

**A PREDICTIVE MODEL FOR QUALITY-ASSURANCE IN NORMAL LEARNING
THROUGH STUDENTS FEEDBACK IN KIAMBU TECHNICAL COLLEGES**

**BY: STEPHEN M. WAWERU
MASTER OF SCIENCE DATA ANALYTICS
KCA UNIVERSITY 2021**


SUPERVISOR: STEPHEN NJENGA

**DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE AWARD OF MSC DATA ANALYTICS IN THE
COLLEGE OF TECHNOLOGY AT KCA UNIVERSITY**

NOVEMBER 6TH, 2021

DECLARATION

This project is my own work and it has not been in part or fully been submitted or presented for the award of degree or any other academic work.

Signature 

Date ...6/11/21...

STEPHEN MAGU WAWERU

12/03157

This project has been submitted for examination with my approval as the appointed university supervisor

DR. STEPHEN NJENGA

Signature..... 

Date.....

ABSTRACT

Quality assurance (QA) in any technical college requires monitoring and evaluation for quality learning. Feedback from students when adopted can promote QA to ensure the learning environment is enhanced continuously and consistently. The adoption of the Competence-Based Education Training (CBET) in Kenya has come with new challenges of handling data. The change of the structure for learning delivery and assessment requires monitoring and evaluation to facilitate trainees, trainers, and managers to adapt to the requirements of this approach of training. Technical and Vocational Education Training Authority (TVETA) as the regulatory agency for technical colleges has standards for monitoring the successful and full implementation of competence-based training in Kenya. The County of Kiambu is among the counties in Kenya endowed with reliable communication infrastructure and should lead the way in leveraging on Internet, mobile telephony as well as the modern data repository and analytical technologies in the execution of processes in technical colleges. The predictive model for IQA from this study will be useful in adjustments of policies and methodologies, and facilitate instituting changes or affirming instructional roles. A sample population of technical colleges in Kiambu County was used to identify the current feedback mechanisms implemented in the colleges. The variables identified for the predictive model are Instructional delivery, resource provisioning, assessment process and technology usage. The phenomenon of participatory sensing is extended in this study where the student, the mobile phone, an application embedded in the phone and the internet are combined to form a sensing mechanism to support learning. Technologies employed in this model are VADER, spark, HDFS mobile telephony and the internet, while visualization graph for multivariate data is done by Parallel coordinates graph. This model and system derived thereof will support the digital survey student accustomed to use of mobile telephony for communication and cyberspace for seeking information, in order to interact in the learning ecosystem in the phenomenon of participatory sensing. The future of quality education is to leverage technology to help both the student and the instructor or trainer, to promote Kenya's industrial and national growth by the development of a skilled workforce.

Key Words: Semantic analysis, Feedback, Repository, Quality-Assurance

ACKNOWLEDGMENT

I would like to acknowledge my supervisor Dr. Stephen Njenga for guidance in realizing this Dissertation. In addition, I acknowledge the support and encouragement of my spouse and children in undertaking this research.

ACRONYMS AND ABBREVIATIONS

CBET	Competence-Based Education Training
IQA	Internal Quality Assurance Officer
SDG	Sustainable Development Goal
GDP	Gross Domestic Product
TVETA	Technical and Vocational Education Training Authority
TVET	Technical and Vocational Education Training
KBET	Knowledge-Based Education Training
UK	United Kingdom
UN	United Nations
AOD	Asynchronous Online Discussion Boards
NLP	Natural Language Processing
LDA	Latent Dirichlet Allocation
LSA	Latent Semantic Analysis
LSA	Latent Semantic Indexing
CAK	Communication Commission of Kenya
ICT	Information and Communication Technology
KDR	Kenyan Data Report
QA	Quality Assurance
QAS	Quality Assurance Systems
HDFS	Hadoop distributed file system.
HEI	Higher Education Institution.
STES	Student Teaching Evaluations System
TQAS	Teaching Quality Assurance System
CDACC	Curriculum Development, Assessment and Certification Council
KNQA	Kenya National Qualification Authority
TQM	Total Quality Management
NACCOSTI	National Commission for Science, Technology and Innovation
NLTK	Natural Language Toolkit (NLTK) python libraries and programs
SA	Sentimental analysis

GLOSSARY

Students' Feedback: These are opinions expressed on services received and facilities provided.

Competence: This is complex knowledge of how to act built on the effective mobilization of a variety of internal and external resources within related situations.

Internet of Everything: This networked connection of people, processes, data, and things.

Participatory sensing. This is sensing the environment by a combination of mobile devices and their owners

Purposive sampling: This a method for identification of participants of a research that does not use probability criteria but prior qualified knowledge and wisdom of the person undertaking a research to identify primary participants of a population leveraging information-rich cases for a phenomenon.

Culling: This removing redundant or irrelevant data and segmenting the data into meaningful and useful units

Sentimental analysis: Natural language processing to capture the attitude or tendency from textual or audial communication by a person through the contextual polarity.

