NB for Chapter Three

- The chapter is written in future tense when developing the proposal. However when writing the final document this chapter tenses change to past tense
- From section 3.1 to 3.8 you need to introduce the section by defining the section heading, explain why it is important. Briefly explain the different categories/types. Identify and justify the choice made.

Components in Chapter Three

3.1 Research Design

Define research design

Discuss types of research design

3.2 Population, Sample and Sampling

You are expected to explicitly state the size of population, what is the population? where is it defined? What is the sample size? How do you calculate the sample size? And what is the sampling technique use?

Remember you need to define what population, sampling is? Briefly discuss other sampling techniques, justify why the chosen technique is the best for the study.

3.3 Data Collection, Instrumentation and Procedure

Define what data collection tools is

Briefly discuss the different data collection tools

Justify the choice of the data collection tool

Explain how the data collection tool will be used (procedure)

Attach evidence as an appendix

3.4 Development tools and Material

- Identify all the software's that will be used to implement the system you intend to develop
- Identify the environment/ platform the system will run in. e.g windows 10.
- Identify all machine specifications that will be required for the system to run e.g. 10GB RAM
- The information provided should be written in continuous pros, NOT in point form.

3.5 System Development Methodology

- Define what System development methodology is.
- Briefly explain the SDM that exist, giving a citation in each case e.g. RAD, Incremental, Agile, Spiral, Water fall e.t.c
- Identify a suitable methodology and justify why that is the best methodology for your system
- **Nb**: Do not explain the phases in the methodology

3.6 Data Processing and Analysis

Data processing: Data is raw facts that need to be processed to give meaning. E.g. the data collected using questionnaires needs to be organized in a way that it will give meaning. Processing involves editing, coding, classification and tabulation.

Editing: The purpose of editing is that careful scrutiny of all collected questionnaires to produce completeness, error-free and readability

Coding: The purpose of coding is the assigning codes (numbers) for each category of answers, for example the code No 1 for the answer less than 25%, the code No 2 for the answer 26% up to 50% and so

Classification: The purpose of classification is to divide the received questionnaires on the basis of their groups

Tabulation: The purpose of tabulation is the process of summarizing data and displaying them in the appropriate tables that further analysis are to be facilitated (explain how tabulation will be done).

Data Analysis

This step has vital impact on research process so that the testing of pre-determined hypotheses would be implemented. So far we have collected a mass of data that through the previous steps has been proceed, however they are unable to generalize any information. In other words whenever the mass of data is collected the statistics comes into account and it creates the procedures to support processing of data and also analysis of data.

There are two statistical measures that a researcher can use to analyze data. i.e. descriptive and inferential statistics to be able to analyze data.

Descriptive statistics: **Descriptive statistics** are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics **analysis**, they form the basis of virtually every quantitative **analysis** of data. **There are four major types of descriptive statistics:**

- Measures of Frequency: * Count, Percent, Frequency. ...
- Measures of Central Tendency. * Mean, Median, and Mode. ...
- Measures of Dispersion or Variation. * Range, Variance, Standard Deviation. ...
- Measures of Position. * Percentile Ranks, Quartile Ranks.

For example, suppose a pet shop sells cats, dogs, birds and fish. If 100 pets are sold, and 40 out of the 100 were dogs, then one description of the data on the pets sold would be that 40% were dogs.

This same pet shop may conduct a study on the number of fish sold each day for one month and determine that an average of 10 fish were sold each day. The average is an example of descriptive statistics.

A graphical representation of data is another method of descriptive statistics. Examples of this visual representation are histograms, bar graphs and pie graphs, to name a few. Using these methods, the data is described by compiling it into a graph, table or other visual representation.

This provides a quick method to make comparisons between different data sets and to spot the smallest and largest values and trends or changes over a period of time. If the pet shop owner wanted to know what type of pet was purchased most in the summer, a graph might be a good medium to compare the number of each type of pet sold and the months of the year.

Inferential statistics use a random sample of data taken from a population to describe and make inferences about the population. ... You can measure the diameters of a representative random sample of nails.

