types of reasoning: **deductive** reasoning and **inductive** reasoning.

- ***Deductive reasoning:***

✓ Begins with one or more premises—statements or assumptions that the researcher initially takes to be true; valuable for generating research hypotheses and testing theories.

✓ Deductive reasoning is a form of reasoning in which the conclusion must necessarily follow from the reasons given. A deduction is valid if it is impossible for the conclusion to be false if the premises are true. The premises provide convincing evidence for the conclusion (Cooper & Schindler, 2006:32-33).

For example: if a = b and b = c then a = c.

- ***Inductive Reasoning:***

Begins with an observation of a specific event to draw conclusions about entire classes of objects or events (i.e., observe a sample and then draw conclusions about the population from which the sample has been taken).

Variables

Variables are a class of outcomes that can take on more than one value that is, variable refers to anything in the universe that has the capacity to vary in attributes or characteristics. It symbolises the property being studied or an event, act, characteristic, trait, or attribute that can be measured and to which we assign categorical values. For purposes of data entry and analysis we assign a numerical value to a variable based on the variable's properties.

Dependent variable (DV)

The **dependent variable** (DV) is predicted, measured, or otherwise monitored and is expected to be affected by manipulation of an **independent variable** (IV). It is the outcome of the changes in the independent variable hence it is also called the **outcome variable**. The

dependent variable (usually symbolised as Y) is the 'effect' (or outcome) in which the researcher is interested while the independent variable (usually symbolised as X) is the **presumed 'cause'**. Changes in the independent variable lead to changes in the dependent variable.

Independent variable (IV)

- Treatments or conditions under control of the researcher;
- The independent variable (IV) is the variable that is manipulated by the researcher;
- The IV is the factor that affects the dependent variable;
- In experimental research the IV is manipulated to demonstrate a causal influence on another variable so that predictions can be made as to the nature of the effect.

Examples of DV and IV:

- Jogging makes you feel better.

Independent variable = jogging or not jogging

Dependent variable = feeling better or not feeling better

- Fire causes smoke in nature

Fire	*Smoke*
Independent variable	Dependent variable
X	*Y*

To be able to claim that changing the value or level of one variable (IV) produces predictable changes (causal effect) in another, certain conditions must be met:

- The IV (cause) must always occur before the DV(effect)
- The two events must co-vary (change in value together)
- There can be no other explanations for the observed effect of X on Y (X = IV and Y = DV)

Non-experimental research does not show causal relations between variables but only the degree to which a variable changes in accordance with changes in another variable. The

independent (predictor) variables can predict an influence but cannot claim that the one causes the other, e.g.: *Longevity and higher income have a strong positive relationship.* Certainly, higher income does not "cause" people to live longer; it is what money buys namely –medical services, good food, good living conditions.

Other important variables (the third variable or confounding variable)

- **Extraneous Variable:** Has an unpredictable impact on the dependent variable.

Extraneous variables (EV) are variables that are not a part of the research, yet may have an influence on the variables included for research. These variables may affect the changes attributed to the independent variables. The extraneous variables are usually not measured by the study but may decrease or increase the magnitude or strength of the relationship between the independent and dependent variables.

- **Moderator Variable:** Variables related to independent or dependent variables, and hiding the true relationship between independent and dependent variables. A moderating variable (MV) is a second independent variable that is included because it is believed to have a significant contributory or contingent effect on the originally stated IV-DV relationship. The MV interacts in such a way that the effect on the response by one variable depends on the level or value of another IV. Two independent variables have not only a separate effect but a combined effect. Moderating variables are factors that alter the magnitude of a given relationship. The one independent variable together with the other independent variable has a bigger effect on the dependent variable.

Example: *Smoking and high cholesterol may contribute to coronary heart disease.* Smoking (X) = IV; Cholesterol (Z) = IV = MV; Coronary heart disease (V) = DV; Smoking together with high levels of cholesterol will increase the chances of coronary heart disease even more. Smoking (IV) may contribute to coronary heart disease (DV); especially if the person has high cholesterol levels (MV). A moderating variable is therefore a second independent variable that is included because it has a significant contributory effect on the IV-DV relationship. **Example:** *The introduction of a four-day working week (IV); will lead to higher productivity (DV);*

especially among younger workers (MV). Four-day working week = (X) =IV; Age = (Z) = IV = MV; Productivity = (V) = DV

Summary

Variable	Definition	Other Terms
Dependent	A variable that is measured to see whether the treatment or manipulation of the independent variable had an effect	Outcome/ Criterion variable
Independent	A variable that is manipulated to examine its impact on a dependent variable	Predictor variable
Control	A variable that is related to the dependent variable, the influence of which needs to be removed	Restricting variable
Extraneous	A variable that is related to the dependent variable or independent variable that is not part of the experiment	Threatening variable
Moderator	A variable that is related to the dependent variable or independent variable and has an impact on the dependent variable	Interacting variable

Validity and Reliability of Measurement

- ***Validity =*** the extent to which a measurement instrument measures what it is intended to measure.
- ***Reliability =*** the consistency with which a measurement instrument yields a certain result when the entity being measured hasn't changed.

Classifying variables for measurement - The Four Scales of Measurement

Variables can be classified into two broad categories as: (1) categorical also known as discrete or qualitative variables and (2) continuous or quantitative variables. Categorical variables can be further categorized as nominal, ordinal or dichotomous. Continuous variables can be further categorized as either interval or ratio variables.

Nominal

Nominal variables are variables that have two or more categories but which do not have an intrinsic order. For example, a real estate agent could classify their types of property into distinct categories such as houses, condos, co-ops or bungalows. So "type of property" is a

nominal variable with 4 categories called houses, condos, co-ops and bungalows. Of note, the different categories of a nominal variable can also be referred to as groups or levels of the nominal variable. Another example of a nominal variable would be classifying where people live in South Africa by province.

In simple terms, nominal variables:

- Measure data by assigning names/labels to them
- Divide data into discrete categories e.g. Male/Female
- One object is different from another.

Dichotomous variables are nominal variables which have only two categories or levels. For example, gender = either "male" or "female". Another example might be if we asked a person if they owned a mobile phone. Here, mobile phone ownership = "Yes" or "No". In the real estate agent example, if type of property had been classified as either residential or commercial then "type of property" would be a dichotomous variable.

Ordinal

Ordinal variables are variables that have two or more categories just like nominal variables only that the categories can also be ordered or ranked. So if you asked someone if they liked the policies of the Democratic Party and they could answer either "Not very much", "They are OK" or "Yes, a lot" then you have an ordinal variable with three categories which you can rank from the most positive (Yes, a lot), to the middle response (They are OK), to the least positive (Not very much). However, whilst we can rank the levels, we cannot place a "value" to them; we cannot say that "They are OK" is twice as positive as "Not very much".

In summary, for ordinal variable,

- Think in terms of symbols (>; <)
- Allow data to be rank-ordered
- One object is bigger or better or there is more of one thing than another.