Business Intelligence: Capture, storage, and analysis of internal and competitive information in the management of knowledge by organizations for robust decisions and informed actions in efficient and effective implementation of processes.

TABLE OF CONTENTS

DECLARATION	2
ABSTRACT	3
ACKNOWLEDGMENT	4
ACRONYMS AND ABBREVIATIONS	5
GLOSSARY	6
TABLE OF CONTENTS	7
LIST OF TABLES	9
LIST OF FIGURES	10
CHAPTER ONE	11
INTRODUCTION	11
1.1 Background of the study	11
1.2 Statement of the Problem.....	16
1.3 Main Objective.....	18
1.4 Specific Objectives	18
1.5 Hypothesis.....	18
1.6 Proposed Solution	18
1.7 Significance of the study.....	19
1.8 The motivation of the study	19
1.9 The Scope of the research	19
CHAPTER TWO	21
LITERATURE REVIEW	21
2.1 Introduction.....	21
2.2 Theoretical Review	21
2.3 Empirical Review.....	28
2.4 Conceptual Framework.....	30
2.5 Operationalization of variables	32
2.6 Summary	32
CHAPTER THREE	33
METHODOLOGY	33
3.1 Introduction.....	33
3.2 Research Design.....	33
3.2.1 Research strategy	36
3.3 Study Population.....	39

3.4 Sample and Sampling Procedure	40
3.5 Research Instrument.....	41
3.6 Instruments' Validity and Reliability.....	42
3.7 Data Collection Procedure	42
3.7.1 Legal and Ethical Consideration	43
3.8 Data Processing and Analysis	43
CHAPTER FOUR.....	44
DATA ANALYSIS FINDINGS AND CONCLUSIONS.....	44
4.1 Introduction.....	44
4.2 Response Rate.....	44
4.3 Description of Participants.....	46
4.4 Research Findings	46
CHAPTER FIVE	56
CONCLUSIONS AND RECOMMENDATIONS.....	56
5.1 Introduction.....	56
5.2 Conclusions.....	56
5.3 Contribution to the study.....	56
5.4 Recommendations for future research	57
REFERENCES.....	59
APPENDIX 1.....	67
APPENDIX 2.....	93
Research Questionnaire.....	93
APPENDIX 3.....	95
NACCOSTI Research Permit.....	95
APPENDIX 5.....	97
Research Schedule	97

LIST OF TABLES

Table 1: Tools Performance on Students' Feedback	29
Table 2: Preselected Variables	32
Table 3: Sample size.....	41
Table 4: Response Rates.	45
Table 5: Survey Respondents from Kiambu County	46
Table 6: Positivity Rating.	47
Table 7: Website and Mobile application usage.....	47
Table 8: Internet usage and Instructional delivery rating	48
Table 9: Resource Awareness and Technology Requirement	48
Table 10: Feedback capture frequency.....	49
Table 11: Model Data Files.....	49
Table 12: VADER and TextBlob Scores by column content	51
Table 13: Scores on Linear scale Importance measurement.	51
Table 14: Scores on Linear scale Efficiency measurement.....	51
Table 15: Scores on Likert scale Agree measurement.....	52

LIST OF FIGURES

Figure 1: The structure of Quality of Assurance.....	26
Figure 2: Conceptual Framework.....	30
Figure 3: Milestones and Deliverables.....	34
Figure 4: Text Analytic Model.....	38
Figure 5: Topics and sentiments.....	38
Figure 6: Opinion Mining Architecture	39
Figure 7: Students' Feedback Analytic pipeline.....	39
Figure 8: Visualization of independent variables	50
Figure 9: Positivity Ratings for Independent Variables.....	50
Figure 10: Visualization of Model Output.....	53

CHAPTER ONE

INTRODUCTION

1.1 Background of the study

Skilled personnel will facilitate attainment of Sustainable Development according to Global 2030 agenda (UN, 2015), as captured in SDG 9 target 9.2, by increasing the proportion of industry's contribution to Gross Domestic Product (GDP) and Employment. Kenya through TVETA is transforming education from Knowledge-Based Education Training (KBET) to Competence-Based Training (CBET) to meet her industry and business needs in line with SDG 4, target 4.7 that addresses among other items, curricula and education policies (UN, 2015). Coordination of all TVET activities in Kenya are done by TVETA, while accreditation, quality assurance, curriculum, assessment, certification, and resourcing are executed by its council, the Curriculum Development Assessment and Certification Council (CDACC). On the other hand the national qualifications framework is established and regulated by Kenya National Qualification Authority (KNQA), as echoed by Akala, Winston, Changilwa and Peter (2018).