Types of inferential statistics include:

T-test: used when comparing two groups on a dependent variable, the data must be sorted according to the independent variable

Analysis of Variance (ANOVA): When comparing three or more groups on one dependent variable

Correlation: Correlations should be calculated to examine the relationship between two variables within the same group of participants. It is useful when a researcher wants to establish if there are possible connections between variables. For example the relationship between academic achievement and achievement motivation, height and weight

Regression analysis: Regression analysis is used to model the relationship between a response variable and one or more predictor variables.

3.7 Ethical Consideration

Ethics are the norms or standards for conduct that distinguish between right and wrong. ... First, **ethical** standards prevent against the fabrication or falsifying of data and therefore, promote the pursuit of knowledge and truth which is the primary goal of **research**.

An accumulation of values and principles that address questions of what is good or bad in human affairs. **Ethics** searches for reasons for acting or refraining from acting; for approving or not approving conduct; for believing or denying something about virtuous or vicious conduct or good or evil rules.

3.8 Summary

Do a summary of the entire chapter

CHAPTER THREE
METHODOLOGY

3.0 Overview

In this research a system for allowing students to register online will be developed. The new system will aim to bridge the gap identified during literature review. The chapter outlines the manner in which the study will be conducted. The key components are research design, population and sampling, data collection methods, development tools and material, system development methodology, system design and data processing and analysis.

3.1 Sample Research Design

A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure *(citation)*. It constitutes the blueprint for the collection, measurement and

analysis of data *(citation)*. As such the design includes an outline of what the researcher will do from writing the research objectives and its operational implications to the final analysis of data. A research design can be explanatory in which a researcher uses his own imaginations and ideas. It is based on the researcher personal judgment and obtaining information about something (*citation)*.A research design can also be descriptive where a researcher is interested in describing a particular situation or phenomena under his study. It is a theoretical type of researcher design based on the collection designing and presentation of the collected data (*citation)*. A research can also be diagnostic in which the researcher wants to know about the root causes of the problem. He describes the factors responsible for the problematic situation (*citation)*. Finally a research design can be experimental which is the direction taken by this study. An experimental design uses in natural science but it is different in social sciences. Human behavior cannot be measured through test-tubes and microscopes. The social researcher use a method of experiment in that type of research design (*citation)*.. One group is subjected to experiment called independent variables while other is considered as control group called dependent variable. The result obtained by the comparison of both the two groups (*citation)*.. Both have the cause and effect relationship between each other. The research design will be appropriate for developing the student online registration system because……*(provide the justification for choosing the research design).* A quantitative approach will be adopted to be able to effectively show how the system will improve the number of students registering per day.

3.2 Population, Sample and Sampling

Population is a well-defined collection of individuals or objects known to have similar characteristics (*citation)* while a target population is the *…..(Citation)*. The target population in this research will comprise of all 2500 registered students at Meru University of Science and Technology. The choice of the target population was based on the fact that it is convenient for researcher to gather data *(this is the justification)*.

Sampling is a process of selecting a portion of objects/ individuals from a group or population to become the foundation on estimating and predicting outcome of the population *(citation)*. Also (citation), defined sampling as a reduction technique which allows a large dataset to be represented by a small subset of data.

There are different sampling methods, including random sampling, stratified sampling, systematic sampling, and cluster sampling. Random sampling….. (Briefly explain each of them and don't forget to cite). In this study, the researchers will use stratified sampling.

Stratified sampling allows the researcher to divide the population in defined classes (*citation*). The reason for using stratified sampling is because, the students can be divided into years of study which will ensure all students are properly represented (*this is the justification*). The sample will be calculated based on the formula…… (*Give the formula here*). Based on the sampling technique and sampling formula applied, table 3.1 shows the population and the sample size.

Table 3.1: Population and Sample size

Year of study	Population	Sample Size
One	500	210
Two	1000	670
Three	600	240
Four	400	135

3.3 Data Collection, Instrumentation and Procedure

According to (citation), data collecting is …… (*explain*). There are several data collection tools that can be used in research. This includes Questionnaire, Interviews

(structured, phone, group), Record Inspection, Observation, Prototyping and JAD.(*For each using one sentence, discuss*).