Interval

Interval variables are variables for which their central characteristic is that they can be measured along a continuum and they have a numerical value (for example, temperature measured in degrees Celsius or Fahrenheit). So the difference between 0C and 20C is the same as 20C to 40C.

To surmise, interval variable

- Has equal units of measurement
- Zero point is established arbitrarily
- Rating scales, such as surveys, assumed to fall on interval scales (beware of differing view)
- One object is so many units (degrees, inches) more than another.

Ratio

Ratio variables are interval variables but with the added condition that 0 (zero) of the measurement indicates that there is none of that variable. Examples of ratio variables include height, mass, distance and many more. The name "ratio" reflects the fact that you can use the ratio of measurements. So, for example, a distance of ten metres is twice the distance of 5 metres.

In short, ratio variables are:

- Characterized by equal measurement units (similar to an interval scale)
- Has an absolute zero point (0 = total absence of the quality being measured)
- Can express values in terms of multiples and fractional parts
- Ratios are true ratios
- One object is so many times as big or bright or tall or heavy as another.

NB: For more on statistics please visit **http://stattrek.com/**

A word of caution when classifying variables

In some cases, the measurement scale for data is ordinal but the variable is treated as interval. For example, a Likert scale that contains five values - strongly agree, agree, neither agree nor

disagree, disagree, and strongly disagree - is ordinal. However, where a Likert scale contains seven or more value - strongly agree, moderately agree, agree, neither agree nor disagree, disagree, moderately disagree, and strongly disagree - the underlying scale is sometimes treated as continuous although where you should do this is a cause of great dispute.

It is worth noting that how we categorise variables is somewhat of a choice. Whilst we categorised gender as a dichotomous variable (you are either male or female), social scientists may disagree with this, arguing that gender is a more complex variable involving more than two distinctions, but also including measurement levels like gender queer, intersex, and transgender. At the same time, some researchers would argue that a Likert scale, even with seven values, should never be treated as a continuous variable.

1.6 BASIC PRINCIPLES OF ETHICAL RESEARCH

Ethics in research are moral principles and values that influence the way a researcher conducts his/her research activities (Ghauri & Gronhaug, 2005:19). There are many ethical dilemmas that confront researchers in their research activities and which they will have to manage. Some of the prominent ones are:

- Protection from harm
- Maintenance of privacy
- Anonymity
- Coercion
- Informed consent
- Confidentiality
- Deception
- Intrusive questioning

Consult your textbooks and other sources to comprehend the nature of these ethical dilemmas and how to deal with them.

1.7 ACTIVITIES

1.7.1. It is argued that the positivist research paradigm lends itself to highly valid and highly reliable research. Bearing this in mind, fully explain why researchers ever follow any other paradigm.

1.7.2. 'Original contribution to knowledge' is a basic condition for a scientific study. But this idea can be misunderstood. Students normally believe that topics used by others in their theses should not be studied because by doing so, they would lose originality. Explain why this may not be the case.

1.7.3 Sensitive questions must sometimes be posed in research. What are the issues when trying to obtain information respondents are reluctant to give? How will you work with sensitive questions so as to get response from respondents?

Suggested Outline Answer:

I) Examples of sensitive information

II) Techniques to retrieve sensitive information

III) Conclusions

UNIT 2: FOCUSING YOUR RESEARCH

WEEK 3

2.1 THE RESEARCH PROBLEM

A research problem refers to some difficulty which the researcher experiences in the context of either a theoretical or practical situation and to which (s)he wants to obtain a solution.

Among others:

- The research problem should address an important question so that the answer will make a difference.
- The research problem should advance the frontiers of knowledge by leading to new ways of thinking, suggesting possible applications, or paving the way for further research in the field.
- Problems that result in a yes or no answer are not suitable problems for research.

Guidelines for stating the research problem

- State the problem clearly and completely.
- Think through the feasibility of the project that the problem implies.
- Say precisely what you mean.
- State the problem in a way that reflects an open mind about its solution.

Guidelines for dividing research problems into sub-problem

Sub problems are the subparts of the main problem that are an integral part of the main problem.

- Each sub problem should be a completely researchable unit.
- Each sub problem must be clearly tied to the interpretation of the data.
- The sub problems must add up to the totality of the problem.
- Sub problems should be few in number.

The research objectives, specific research questions, &/or hypotheses

After the problem is formulated, you should state the goal for your research. The goal is also known as the objective or aim of your study. These objectives specify the focus of the research and can be a brief statement of what the research plans to investigate. The rationale of the research provides reasons why the research is being conducted.

- Objectives are essential to any research.
- Hypotheses are essential to experimental research; research questions are more common in qualitative research.
- Both hypotheses and research questions provide guidance for the kind of data that should be collected.
- Both hypotheses and research questions suggest how data should be analysed and interpreted.
- Hypotheses and research questions may originate in the sub problems.
- Hypotheses and research questions provide a position from which the researcher may initiate an exploration of the problem.
- Hypotheses and research questions act as checkpoints against which to test the findings that the data reveal.

2.2 DELIMITING THE STUDY

- Delimitations of the research are statements about what the researcher is not going to do.
- What the researcher will not do is to become involved in data extraneous (outside) to the research problem.
- The researcher must distinguish between what is and is not relevant to the problem.

2.3 OPERATIONAL DEFINITION FOR KEY TERMS USED IN THE RESEARCH

- ***Operational definition*** is the definition of a characteristic or variable in terms of how it will be measured in the study.
- The researcher must precisely define the terms in the problem and the sub problems.
- Each term should be defined as it will be used in the researcher's project.
- In defining a term, the researcher makes the term mean whatever he or she wishes it to mean within the context of the problem and its sub problems.

2.4 ASSUMPTIONS UNDERLYING THE RESEARCH

- *Assumptions are what the researcher may be taking for granted with respect to the problem.*
- Assumptions are basic to the research problem.
- All assumptions that have a material bearing on the problem should be openly and unreservedly set forth.
- A statement of the assumptions is necessary for others to evaluate the conclusions of the study.

2.5 ACTIVITIES

2.5.1. A lecturer asks 100 students to complete a maths test. The lecturer wants to know why some students perform better than others. Whilst the lecturer does not know the answer to this, (s)he thinks that it might be because of two reasons: (1) some students spend more time revising for their test; and (2) some students are naturally more intelligent than others. As such, the lecturer decides to investigate the effect of revision time and intelligence on the test performance of the 100 students. What are the dependent and independent variables? Justify your answer.

2.5.2. Two statistics lecturers, Sarah and Mike, think that they use the best method to teach their students. Each lecturer has 50 statistics students that are studying for a degree in management. In Sarah's class, students have to attend one lecture and one seminar class every week, whilst in Mike's class students only have to attend one lecture. Sarah thinks that seminars, in addition to lectures, are an important teaching method in statistics, whilst Mike believes that lectures are sufficient by themselves and thinks that students are better off solving problems by themselves in their own time. This is the first year that Sarah has given seminars, but since they take up a lot of her time, she wants to make sure that she is not wasting her time and that seminars improve her students' performance.