Competence refers to complex knowledge of how to act built on the effective mobilization of a variety of internal and external resources within related situations, Tardif(2006) as cited by (Cote, 2016). Attitude is an attribute of competency just as knowledge and skills, and hence should be assessed. In Competence Based Education Training, the learning process is more practice-oriented and the formative assessment is dependent on each student's readiness, making coordination of a class for learning and assessment a challenge. The KBET emphasizes on knowledge impartation although the skills and attitude are also taught and assessed. While Skills is about the performance of some tasks, professional attitude is transformation into a work culture or work style according to Bakarman, (2005). This includes ability to work and fit in a team, motivation, concentration, due diligence, focus on quality, technical approach in executing tasks, teamwork, responsibility of outcomes, time management etc. Attitude is depicted by behavior and assessment may involve observance of the students on performance on aspects of resources utilization, safe work habits, teamwork, problem-solving, service quality etc.

With the paradigm shift of the training approach from KBET to CBET, students' feedback can be an integral part of a college self-evaluation or internal quality assurance mechanism to capture the challenges and concerns as both the students and the lecturer adapt to the new approach. The students in this environment are bound to have concerns about many aspects of learning which might include contact hours, lesson delivery, trainer performance,

etc. The instructor is also adopting or in the process of embracing the paradigm shift in training methodology and therefore a measure to which the CBET approach is being realized should be captured through students' feedback on instruction. But even as the student responds to the instructional process, extreme cases or main areas of concern will be captured by themes revealed through topic modeling and semantic analysis. In the digitization of various aspects of the learning environment, the feedback mechanism should not be left behind. This requires information repositories and processing mechanisms for reporting and decision-making purposes. In addressing the digitization of libraries, Amollo (2011) referred to an information repository as a specialized kind of a digital library, for the management and easy availability of all digital materials in an organization.

In meeting the objectives as stipulated in the TVETA Quality policy, (Ministry of education, 2021) adoption of new technologies and their attendant infrastructure is an imperative. This should be accompanied by skilled personnel that promotes the use of these technologies, and therefore a paradigm shift on how data and information is used in learning institution should be accelerated. This has come in the wake of the fast adoption of Mobile telephony and networking technologies for communication, financial transaction, and learning among other processes in Kenya. With the shift in training approach and the technologies in our environment, there is a need to adjust to match these new developments with the learning requirements. Technical Training institutes that are transitioning to CBET should take cognizant of data requirements as dictated by these factors and leverage the new data capture, storage, and analysis technologies in creating a conducive environment for the learner.

Students' Feedback is opinions expressed on services received and facilities provided, (Harvey, 2003). These reflect their satisfaction with the teaching and learning process, and on the provision of the requisite facilities for learning. A feedback study on American higher education revealed instructional inadequacy, requiring improvement according to Boyer (1987, as cited by Brinko, 1993). A proposal to use feedback for improvement of teacher performance was made with a rider, 'delivery communication in the right way'. Students being the source of feedback was accorded higher ranking on the attributes of validity and reliability, and many instances of feedback are required to change an individual's behavior and self-perceptions according to Hunsacker, (1983, as cited by Brinko, 1993). Feedback is more effective when it focuses on behavior rather than the person, and when it is descriptive rather than evaluative, as well as when it is continuous and considered as a process according to according to Batista (1976, as cited by Brinko, 1993). Feedback serves for improvement on various aspects in the training environment. Learner-centered feedback, Johnson 2009, (as cited by Sosibo, 2010)

should cover goals, activities, assessments, and outcomes in learning. Provision of information on the status of the learning environment can be made effectively by students themselves (Walber, 1976), acting as mediators of the learning process in improvement of their learning environment. When technologies are used, the students apportion more attention and time to learning (Angelo, 2018). Application of technology becomes instrumental in provision of continuous students' feedback for frequent monitoring of processes.

Models for student's feedback are varied, and Sosibo (2010) addressed the views on feedback for course evaluation and curriculum improvement and the challenges encountered in course evaluations. Predictive models proposed by Srungaram, (2017) focused on the approaches in teaching, the trainer's performance, and (Sorour et al, 2014) the online environment. Evaluation Methods as applied by seven Southern African Universities included centralized and standardized forms, (Sosibo, 2010). There are divers ways of collecting Student comments that include focus groups, suggestion boxes, questionnaires (Harvey, 2003), website forms, online applications consultative or representative meetings, etc. The hardcopy questionnaires used for feedback can only capture a small spectrum of the learning environment and are mostly employed in appraising lecturers or instructors. Suggestion boxes on the other hand are mostly deployed for external players that interact with colleges like suppliers and vendors. Williams et al, (2013) opined that open-ended questionnaires could serve the purpose for certain education programs but analyses of responses, in this case are immense even with relatively modest samples unless each question is encoded straightforwardly.