(Identify the best data collection method and justify your choice). This study will use a questionnaire to collect data. Questionnaire is the most appropriate data collection tool for this research because…..(*provide a justification*). (*Explain further how the tool will be used*). Two separate semi-structured questionnaires (one for lecturers and one for students) containing closed and open-ended items will be distributed by the researcher to the participants who in turn will fill in the questionnaire at convenient times due to their busy schedules. This will give respondents the opportunity to freely express their views on paper by either ticking the appropriate boxes and/or filling in the blank spaces provided for on the questionnaire. A follow-up to collect the completed questionnaires was done within an interval of one week making it possible to achieve a questionnaire return rate of 85.9% for industrial attachment coordinators, 86.7% for lecturers, and 87.8% for students respectively.

3.4 Development tools and Material

This study will use the MATLAB software package for image processing. It has powerful and easy-to-use features, especially when dealing with images. The algorithm will be developed using MATLAB (R2018a) on a PC equipped with an Intel Core i5 processor, 8GB RAM, and Windows 10 Operating System. The choice of MATLAB was based on its capability to carry out complex computations; it provided a visualization environment and easy to use mathematical notations *(citation)*. Further, it provides inbuilt toolboxes for image processing, computer vision, and batch processing which are appropriate for this research. A detailed explanation of the MATLAB software installation and activation will be attached as appendix… of this document.

Getting actual images to carry out this research will be highly difficult because of privacy and legal issues and technical hurdles. However, there are several datasets accessible on the internet that can be used. This datasets include: (*identify the data set and briefly (in one sentence each) explain each one of them and give citations*). In this research, the MIAS dataset will be used to study the effectiveness of the developed local descriptor as it was a benchmark dataset available online for research. The choice of the dataset was based on ……(*give your justification here*).

3.5 System Development Methodology

According to (*citation*), a system development methodology is….. (*Explain*). The existing system development methodology include… (*Explain each one of them*).The student registration system will be developed used RAD system methodology. RAD is a methodology that allows……. (*Briefly explain RAD*). RAD will be appropriate for this system because it will allow…… (*Justification for using RAD*).

3.6 Data Processing and Analysis

In this stage of research, the collected data will be processed and analyzed. The processing stage will include editing, coding, classification and tabulation of collected data that will be ready for analysis. The purpose of editing is that careful scrutiny of all collected questionnaires to produce completeness, error-free and readability (*citation).* (*Explain how editing will be done*) The purpose of coding is the assigning codes (numbers) for each category of answers (citation).In this research, the code No 1 will be used for the answer less than 25%, the code No 2 for the answer 26% up to 50% and so on…… : The purpose of classification is to divide the received questionnaires on the basis of their groups (*citation*). This study will split the received questionnaire into three groups including, group one (top management and executive managers), group two

(auditors and inspectors) and group three (experts). ……The purpose of tabulation is the process of summarizing data and displaying them in the appropriate tables that further analysis are to be facilitated (*explain how tabulation will be done*).

Once the data has been processed, it will be analyzed. Data analysis is the process of (*citation*). Data analysis enables the researcher to gain insight on the behavior of raw information (*Citation*). There are two statistical measures that a researcher can use to analyze data. i.e. descriptive and inferential statistic (*briefly explain each*)s. This study will analyze data using descriptive and inferential statistics. Descriptive statistic function such as standard deviation and frequency distribution will be used. Standard deviation will be used to determine on average the number of students who registered by the end of the first day of registration. While frequency distribution will be used to determine fourth year students who have registered within the first two days of registration process.

3.7 Ethical Consideration

The researchers understand that mammographic images contain sensitive information and that they are properties of the research community. The researcher will exercise utmost confidentiality by using the mammographic images purely for academic research. Since the mammographic images will anonymously be coded with no patient demographic information, they will not be traced back to individual patients. Also, an effort will be made to source for research permit approval from the National Commission of Science Technology and Innovation (NACOSTI) before proceeding with data collection.

3.8 Summary

Do a summary of the chapter content