Required: (a) State an appropriate aim for the study; (b) State appropriate null and alternate hypotheses.

WEEK 4

3.1 THE PURPOSE OF LITERATURE REVIEW

The literature review enables the researcher to gain insight into the topic and to identify the key issues that need to be explored. The literature review enables us to identify pitfalls and helps us to avoid duplication of a study. It also helps us to identify the gaps in existing knowledge. Literature review is a continuous process that begins before the research problem is finalised and continues until the report is finished. In short we can say that literature review enables us to:

- Demonstrate a familiarity
- Show the path of prior research
- Integrate and summarize
- Learn from others

The benefits of literature reviews

- It can offer new ideas, perspectives, and approaches that may not have occurred to you.
- It can inform you about other researchers who conduct work in the same area.
- It can show you how others have handled methodological and design issues in studies similar to your own.
- It can reveal sources of data that you may not have known existed.
- It can introduce you to measurement tools that other researchers have developed and used effectively.

- It can reveal methods of dealing with problem situations that may be similar to difficulties you are facing.
- It can help you interpret and make sense of your findings and, ultimately, help you tie your results to the work of those who have preceded you.
- It will bolster your confidence that your topic is one worth studying, because you will find that others have invested considerable time, effort, and resources in studying it.

3.2 PRIMARY AND SECONDARY DATA

Information Source	What it Does	Example
General Sources	Provides an overview of a topic and provides leads to where more information can be found.	Daily newspapers, news weeklies, popular periodicals and magazines, trade books, *Reader's Digest Guide to Periodical Literature*
Secondary Sources	Provides a level of information "once removed" from the original work.	Books on specific subjects and reviews of research
Primary Sources	The original reports of the original work or experience.	Journals, abstracts and scholarly books

3.3 SOURCES OF SECONDARY DATA AND HOW TO SELECT THEM

- Identify one or more *keywords*.
- Use the library catalogue.
- Use indexes, abstracts, and other general references.
- Use the library's online database such as:

- Emerald
- Proquest.
- EbscoHost
- SA Citations

- Explore government publications
- Search the World Wide Web.
- Examine citations and reference lists of those who have gone before you.
- Periodicals
- Scholarly journals
- Books
- Dissertations
- Government documents
- Policy reports
- Presented paper

3.4 CONDUCTING LITERATURE REVIEW

- Never take other people's conclusions at face value; determine for yourself whether their conclusions are justified based on the data presented.
- Organize the ideas you encounter during your review.
- Synthesize what you've learned from your review:
    - Compare and contrast varying theoretical positions on the topic.
    - Show how approaches to the topic have changed over time.
    - Describe general trends in the research findings.
    - Identify and explain discrepant or contradictory findings.

- Identify general themes that run throughout the literature.

3.5 REFERENCING

Whenever you write, you are bound to refer to the work of other authors. Each time you do so, it is necessary to identify their work by making reference to it - both in text and in the *reference list* or *bibliography* (please familiarise yourself with the difference in the two terminologies) at the end of your work. This practice of acknowledging authors is known as referencing.

- References must be provided whenever you use someone else's opinions, theories, data, or organisation of material.
- You need to reference information from books, articles, videos, computers, other print or electronic sources, and personal communications.
- A reference is required if you: Quote (use someone else's exact words); copy (use figures, tables, or structure); paraphrase (convert someone else's ideas into your own words); Summarize (use a brief account of someone else's ideas).
- here are a number of different referencing systems used in academic writing.
- **CUT acknowledges the Harvard referencing style for all faculties**

Difference between a reference list and bibliography

The reference list only identifies sources referred to (cite) in the text of your research. You may also be required to provide a bibliography. A bibliography is presented in the same format as a reference list but it includes all material consulted in the preparation of your research. In other words, a bibliography presents the same items as a reference list but it also includes all other sources which you read or consulted but did not cite.

3.6 ACTIVITY

As a group, identify at least ten (10) academically credible articles relevant to your research project for the second semester. Use these articles as the basis of your literature review. Attach these articles to your questionnaire when you submit for evaluation.

Make sure you follow the Harvard method of referencing!

planning and design. Pearson, and (2) Saikkia, J.V. 2009. *Exploring Research*. Pearson.

UNIT 4: SOME KEY ISSUES RELATED TO DATA GATHERING

NB: For more on statistics please visit **http://stattrek.com/**

WEEK 5 & 6

4.1 INTRODUCTION

This unit focuses on three key issues related to data gathering in social research – *measurement*, *reliability* and *validity*. But to begin with, we take look a look at related issues of research planning and methodology to understand what each terminology means and how they invariably affect our three key issues of measurement, reliability and validity. Thereafter, we identify characteristics/main features of research projects. We then spend some considerable time in examining measurement, reliability and validity as well as other issues related to data gathering – including the nature and types of data, sampling, and generalizability.

4.2 COMMON FEATURES OF RESEARCH PROJECTS

- **Universality:** the research project should be one that might be carried out by any competent person. The researcher is a catalyst who collects, organizes, and reports what the collected data seem to indicate.
- **Replication:** the research should be *repeatable*; any other competent person should be able to take the problem and, collecting data under the same circumstances and within the same parameters you have used, achieve results comparable to yours.
- **Control:** the researcher must isolate, or *control*, factors that are central to the research problem; control is important for replication and consistency within the research design.
- **Measurement:** the data should be able to be measured in some way.

4.3 THE NATURE OF RESEARCH DATA

- Data is primary or secondary.
- ***Primary data*** is the layer closest to the truth.
- ***Secondary data*** are derived, not from the truth, but from primary data.
- Data must meet certain criteria to be admitted to study; any data not meeting the criteria are excluded from the study.

4.4 GENERALIZABILITY

Generalizability means—the ability to infer population characteristics based on the sample.

4.5 SAMPLE AND POPULATION

In conducting your research, you might want to involve the whole population or a sample.

- A **sample** is a **representative** subset of the population;
- The **population** is the entire set of participants of interest;
- **Sampling frame** is the correct and complete list of population members available to sample from. This list may or may not differ from the theoretical population.
- A **good sample** has both **accuracy** and **precision**.
- An **accurate sample** is one in which there is **little or no bias**
- A sample with **adequate precision** is one that has **sampling error** that is **within acceptable limit.**

4.6 SAMPLING METHODS

Probability sampling

- The likelihood of any member of the population being selected is known;
- The researcher specifies in advance that each segment of the population is represented in the sample.

Nonprobability sampling

- The likelihood of any member of the population being selected is unknown;
- The researcher has no way of forecasting that each element of the population will be represented in the sample. Some members of the population have little or no chance of being sampled.

Four (4) probability sampling strategies:

- **Simple Random Sampling**
    - Each member of the population has an **equal** and **independent** chance of being chosen
    - The sample is representative of the population

- **Systematic Sampling**
    - Divide the population by the size of the desired sample: e.g., 50/10 = 5
    - Select a starting point at random.
    - Select every 5^{th} name from the starting point

- **Stratified Sampling**

  The researcher samples equally from each of the layers in an overall population.