Some Institutions spend a lot of time in physical meetings, brainstorming to address learning processes and students' concerns. In Technical colleges in Kenya, boycotting and threats to boycott internal examination has been occasioned perhaps by insufficient or lack of information on various aspects of learning. Detecting growing discontent among student on aspects of learning is key to smooth learning of Institutions. In such cases, with either wrong or justified perceptions on aspects like processes, events or resources, the students' comments if analyzed can form the basis for timely and appropriate pre-emptive actions through communication and resource provisioning. As students' comment on either the appropriate or unacceptable behavior of a trainer, as well as their liking or dislike of the behavior, this may prompt investigation on performance, competence in delivery or redistribution of instructional roles. Where performance in some programs is compromised, analysis of students' feedback may prompt further investigation to reveal the cause, which could be caused by workload among other factors. In cases where students' enrolments in an institution are increased, any imbalance in processes and constituents of the learning ecosystem can be identified as students'

comments on their experiences in order to identify the areas affected adversely. In these case, students feedback should be descriptive, captured continuously, focus on behavior, and implemented as a process Batista (1976, as cited by Brinko, 1993) for it more relevant and effective. Additionally, feedback from students can serve in identifying possible failures and inadequacies in processes to generate appropriate exploratory or confirmatory action in a variety of facets of learning that might include instructional role assignments, resource provisioning and assessments and other training processes.

The flexibility of online platforms provides the ease of restructuring as needs arise and can allow initiating discussions to solicit feedback on areas of concern. Lack of consistency in evaluation, gaps in data analysis, and inadequate skills in interpretation led to misinterpretation and inappropriate application of the feedback outcomes, Sosibo (2010). Recommendations were done to have a well-coordinated evaluation system for determination of the effectiveness of that evaluation system. Some evaluation models emphasize teacher-student relationship to make the students feel included. The new quality assurance requirements have evolved from the adoption of CBET needs as advocated by TVETA, and the behavior of the modern student accustomed to networking and mobile telephony technology's needs. It is imperative to consider new strategies for managing the learning environment by fully employing appropriate data capture technologies and storage systems commensurate with the data requirements. Implementation of CBET is meant to meet the quality and standards as dictated by the industry and business needs in Kenya according to Amina (Ministry of Education, 2018). This arose from the inability of the current knowledge-oriented system to provide a skilled workforce hence the need of the flexible, outcome-based, industrial and business-centered system (Ministry of Education, 2018) to impart skills appropriate for the labor market. The feedback mechanism should monitor the learning environment to ensure meeting the requirements of the new curriculum. The appropriate feedback presented in the right way will therefore be useful to the individual instructor or trainer and to the managers in the overall status of the learning environment.

There's rapid growth in acceptance and use of the Internet of Everything (IoE). This is vital in enhancing educational and academic activities concerning learning processes and information management. The penetration of mobile telephony and internet technology, according to Kemp (2021) has attracted 22.86 million internet users (43% of the population of Kenya population, with 97 % is accessing the internet via mobile phone and average time of usage of internet of 4 hours 36 minutes. This depicts a population adopting these two technologies in every facet of their lives. In this regard, the students at universities and tertiary

colleges use these devices, most of them at age of 18 to 34 years and easily reachable online according to Kemp (2021) on the Kenyan Data report. With this growing level of exposure and usage of these technologies, the old forms of relaying data and information in hardcopy form are being rendered irrelevant and inapplicable in various aspects of college activities and procedures.

Mobile telephony and Network technologies are being deployed in diverse ways in many facets of our lives. The Mobile devices in conjunction with the owners of the devices are being exploited as sensors of environments in a new phenomenon referred to as participatory sensing (Burkey, 2006). In the learning environment, the areas of interest can be captured by getting students' sentiments through text messages on various aspects of learning including industrial tours and attachments, learning assessments, learning-teaching process. These can promote agenda generation and decision-making in colleges. Changes in Policy and eventual formulation can then be used not only on the financial capacity but based on the evolving needs of the trainees to ensure provision of the right environment for learning. The students, each with a Mobile phone can form a college participatory sensor network which underpins the role of mobile telephony and networking technologies in technical colleges today. The capture of feedback from students can be done by a Mobile application tailored for a college and sent to the students' phone upon registration for installation. These can be a stand-alone application or as a collection of applications for various tasks including assessment feedback, assignment submission, etc.

Quality assurance according to Anderson et al. (2000 as cited by, Kelum,2020), is confirmatory means to meet standards set by a learning institutions itself or other regulating agencies, to maintain and enhance the quality of learning and teaching. Kelum (2020) argues that the factors that influence QA systems in many countries are their systems in higher education, organizational culture, social and traditional backgrounds. Therefore, the level of development of education system in Kenya, the KBET and CBET curricula, the students' culture in communication and information seeking, and the level of adoption of technologies should inform how Quality Assurance Systems (QAS) are designed and developed to much the prevailing local needs. The needs of various curricula and training delivery approaches have their unique requirements and therefore considerations of policy changes and implementation methods should start in Colleges as Kenya seeks to transition to CBET. IQA is meant to review service delivery in a learning institution, assess adherence to standards as well as maintenance and enhancement of standards (Daguang, Zuoxu, Fan, & Yanjie, 2017). The Trainee or student should play a pivotal role in ensuring a good environment that match the rigors in practice and