  The goal of sampling is to select a sample that is representative of the population

But:

- People in a population may differ systematically along some characteristic
- And this characteristic may relate to the factors being studied.

Then stratified sampling is one solution. In stratified sampling,

- The characteristic(s) of interest are identified (e.g., gender)
- The individuals in the population are listed separately according to their classification (e.g., females and males)
- The proportional representation of each class is determined (e.g., 40% females & 60% males)
- A random sample is selected that reflects the proportions in the population(e.g., 4 females & 6 males)

- **Cluster Sampling**

  Instead of randomly selecting individuals:

  - Units (groups) of individuals are identified
  - A random sample of units is then selected
  - All individuals in each unit are assigned to one of the treatment conditions
  - Units must be homogeneous in order to avoid bias
  - Cluster sampling is ideal for populations that are spread out over a large area;
  - The area is subdivided into smaller units; a unit is randomly selected.

Two (2) nonprobability sampling strategies:

- **Convenience Sampling -** Not random and weak representativeness
- **Quota Sampling -** Participants with the characteristic of interest are non-randomly selected until a set quota is met.

Type of Sampling	When to Use It	Advantages	Disadvantages
Probability Strategies			
Simple random sampling	When the population members are similar to one another on important variables	Ensures a high degree of representatives	Time consuming and tedious
Systematic sampling	When the population members are similar to one another on important variables	Ensures a high degree of representatives and no need to use a table of random numbers	Less random than simple random sampling
Stratified random sampling	When the population is heterogeneous and contains several different groups, some of which are related to the topic of study	Ensures a high degree of representatives of all the strata or layers in the population	Time consuming and tedious
Cluster sampling	When the population consists of units rather than individuals	Easy and convenient	Possibly members of units are different from one another, decreasing the technique's effectiveness
Nonprobability Sampling Strategies			
Convenience sampling	When the members of the population are convenient to sample	Convenient and inexpensive	Degree of generalizability is questionable
Quota sampling	When strata are present and stratified sampling is not possible	Ensures some degree of reprentativeness of all the strata in the population	Degree of generalizability is questionable

Table 4.5 Summary of the different types of probability and nonprobability strategies

10004.7 SAMPLING ERROR

- **Sampling error** = difference between sample and population characteristics
- **Bias:** any influence, condition, or set of conditions that singly or in combination distort the data.
- **Sampling Bias**: is any influence that may disturb the randomness by which the choice of a sample population has been selected.

Reducing sampling error

- As sample size increases, sampling error decreases. That is, sampling error is a function of the size of a sample with the error being largest when the sample is small. Therefore, it is important to select a sample of adequate size
- Scrutinize the questionnaire for items that may be influenced by factors that distinguish respondents from non-respondents.
- Compare responses that were returned quickly with those that were returned later (may reflect the kinds of responses that non-respondents would have given.
- Randomly select a small number of non-respondents and match their answers against those of respondents.

4.8 ESTIMATING APPROPRIATE SAMPLE SIZE

- In general, we need a large sample to represent the population accurately when the amount of variability in the population is great – that is there is more diversity in the population.
- In general, the larger the sample, the better because, sampling error tends to be smaller. However, larger samples ignore the power of scientific inference
- For smaller populations (N=100 or fewer), survey the entire population.
- If population is around 500, sample 50%.

- If population is around 1,500, sample 20%.
- If population is over 5,000, a sample size of 400 is fine.
- *The larger the population, the smaller the percentage needed for a representative sample.*
- Larger samples are usually more representative but it is more expensive
- Generally, larger samples are needed when
    - Variability within each group is great
    - Differences between groups are smaller
- Because
    - As a group becomes more diverse, more data points are needed to represent the group
    - As the difference between groups becomes smaller, more participants are needed to reach "critical mass" to detect the difference

4.9 THE DIFFERENT LEVELS OF MEASUREMENT

Level of Measurement	**For Example**	**Quality of Level**
Ratio	Rachael is 5' 10" and Gregory is 5' 5"	Absolute zero
Interval	Rachael is 5" taller than Gregory	An inch is an inch is an inch
Ordinal	Rachael is taller than Gregory	Greater than
Nominal	Rachael is tall and Gregory is short	Different from

- Variables are measured at one of these four levels
- Qualities of one level are characteristic of the next level up
- The more precise (higher) the level of measurement, the more accurate is the measurement process

4.10 VALIDITY AND RELIABILITY

- **Reliability**—tool is consistent
- **Validity**—tool measures "what-it-should"
    - A valid test does what it was designed to do
    - A valid test measures what it was designed to measure

4.12.1 CLASSIFYING VALIDITY

There are various ways of classifying validity. One way of doing it is either as:

- **Internal Validity:** the extent to which the design and data of a research study allow the researcher to draw accurate conclusions about cause-and-effect and other relationships within the data.

  Or

- **External Validity:** the extent to which the results of a research study apply to situations beyond the study itself; the extent to which conclusions can be generalized.

  It can also be classified as follows:

- **Face Validity:** the extent to which an instrument looks like it's measuring a particular characteristic; relies on subjective judgment.
- **Content Validity:** the extent to which a measurement instrument is a representative sample of the content area being measured.

- **<u>Criterion Validity</u>:** the extent to which the results of an assessment correlate with another, related measure.
- **<u>Construct Validity</u>:** the extent to which an instrument measures a characteristic that cannot be directly observed but is assumed to exist (a construct, such as intelligence).

Type of Validity	What is it?	How do you establish it?
Content	A measure of how well the items represent the entire universe of items	Ask an expert if the items assess what you want them to
Criterion	A measure of the extent to which the results of an assessment correlate with another, related measure	Compare your measure with another measure of the same phenomenon
Concurrent	A measure of how well a test estimates a criterion	Select a criterion and correlate scores on the test with scores on the criterion in the present
Predictive	A measure of how well a test predicts a criterion	Select a criterion and correlate scores on the test with scores on the criterion in the future
Construct	A measure of how well a test assesses some underlying construct	Assess the underlying construct on which the test is based and correlate these scores with the test scores

4.12.2 CLASSIFYING RELIABILITY

- **Interrater reliability**: the extent to which two or more individuals evaluating the same product or performance give identical judgments.
- **Internal consistency reliability:** the extent to which all of the items within a single instrument yield similar results.
- **Equivalent forms reliability:** the extent to which two different versions of the same instrument yield similar results.
- **Test-retest reliability:** the extent to which a single instrument yields the same results for the same people on two different occasions.

Type of Reliability	**What It Is**	**How You Do It**	**What the Reliability Coefficient Looks Like**
Test-Retest	A measure of stability	Administer the same test/measure at two different times to the same group of participants	$r_{test1 \bullet test1}$
Parallel Forms	A measure of equivalence	Administer two different forms of the same test to the same group of participants	$r_{form1 \bullet form2}$
Inter-Rater	A measure of agreement	Have two raters rate behaviours and then determine the amount of agreement between them	Percentage of agreements
Internal Consistency	A measure of how consistently each item measures the same underlying construct	Correlate performance on each item with overall performance across participants	Cronbach's alpha Kuder-Richardson

4.12.3 HOW TO INCREASE RELIABILITY

- Increase sample size
- Eliminate unclear questions
- Standardize testing conditions
- Moderate the degree of difficulty of the tests
- Minimize the effects of external events
- Standardize instructions
- Maintain consistent scoring procedures

4.12.4 RELATIONSHIP BETWEEN RELIABILITY AND VALIDITY

A valid test must be reliable but a reliable test need not be valid

NB: For more on statistics please visit **http://stattrek.com/**

4.13 Activity

4.13.1 There are two major criteria for evaluating measurements that you have covered in this unit. Fully explain how you would ensure that in your research endeavour you abide by these criteria.

4.13.2 Are large samples always necessary? Discuss.

Suggested outline for your answer:

I) The rationale of sampling

II) Determinants of sample size
   a) Desired level of accuracy
   b) Heterogeneity of the population

III) Sample size and costs
   a) Marginal accuracy benefits
   b) Increasing the marginal benefits through advanced probability samples
   c) Reduced costs of non-probability samples

UNIT 5: QUALITATIVE RESEARCH

WEEK 7

5.1 THE MEANING OF RESEARCH DESIGN

Research design is defined in several ways including but not limited to:

1. A plan outlining how data is to be collected, the data gathering methods, the instruments to be used, how the instruments will be administered, and how the data will be organized and analysed.
2. A plan listing type of data to gather, from whom, how and when to collect the data, and how to analyse the data obtained.
3. A systematic plan that guides research according to the scientific method.

Research can be classified under two broad categories namely **quantitative** and **qualitative** research based on certain characteristics.

5.2. THE NATURE OF QUALITATIVE RESEARCH

The aim of qualitative research is to deepen our understanding about something, and usually this means going beyond the numbers and the statistics. Qualitative research helps us to give reasons why the numbers tell us what they do. **Qualitative** research involves looking at characteristics, or qualities, that cannot easily be reduced to numerical values. In this type of research, one attempts to examine nuances and complexities of a particular phenomenon.

4. Focus is on phenomena that occur in natural settings
5. Involves studying those phenomena in all of their complexity
6. Researchers recognize the multifaceted form of the issue they are studying
7. The researcher is the instrument
8. Qualitative researchers formulate only general research problems and ask only general questions about the phenomena they study.

9. Qualitative researchers may not identify ahead of time the exact methods they will use.
10. The methodology may continue to evolve over the course of the study.
11. Qualitative research requires considerable preparation and planning.
12. Qualitative studies do not allow the researcher to identify cause-and-effect relationships.
13. Qualitative research uses descriptions rather than numerical measurements.

5.3 TYPES OF QUALITATIVE RESEARCH DESIGN

There are several types of qualitative research but in this unit, we focus on:

- Case Study
- Ethnography
- Phenomenological Study
- Grounded Theory Study
- Content Analysis

Case Study

- Definition: In a case study, a particular individual, program or event is studied in depth; findings may not be generalizable.
- Method: The researcher collects extensive data on the individual(s), program(s), or event(s) on which the investigation is focused.
- Data Analysis: involves (1) organization of details about the case, (2) categorization of data, (3) interpretation of single instances, (4) identification of patterns, and (5) synthesis and generalizations.

Ethnography

- Definition: In ethnography, the researcher looks in depth at an entire group that shares a common culture; it is especially useful for understanding complexities of a particular, intact sociocultural group.
- Method: Site-based fieldwork is the essence of any ethnography. The researcher depends on a gatekeeper and key informants and is a careful observer.
- Data Analysis: Data collection and data analysis occur somewhat simultaneously. Steps in data collection include (1) description, (2) analysis, and (3) interpretation. The research strives for rigorous subjectivity.

Phenomenology

- Definition: A phenomenological study attempts to understand people's perceptions, perspectives, and understandings of a particular situation. The researcher tries to answer the question, "What is it like to experience such-and-such?"
- Method: Phenomenological researchers depend almost exclusively on lengthy interviews with a carefully selected sample of participants. The phenomenological interview is often unstructured in which the researcher and participants to arrive at "the heart of the matter."
- Data Analysis: The central task in data analysis is to identify common themes in people's descriptions of their experiences. Steps include (1) identifying relevant statements, (2) grouping statements into "meaning units," (3) seeking divergent perspectives, and (4) constructing a composite.

Grounded theory

- Definition: The major purpose of a grounded theory study is to begin with the data and use them to develop a theory. The study uses a prescribed set of procedures for analysing data and constructing a theoretical model from them—the theory is "grounded" in the data.

- Method: Data collection is field-based, flexible, and likely to change over the course of the study. Interviews typically play a major role. A *constant comparative method* is used in that data analysis drives later data collection.
- Data Analysis: Data analysis includes (1) open coding, (2) axial coding, (3) selective coding, and (4) development of a theory. No matter what form the theory takes, it is based entirely on the data collected.

Content Analysis

- Definition: A content analysis is a detailed and systematic examination of the contents of a particular body of material for the purpose of identifying patterns, themes, or biases. A content analysis is typically performed on forms of human communication and involves the greatest amount of planning at the front end of the project.
- Method: A content analysis is systematic and includes (1) identification of the material to be studied, (2) definition of the characteristics to be studied, (3) a breakdown of complex items into smaller segments, and (4) scrutiny of material for identified characteristics under study.
- Data Analysis: Data analysis involves tabulation of the frequency of each characteristic found in the material studied. Tabulations and statistical analyses are used to interpret the data.

5.4 DATA COLLECTION IN QUALITATIVE RESEARCH

- Multiple forms of data are used.
- Data collected early in the investigation may influence subsequent data.
- Potential sources of data are unlimited.
- Data collection takes a great deal of time.
- Data collection methods should be consistent with ethical principles of research studies.