assessments in CBET, by feedback channeled, analyzed and stored by application of appropriate technologies. The focus of these undertaking should address the trainee positive interests in order to meet both student satisfaction and the requirements for the new curriculum thereby uplifting the requisite Skills level as dictated by industries and Businesses. In Kenya TVET institutions, some of pedagogical issues CBET include weak QA mechanisms and insufficient skilled trainers which results in ineffectiveness in curriculum delivery according to Akala et al (2018). The modern data and communication technologies may fill the gap of this inadequacy to ensure there is a focus on the overall monitoring and management of the learning environment. There is a need to restructure the feedback mechanism by soliciting more information which can be achieved by allowing trainees to freely provide information on their environment and from their perspective. The requirements for quality assurance in Australia and UK governments include active involvement of students through feedback, which is analyzed and then results published, according to Shah and Nair 2012, (as cited by Shah, 2017). Quality assurance involves putting the right mechanism in place and consistently informing on implementations. This promotes provision of service or product to the acceptable level of quality and fulfills the expectations of the customer. Standards specified should be met consistently for a product or service (Ellis, 1995). In the context of technical colleges, the learning needs should be met consistently and where inadequacies exist, collective measures should be undertaken promptly and communicated to students. The status of the learning environment can best be communicated and advocated for by the learner through appropriate feedback mechanisms for channeling their opinions, observations and concerns. This promotes the student as a major constituents of the learning ecosystems to participation in its management.

Development of a predictive model and provision of necessary infrastructure for student capture and analysis of feedback is an imperative to support the student and promote quality assurance for the CBET learning environment.

1.2 Statement of the Problem

Internal Quality Assurance is meant to review service delivery in learning institution to assess adherence to standards and maintenance and enhancement (Daguang, Zuoxu, Fan, & Yanjie, 2017).

Exclusion of the main stakeholder or internal customer (the student) in Many QA and IQA models, and giving prominence to the institution experience instead of the students' learning experience and satisfaction, (Douglas, McClelland, and Davies (2007), Casotti et al,

(2013) and Glaser, (2006, as cited by, Massey et al, 2019) may fail to capture important aspects or details that can only be communicated effectively by student.

There is complexity in some models by having several standards and capturing many variables as suggested in the study by Daromes and Suwandi (2015) and Kelum, Pradeep, Najdanovic-Visak and Gunawardhana (2020). Some models employ intermittent or periodic data capture usually at the end of a semester instead of daily and Continuous data capture. These require streamlining for appropriate analysis and decision making.

Timely and updated information lacks with some models like an Australian university capturing data at the end of semester (Nair, 2011), which may fail to detect growing discontent among student on aspects of learning early enough to deter adverse effects on the learning processes.

The modern student is digital savvy and accustomed to technology for communication and information seeking, hence cannot effectively participate in moderating the environment using traditional methods of questionnaires at the end of Semester. This was suggested by the study by James et al (2021) indicating students' engagement on mobile platforms in Laikipia campus with WhatsApp users (93.7%), Facebook (67.4%), among other social media platforms, spending at least two hours on daily basis. Many models don't employ the modern data capture, analysis and technologies for efficient monitoring of QA.

Data analysis at the end of a module or semester, leaves several aspects unaddressed for a considerable time e.g., a month or three months. These still impact on the learning processes and student's learning experience even if they will be addressed. Analysis of textual data from students' feedback require Natural language processing by machine learning, Lexicon and rule-based algorithms like VADER to get insights from the data. A modern system should cover daily and time series feedback for analyses of trends as well as daily instances.

Pedagogical issues in Kenya like weak QA mechanisms and insufficient skilled trainers as reported by Akala et al (2018) should be monitored, identified and addresses early enough to maintain effectiveness in training delivery.

The Trainee or student should play a pivotal role in ensuring a good environment that match the rigors in practice and assessments in CBET, by feedback channeled, analyzed and stored using appropriate technologies.

Some Trainers will take time to adopt new ways and adapt to training delivery methods in transformation from KBET to CBET. The human and physical resource requirements as well as the rigors in skill training and assessments for CBET call for IQA mechanisms commensurate with challenges and needs thereof.

The predictive models that exist focuses on application of technology online environments, feedback to students for self -improvement and evaluation of training programmes, and some using traditional methods of data capture as others adapt the websites for feedback. The normal learning should not be excluded in leveraging on technology to facilitate learning.

1.3 Main Objective

This research is geared toward developing a Predictive Model to facilitate Quality Assurance in Competence-Based Education Training in Kiambu County through students' feedback on the learning environment.

1.4 Specific Objectives

- To identify the quality assurance features under which students' feedback will be classified from data collected.
- To develop a Predictive model for quality assurance through students' feedback.
- To evaluate and validate the Predictive Model for CBET.

1.5 Hypothesis

Deployment of data capture, repository, and analysis technologies for students' feedback on the learning environment will enhance internal quality assurance in competence-based education training.