Sampling

- Choice of sample depends on the research question
- Tends to be non-random in selection of data sources
- Is purposeful

Observations

- Researcher may be an outsider or a participant observer
- Intentionally unstructured and free-flowing
- Can be problematic and lack objectivity

Interviews

The most common forms of qualitative research are *face-to-face interviews* and *focus groups*. Face-to-face interviews are just that: Meeting someone in person and discussing various issues. The *informant* – or person you are interviewing – may be an expert in a particular field (e.g. the editor of a newspaper) or they may be someone who is affected by the issues you are researching (e.g. someone who is HIV positive or who reads the media). Although it is very important to develop a list of questions you want to ask someone, face-to-face interviews usually involve more than 'yes' or 'no' answers. The point is to try to understand the *complexity* of the issues you are researching. The nature of face-to-face interviews is that they are usually quite discursive. Interviews:

- Can yield a great deal of information
- Tend to be unstructured and open-ended

Research tip: Face-to-face interviews

- Always prepare a set of questions to ask the informant;
- It is a good idea to record your interviews, so that you can check your facts later. Take notes during the interview, if you feel comfortable doing this;
- Remember: Interviews take time, and the informant is giving you his or her time for free. Interviews shouldn't really take more than an hour, unless the informant wants

to spend more time talking to you. Usually 10 questions are enough for this amount of time;

- Sometimes people transcribe the interview recordings. This usually makes analysing the results easier, but it also takes time and can be quite an effort. Consider including a budget in your research proposal for transcription, and then pay someone else to do it;
- Sometimes you may need to ask the informant if they are prepared to be identified in your research, or if they would like to be quoted anonymously. This is usually the case if their identity needs to be protected, and sometimes if you are researching a controversial topic;
- You may want to consider letting your informant review any direct quotes you use before publishing the research report. However, this takes time and sometimes can delay the research process;
- Ask the informant if you can include their contact details in your research report;
- Ask the informant if he or she would like to be alerted when the research is published, and let them know where they can read the report if it is publicly available.

Advantages of face-to-face interviews

- Can allow for in-depth knowledge sharing;
- Helps to develop the bigger picture;
- Helps with analysis of results;
- Good for networking (e.g. you may be referred to other people to interview).

Disadvantages of face-to-face interviews

- Can be time consuming;
- May be difficult to arrange an interview time;
- Can be difficult to compare and analyse information.

Focus groups

Focus groups involve discussions with two or more participants. While questions for focus groups need to be prepared to guide and focus the discussions, the responses are often free-ranging, as the participants are encouraged to explore the issues at hand in an in-depth way. While focus groups and interviews will help you develop explanations for quantitative data, sometimes they can provide you with quantitative data themselves. For example, you might find that 20% of the participants in a focus group discussion did not like the way HIV/AIDS positive people were portrayed in the media (quantitative data). Then you might find that the reasons (qualitative information) they gave were that:

- They found it demeaning;
- They thought that it was insensitive;
- They thought that HIV/AIDS positive people were treated as 'others'.

With focus groups and interviews, it is usual to write up the responses to your questions, to arrange and analyse the responses in a careful and meaningful way, and to include the most relevant ones in your *research report.*

Tip: Focus groups

- Focus groups can sometimes take time to arrange, so prepare in advance. Try to find an intermediary to help you (an organization or individual in close contact with the potential focus group participants);
- Think about who you want to participate in the focus group by referring to your *research question*. What age group should they be? Should they be male or female? Should they come from a particular income bracket? You may want to consider holding separate focus groups for different age groups, or for different genders. For example, it may be important to hold a separate focus group for males and females if you are discussing sex and sexuality;
- Issues of power: The focus group facilitator holds an immense amount of power in the discussions. You need to keep this in mind. If a male facilitator questions young girls about sexuality, will that affect the research results? Some researchers will not let the people commissioning the research (e.g. government) be present in the focus group

sessions. In some cases, special rooms are built with one-way glass so that the session can be observed unobtrusively;

- You need to find out if it is normal to pay focus group participants, and what the going rate is. Often focus group participants come from poorer communities. If it is not normal to pay participants, you may want to consider it and set a trend!

Advantages of focus groups

- Good for community participation (grassroots input);
- Helpful in developing ideas and sharing latent, or hidden, knowledge spontaneously;
- Enables you to get information from a number of individuals simultaneously.

Disadvantages of focus groups

- Can be difficult to set up;
- Participants may need to be paid;
- Need to be sensitive to who the facilitator is;
- May need a translator;
- Sometimes difficult to organize and analyse information.

5.4. DATA ANALYSIS IN QUALITATIVE RESEARCH

1. Peruse the data set to get a sense of what it contains as a whole, making notes of associations or ideas that occur to you;
2. Organize the data into similar categories;
3. Attempt to identify patterns (general categories/themes and subthemes). Then classify data accordingly.
4. If you have done quantitative research at the same time, try to match some of your qualitative results to your quantitative results. Where are the links? How do the results 'speak to' or explain each other? What conclusions can be reached that aren't obvious at first glance?

5. Be methodical. Think step by step, and explain your assumptions if you have to. Remember, your research results might not tell you everything you want to know but don't be afraid to say what they *do* tell you. Be careful with your words, and be specific.
6. And summarize the data.

Activity 5.1 Case studies in qualitative research offer unique opportunities but can also pose major problems**. Please answer the following questions**

(a) In terms of research what does case study involve?

(b) Explain the major advantages and limitations of case studies

UNIT 6: CLASSIFYING RESEARCH- QUANTITATIVE RESEARCH

WEEK 8 & 9

6.1 QUANTITATIVE RESEARCH DESIGNS

Quantitative research involves quantifying relationships between variables. In this type of study, you measure variables on subjects, and express the relationship between variables using statistics, such as correlations, relative frequencies, or differences between means. In this unit, and the next one, we focus on the design of quantitative research.

Studies aimed at quantifying relationships can be seen to be of two broad types: **descriptive** and **experimental**. In a descriptive study, you measure things as they are. In an experimental study, you first take measurements, try some sort of intervention, and then take measurements again to see what happened. **The quantitative research process** basically involves:

⇒ formation of one or more hypotheses

⇒ review of the related literature

⇒ Collection and analysis of data.

6.2 DESCRIPTIVE STUDIES

Descriptive studies are carried out on subjects without intervening. Although there are many types, in this unit, we focus on the following types of descriptive studies: observation Studies; correlational research; and survey research.

6.2.1 OBSERVATION STUDIES

Characteristics:

- Usually involves humans, animals, plants, non-living objects

NB: The material in this study guide is largely based on: (1) Leedy, PD & Ormorod, JE. 2010. *Practical Research, planning and design*. Pearson; and (2) Salkind, JN. 2009. *Exploring Research.* Pearson.

- Tends to have a particular pre-specified focus
- Phenomenon being studied is quantified in some way
- Involves meticulous attention to detail

Maintaining objectivity in observation studies

- Define the phenomenon studied precisely and concretely so that it is easily recognized.
- Divide the observation period into small segments and record whether the phenomenon does or does not occur in each segment.
- Use a rating scale to evaluate the phenomenon in terms of specific dimensions.
- Have two or three people rate the same phenomenon independently, without knowledge of one another's ratings.
- Train the raters to use specific criteria when counting or evaluating the phenomenon; continue training until consistent ratings are obtained for any single occurrence of the phenomenon.

6.2.2 CORRELATIONAL RESEARCH

- A correlational study examines the extent to which differences in one variable or characteristic are related to differences in one or more other variables or characteristics.
- In correlational studies, researchers gather data about two or more characteristics for particular units of study in order to determine whether and in what way these characteristics might be interrelated.
- Correlational data is plotted on a *scatter plot*.
- Correlation does not, in and of itself, indicate causation.

6.2.3 SURVEY RESEARCH

- Involves acquiring information about one or more groups of the study unit - about their characteristics, opinions, attitudes, etc. – by asking them questions and tabulating the answers.
- The goal is to learn about a large population by surveying a sample of that population.
- Also called a *descriptive survey* or *normative survey.*
- Simple design – the researcher poses a series of questions, quantifies the responses, and draws inferences about a particular population from the responses of the sample.
- Captures a moment of time; by drawing conclusions from the transitory collection of data, extrapolation can be made about state of affairs over a longer period of time.
- Relies on self-reported data.

6.2.3.1 Face-to-face interviews in survey research

- Structured
- Enables the researcher to establish rapport with participants
- Yields the highest response rates in survey research
- Time and expense involved may be prohibitive

6.2.3.2 Guidelines for: conducting interviews in a quantitative Study

1. Identify questions in advance.
2. Consider how participants' background may influence responses.
3. Make sure interviewees are representative of the group.
4. Find a suitable location.
5. Get written permission.
6. Establish and maintain rapport.

7. Don't put words in people's mouths.
8. Record responses verbatim.
9. Keep your reactions to yourself.
10. As you write questions, think about how to quantify responses.
11. Pilot-test the questions.
12. Restrict each question to a single idea.
13. Save controversial questions for the latter part of the interview.
14. Seek clarifying information when necessary.

6.2.3.3 Measurement instruments (questionnaires)

The most popular instrument used to collect data from subjects comes in the form of a questionnaire. Questionnaires are most often, paper-and-pencil set of structured and focused questions. Unlike in face-face interviews, a questionnaire:

- Save time because individuals can complete them without any direct assistance (they can be self-administered).
- Can be sent out to large groups of people over a large geographical area
- Participants can respond to questions with assurance of remaining anonymous and thus may be more truthful than in face-to-face or telephone interviews
- Sometimes, have a low return rate
- Often make use of **checklists** or a form of **rating scales.**

Checklists and Rating Scales

- Checklist: a list of behaviours, characteristics, or other entities under investigation.
- Rating Scale: used when a behaviour, attitude, or other phenomenon of interest needs to be evaluated on a continuum ("never" to "always")

Some of the notable rating scales are the **Thurnstone scale** and the **Likert scale**.

Thurnstone scale:

- It is a method used of measuring attitudes. They are close to measuring at interval level
- For fuller discussion of a Thurnstone scale, consult Salkind page 136 onwards.

Likert scale

- It is simple to develop and widely used
- Its development is similar to Thurnstone scale but is much simpler in design

NB: You must thoroughly read and understand (1) what makes questionnaires work; (2) the basic assumptions of the questionnaire; (3) format of questions; and (4) format of the questionnaire by consulting Salkind page 142 onwards

6.2.3.4 Guidelines for constructing a questionnaire

1. Keep it short and simple.
2. Provide clear instructions by using simple, clear, unambiguous language.
3. Check for unwarranted assumptions implicit in the question.
4. Word questions in ways that don't give clues about desirable responses.
5. Determine in advance how you will code the responses.
6. Check for consistency.
7. Conduct one or more pilot tests to determine the validity of your questionnaire.
11. Scrutinize the almost-final product to make sure it addresses your needs.
12. Make the questionnaire attractive and professional looking.

6.3 Activities

Compare quantitative and qualitative research using the structure below for your answer.

I) How to determine whether research is qualitative or quantitative
II) Problems that are more suited for qualitative and quantitative research respectively
III) Objectives in qualitative and quantitative research respectively
   a) Description
   b) Exploration
   c) Explanation
   d) Prediction
IV) Conclusions

	Qualitative	Quantitative
Definition		

UNIT 7: STATISTICAL TECHNIQUES FOR DATA ANALYSIS

WEEKS 10 & 11

7.1 PREPARING DATA FOR ANALYSIS

Principles – Exploration of Data:

- Where two variables are concerned, one of the variables becomes dominant and governs meaning that emerges from the other.
- Whatever the researcher does with the data to prepare it for inspection or interpretation will affect the meaning that the data reveal. Therefore, every researcher should be able to provide a clear, logical rationale for the procedure used to arrange and organize the data.

7.2 CLASSIFYING STATISTIC AS EITHER DESCRIPTIVE OR INFERENTIAL

Statistics refers to a group of computational procedures that enable us to find patterns and meaning in numerical data.

There are two major functions of statistics:

- They ***describe*** what the data looks like; this is the function of **descriptive statistics.**
- They allow us to make ***inferences*** about large populations by collecting data on relatively small samples; this is the function **inferential statistics.**

7.3 TYPES OF DESCRIPTIVE STATISTICS

Statisticians use summary measures to describe patterns of data.

Measures of central tendency

Measures of central tendency refer to the summary measures used to describe the most "typical" value in a set of values. The two most common measures of central tendency are the median and the mean, which can be illustrated with an example. Suppose we draw a sample of five women and measure their weights. They weigh 100 pounds, 100 pounds, 130 pounds, 140 pounds, and 150 pounds.

- Mode: the single number or score that occurs most frequently.
- Median: the numerical centre of a set of data.
- Mean: the arithmetic average of the scores within the data set.

To find the **median**, we arrange the observations in order from smallest to largest value. If there is an odd number of observations, the median is the middle value. If there is an even number of observations, the median is the average of the two middle values. Thus, in the sample of five women, the median value would be 130 pounds; since 130 pounds is the middle weight.

The **mean** of a sample or a population is computed by adding all of the observations and dividing by the number of observations. Returning to the example of the five women, the mean weight would equal (100 + 100 + 130 + 140 + 150)/5 = 620/5 = 124 pounds. In the general case, the mean can be calculated, using one of the following equations:

$$\text{Population mean} = \mu = \Sigma X / N \quad \text{OR} \quad \text{Sample mean} = x = \Sigma x / n$$

Where ΣX is the sum of all the population observations, N is the number of population observations, Σx is the sum of all the sample observations, and n is the number of sample observations.

The Mean vs. the Median

As measures of central tendency, the mean and the median each have advantages and disadvantages. Some pros and cons of each measure are summarized below.

- The median may be a better indicator of the most typical value if a set of scores has an **outlier**. An outlier is an extreme value that differs greatly from other values.
- However, when the sample size is large and does not include outliers, the mean score usually provides a better measure of central tendency.

Measures of variability/dispersion

Statisticians use summary measures to describe the amount of variability or spread in a set of data. The most common measures of variability are the range, the interquartile range (IQR), variance, and standard deviation

- **Range:** indicates the spread of data from lowest to highest value (highest score – lowest score)
- **Average Deviation:** the average of differences of each score in a set of scores and the mean score.
- **Standard Deviation:** the measure of variability most commonly used in statistical procedures; the square of the score-mean differences.

The **standard deviation** is the square root of the variance. Thus, the standard deviation of a population is:

$$\sigma = \text{sqrt}\left[\sigma^2 \right] = \text{sqrt}\left[\Sigma \left(X_i - \mu \right)^2 / N \right]$$

Where σ is the population standard deviation, σ^2 is the population variance, μ is the population mean, X_i is the *i*th element from the population, and N is the number of elements in the population.