1.6 Proposed Solution

The study seeks to get the trainers perspective of students feedback, Internal quality assurance and deployment of modern technology for IQA. This will culminate in the development of a Predictive Model for internal quality assurance based on students' feedback, to support implementation of CBET in technical colleges in Kiambu County.

Appropriate evaluation of needs and challenges in the overall learning environment require adoption deployment of Mobile telephony, the Internet, Modern repository system Like Hadoop Distributed File System (HDFS) and fast processes techniques like Valence Aware Dictionary (VADER) should be availed to allow comments to match the needs of the digital savvy student, support inclusion in quality assurance and eventually to promote the learning experience and satisfaction of students. This will help in avoiding misinterpretation and inappropriate application of feedback outcomes, Sosibo (2010). Daily capture of comments from students on rigors aspects of learning form the basis of the overall performance rating of a learning institute. This time series qualitative data can be used to both discover patterns and project outcomes when status quo remains. This will facilitate interrogations of major processes

to detect areas that need improvement as well identification of strengths where there are notable levels on improvements

Exploration of feedback mechanism, data storage, and analysis requirements will enable the formulation of an appropriate model for data capture, storage, and analysis for the learning environment to meet the requirements of CBET training.

1.7 Significance of the study

This study will support the running of technical colleges in Kiambu County by enhancement of internal quality assurance mechanism in TVET member institutions. The learners in this county will have an improved learning experience and satisfaction which will in turn facilitating skill acquisition. The outcomes of the study can be adopted across the Kenyan counties for enhanced training. The model for internal quality assurance will form the basis of enhanced external quality assurance in meeting the TVET National standards in Kenya. If replicated on the global perspective, it will promote realization of the global agenda 2030 through skilled workforce for acceleration of industrial growth.

1.8 The motivation of the study

As a qualified instructor for both KBET and CBET, and working at the Kenya Institute of Surveying and Mapping, I have had the opportunity to interact with the learning environment, observed some behavior trends in students, and noted some inadequacies in learning processes. In transitioning to CBET, it is imperative to have modern systems for a repository, analysis, and feedback, based on these preliminary observations to facilitate full implementation of CBET in TVET member institutions.

1.9 The Scope of the research

The research extends to feedback requirements for capture, storage, and analysis in Technical colleges in Kiambu County. These will facilitate the formulation of a model for students' feedback to promote quality learning in the CBET approach in TVET member institutions in this region. The study will cover a sample population at least 20% of 157 technical colleges in Kiambu County. The data will be covered through interviews and an online questionnaire. Evaluating acceptability, the role, and extent to which descriptive and continuous feedback has been employed to inform learning practices and processes in these tertiary learning institutions is key to the development of a data model that will support full implementation and transitioning from KBET to CBET, which requires closer monitoring with more rigor on practice and assessments. The External environment of learning institutions in which the tertiary institutions are embedded is replete with communicating and networking devices in use to support various aspects of daily lives for the populace. The students being part of this larger

ecosystem are accustomed to the use of mobile phones and networking technologies and have a new culture of communication. This study seeks to uncover the appropriate mechanism for the deployment of these technologies to allow students to effectively respond as actors in their learning environment to moderate it in a way that promotes quality learning. This study will be restricted to students' feedback requirements for model formulation to support quality assurance but will not cover the degree to which this feedback impacts quality assurance, which would require more time and resources.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

Theories and mechanisms on students' feedback have been explored in this chapter. Applications of these theories and approaches in learning institutions are highlighted in the empirical review. In addition, the changing trend of capture and application of feedback capture as well as students' communication behavior is underscored. Variables and their relationships are illustrated in the conceptual framework, their implementation depicted in the operationalization of variables section, and a summary made on the aspects reviewed.

2.2 Theoretical Review

Feedback from students is valid and reliable and promotes the student's learning and experience, (Huxham et al, 2008). Feedback is most suitable when it's descriptive, impartial and geared to provide quality information on work quality, (Hamid & Mahmood, 2010). Insights mined from comments are useful in improving product and services, according to, Negi and Buitelaar (2015, as cited by Gottipati et al, 2018). Student comments can be captured in various ways that include focus groups, suggestion boxes, questionnaires (Harvey, 2003), website forms, online applications consultative or representative meetings, etc. The hardcopy questionnaires used for feedback can only capture a small spectrum of the learning environment and are mostly employed in appraising lecturers or instructors. Open-ended questionnaires apply in certain education programs, but analyses of responses are immense, (Williams et al, 2013).