And the standard deviation of a sample is:

$$s = \text{sqrt}\left[s^2 \right] = \text{sqrt}\left[\Sigma \left(x_i - x \right)^2 / \left(n - 1 \right) \right]$$

Where *s* is the sample standard deviation, s^2 is the sample variance, x is the sample mean, x_i is the *i*th element from the sample, and n is the number of elements in the sample.

- **The Variance**: In a population, **variance** is the average squared deviation from the population mean, as defined by the following formula:

$$\sigma^2 = \Sigma \left(X_i - \mu \right)^2 / N$$

Where σ^2 is the population variance, μ is the population mean, X_i is the *i*th element from the population, and N is the number of elements in the population. The variance of a sample, is defined by slightly different formula, and uses a slightly different notation:

$$s^2 = \Sigma \left(x_i - x \right)^2 / \left(n - 1 \right)$$

Where s^2 is the sample variance, x is the sample mean, x_i is the *i*th element from the sample, and n is the number of elements in the sample. Using this formula, the sample variance can be considered an unbiased estimate of the true population variance. Therefore, if you need to estimate an unknown population variance, based on data from a sample, this is the formula to use.

Measures of association: correlation and correlation coefficient

Correlation: a measure of the relationship between two variables; correlation indicates the strength of the relationship.

Correlation Coefficient: a number between -1 and +1; most correlation coefficients are decimals (positive or negative) somewhere between these two extremes.

Positive Correlation: as one variable increases, the other variable also increases.

Negative Correlation: as one variable increases, the other variable decreases.

Pearson *r*: the most widely used statistic for measuring correlation.

Note: *Correlation does not necessarily indicate causation*.

7.4 TYPES OF INFERENTIAL STATISTICS

Parameter Vs Statistic - *A* value that defines a property of a population is called a parameter. A value that describes a characteristic of a sample is called a statistic. The mean of a population is a parameter but the mean of a sample is a statistic. In analysing quantitative data, we use different symbols to refer to parameters and statistic. Although researchers collect data from ma population, the goal is to explain characteristics of the population using sample data. This goal is achieved through statistical inference.

The two main functions of inferential statistics are:

- To estimate a population parameter from a random sample; and
- To test statistically based hypotheses.

Estimating population parameters

- In conducting research, we use a sample to learn about the larger population from which the sample was drawn.
- ***Inferential statistics*** inform how closely sample statistics approximate parameters of the overall population.
- Statistical estimates of population parameters are based on the assumption that ***the sample is randomly chosen and representative of the total population*.**
- ***Error:*** the difference between the population mean and the sample mean.
- ***Standard Error of the Mean:*** indicates how much a particular mean is likely to vary from one sample to another *when all samples are the same size and are drawn randomly from the same population*.

Hypotheses testing

The greatest problem in hypothesis testing is to find ways to structure hypotheses in such a way that we can test them effectively. Typically, it is important to adhere to the following.

1.	Define the **research hypothesis** and set the **parameters** for the study.
2.	Set out the **null** and **alternative hypothesis** (or more than one hypothesis; in other words, a number of hypotheses).
3.	Explain how you are going to **operationalize** (that is, **measure** or **operationally define**) what you are studying and set out the **variables** to be studied.
4.	Set the **significance level**.
5.	Make a **one-** or **two-tailed prediction**.
6.	Determine whether the distribution that you are studying is **normal** (this has implications for the types of statistical tests that you can run on your data).
7.	**Select** an appropriate **statistical test** based on the variables you have defined and whether the distribution is normal or not.
8.	**Run** the **statistical tests** on your data and **interpret** the **output**.
9.	**Accept** or **reject** the **null hypothesis**.

By the way:

- Null Hypothesis: is a statistical hypothesis which postulates that any result observed is the result of chance alone.
- Testing the Null Hypothesis: means the process of comparing observed data in a research study with what we would expect from chance alone.
- Significance Level: means the probability that researchers use as a cut-off point to decide that a result has not occurred by chance. It asks the question: What is the probability that a score could have arisen by chance? Typically, if there was a 5% or less chance (5 times in 100 or less) that a score from the focal distribution could not have come from the comparison distribution, we would accept the null hypothesis. Alternately, if the chance was greater than 5% (6 times in 100 or more) we would reject the null hypothesis. There is relatively little justification why a significance level of 0.05 is used rather than any other value. However, if we want to be particularly confident in our results, we set a more stringent level of 0.01 (a 1% chance or less; 1 in 100 chance or less).
- The core of hypothesis testing is **comparing distributions- Frequency Distributions**

Making errors in hypothesis testing

- Type I Error: the erroneous conclusion that a result was not due to chance when in fact it was due to chance; incorrectly rejecting the null hypothesis.
- Type II Error: the erroneous conclusion that a result was due to chance when in fact it was not; incorrectly failing to reject a null hypothesis that is actually false; also known as a *beta error*.

7.5 SUGGESTIONS FOR INCREASING THE POWER OF A STATISTICAL TEXT

- Use as large a sample as is reasonably possible.
- Maximize the validity and reliability of your measures.
- Use parametric rather than nonparametric statistics whenever possible.

7.6 INTERPRETATION OF THE DATA

Interpreting the data means several things including:

- Relating the findings to the original research problem and to the specific research questions and hypotheses.
- Relating the findings to pre-existing literature, concepts, theories, and research studies.
- Determining whether the findings have practical significance as well as statistical significance.
- Identifying limitations of the study.

7.7 ACTIVITY

The following distribution of sales was obtained for a company during one week.

Article	Frequency	Percentage
Men‟s T-shirts	8	
Evening dresses	2	
Summer dresses	5	
Sportswear Shoes	3	
Men‟s shorts	4	
Ladies T-shirts	3	
Ladies Shorts	5	
Kid‟s wear	3	
	n =40	

7.7.1

(a) Calculate the percentages for this distribution and complete the last column of this table.

(b) What article was presented by the largest percentage of sales?

(c) Represent graphically this frequency distribution.

7.7.2 Calculate the standard deviation of the following *sample* scores: (**10,12,13,15,18,20,22,25)**

7.7.3 Calculate the standard deviation of the following population scores: (**3,4,4,6,2**)

7.7.4. Nineteen (19) students take part in an experiment and the score they obtained are as follows: **1, 2, 3, 3, 7, 5, 8, 5, 7, 2, 7, 9, 8, 7, 3, 5, 3, 5, 5**

(a) What is the average performance of the sample?

(b) What is the average deviation of scores from the mean?

(c) Calculate the Z-scores for each raw score.

(d) What is the mean of the Z-scores distribution?

(e) What is the variance and standard deviation of the Z-scores distribution?

(f) Produce two frequency distributions (1) for raw scores (2) for Z-scores.

(g) Is there a difference in the shape of the two distributions?