Various Feedback Mechanisms and data capture tools are employed in learning Institutions. When H-forms are used, the group members hold discussions, record the outcome of vote, and individually register their rating a in a specific section of the form according to Guy and Inglis (1999, as cited by Huxham et al, 2008). When this data capture tool is applied in the classroom environment, the students undertaking a particular department in a college learning session participate in monitoring and evaluation of their learning environment by assessing various aspects of a unit of learning. In a Focus group, a leader for a module is appointed, an interview of about twenty minutes is carried out and responses on improvement are recorded, (Huxham et al, 2008). The disadvantage of this method is the lack of anonymity which might hinder capturing of responses from students for fear of reprisal in case the instructors and managers get to know the responder. The questionnaire is the most popular tool for feedback and was compared with other tools by Huxham et al (2008). To overcome the

disadvantage of hardcopy questionnaires, online textual interaction for feedback structured to provide anonymity, can support feedback irrespective of geospatial location and time, especially if an application for feedback is embedded in a mobile phone.

Participatory learning is supported by the Web and the internet through video sharing and discussion boards as collaboration and communication tools, according to Brown and Adler (2008, as cited by Reigeluth et al, 2014). Asynchronous online discussion boards, (AOD) gives flexibility, ease of access, and satisfaction according to Casotti et al, (2013) and Glaser, (2006, as cited by, Massey et al, 2019). Feedback boards structured in the format of asynchronous discussion boards can allow students to communicate anytime. In Technical colleges, Managers and the student's representatives can start up discussions by presenting a topic through Chat and discussions boards to promote interaction with students. The managers can then respond online where the need arises or invite the students' representatives for a meeting based on topics or themes identified through text mining and semantic analysis. As each student responds to the environment and gives feedback, the interactivity attribute is not there at the individual level in the phenomenon of participatory sensing, (Burkey,2006) but this feedback from the global populace of students is instrumental to monitoring, adjustment and maintaining that learning environment. Sensing then in the context of learning is a combination of the student's sentiments, data capture by the mobile telephone device, relaying of the textual message to the repository, processing and display by machine learning technologies. This will reveal the topics of interest and positivity on themes.

The modern-day student is digitally savvy, accustomed to online space and mobile telephony for communication and information, preferring mobile telephony and cyberspace to traditional methods, according to Williams et al, 2011(as cited by Conefrey, 2016). Technical colleges should leverage this new behavior and skills as well as the technologies to get data for quality assurance. Engagement of students in the learning environment through participatory learning (Shen (2014) and participatory sensing (Burkey, 2006) will transform education by enhancing internal quality assurance mechanisms for instructors, students, and managers in the promotion of competence in learning. Zinzou¹ and Rubang (2020) emphasized on exploiting the internet for virtual repository and adoption of modern techniques in application of knowledge for quality assurance in education.

Data processing in the learning context refers to natural language processing (NLP) of textual messages relayed by students in response to events, actions, or processes that are at interplay in the learning ecosystem. Dringus and Ellis (2005) underscored the strategy of text mining in an online learning environment which can be replicated to support physical learning

through engagement when in the precinct of college and when away for feedback purposes. Sahu (2008) noted that text mining was appropriate for evaluating dashboard information objectively and in detail. Revealing topics through analysis of words can be executed by unsupervised machine learning algorithms or statistical methods like Latent Dirichlet Allocation (LDA) to classify sentiments (Negara, 2019). Text Mining and Sentimental analysis when deployed in a technical college or satellite campuses of a college, can give an early pointer on arising needs of the students' populace. Valence Aware Dictionary for sentiment Reasoning (VADER) is a package within the Natural Language Toolkit (NLTK) libraries. It is a tool that uses dictionary features and five heuristics to discover sentiments in text messages. In VADER, the words have been assigned positive or negative labels and the intensity of polarity based on their meaning and five heuristics. The heuristics are; Punctuation, Capitalization, Degree modifiers, Shift of polarity due to Conjunctions, and Negation of Polarity by three words abutting each other, preceding a feature that is heavily laden with sentiments. VADER is suitable in this case since it does not require labelled data for training, and performed better in most instances than other analysis tools like Linguistic Inquiry and Word Count(LIWC), Support Vector Machines (SVMs) according to a comparative study by Hutto and Gibert, (2014). The early sensing of sentiments by LDA and VADER can spur up prompt interventions and actions to promote the smooth running of institutions and to facilitate learning in technical colleges.

Quality Assurance (QA) refers to fitness for purpose and transformation according to Lomas (2001, as cited by Kis, 2005). IQA encompasses processes developed by institutions to monitor and enhance the quality of the learning experience of students, and for realization of stipulated goals. In CBET context Internal Quality Assurance (IQA), is about having the right policies and implementing good practices for improvement, and continuous monitoring of the quality of the education services, (Jingura & Kamusoko, 2019). Internal quality Assurance requires having standards, proper planning, implementing procedures, as well as evaluation of activities geared to drive, and empower those offering services. Positive interrelations of students and teachers involves sufficient address of questions from students, clear outlining of syllabus content and expected outcomes of learning, students participation frequency in lessons' activities (Taurina, 2015). These are among the areas for comments that the student may address in describing the status of the learning environment, to provide the basis for monitoring and management of processes and resources.

In Australia and UK governments, Quality assurance strategies requires students active involvement in external quality assurance through feedback, which are analyzed, and the

results published, according to Shah and Nair 2012, (as cited by shah, 2017). Development of a predictive model and provision of necessary infrastructure for student capture and analysis of feedback is an imperative to support the student, and promote quality assurance for the CBET learning environment. Shuiyun, (2016 as cited by Kelum et al, 2020) argues that it is difficult to evaluate the effect of QA on actual the experience of the learner since students' opinions are excluded, as is the case of learning institutions in China, where evaluation include aspects like mission, staff, learning equipment, learning programs, learning environment and outcomes etc. In many institutions in UK, quality assurance process includes audit of the learning Institutions and inspection of individual departments where Methodologies employed are teaching observation, and interacting with trainers teams for courses, external examining bodies, the learners, previous graduates, the potential employers as well as data collection and analyses of institutional data (Kelum, Pradeep, Najdanovic-Visak and Gunawardhana (2020). The outcomes of the process are evaluations and reports for both internal and external entities according to Filippakou & Tapper (2007, as cited by Kelum 2020). QA serves for evaluation of infrastructure, programs, staff, and internal mechanisms, functions and procedures for overall institutional performance as indicated in **Figure 1**.

Pond, (2002), recommended that student experience should be the basis of evaluation of educational programs, QA processes should focus on the extent of content learning but not its delivery, lecturer interaction with students but not contact hours, the relevance and access to resources, and infrastructure evaluation to capture the extent to which it serves students.\

QA system of universities included Internal audits on quality, with reports preparation involving personnel from various levels and shortcomings from suggestions and recommendations implemented according to Coomaraswamy, Jayatilleke and Abeygunawardena (2014 as cited by, Kelum et al,2020). Every worker in the leaning institution is responsible and accountable in ensuring Internal Quality Assurance as a continuous process (Kelum et al, 2020). At the close of a semester in an Australian university, there is participation of students in quality assurance through paper-based and online surveys administered, which starts by notification of students undertaking a unit via email, after which they access and evaluate each unit electronically through evaluation questionnaire on the portal (Nair, 2011). Teaching evaluation system based on students feedback as deployed in American and European in Higher education Institutions (HEI) have been transformed to effective and standardized Teaching Quality Assurance System (TQAS) according to Dunrong and Meng Fan (2009).

In Kenya, a study by James et al, (2021) revealed usage of mobile platforms by students in Laikipia campus with WhatsApp users (93.7%), Facebook (67.4%). With increase in mobile

telephony and internet technology penetration (Ngugi, 2014), most students in Kenyan universities and technical colleges being young and of 18 to 34 years are easily reachable online according to Kemp (2021) on the Kenyan Data report. Most of Students (83.6%) have messaging applications installed on their phones, 97 % accessing the internet via mobile phone and accessing internet on daily average time 4 hours 36 minutes. This reflects a changes in skills expectations and a paradigm shift of preference of the structure and form students' feedback mechanisms appropriate in colleges in Kenya.

Readiness in adopting technology is a key factor in implementation of an Internal Quality Assurance System, and this ensures that changes in graduate's specification are reflected in learning process, learning technology and curriculum, (Yulherniwati and Ikhsan (2020).

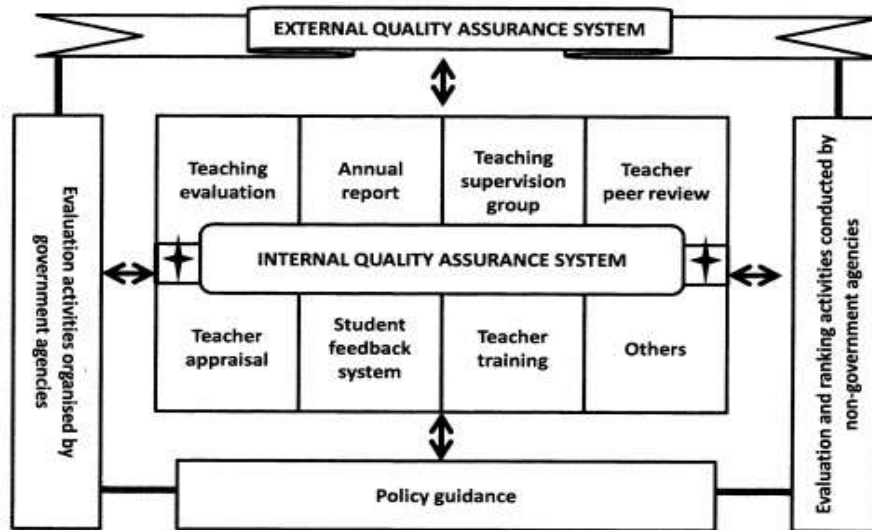
2.2.1 Quality Assurance Models

Programs in education institutions are taken through a systematic and planned review for conformance to standards, and enhancement and maintenance of infrastructure, and mechanisms and procedures for effective quality control. According to Tamaro, (2005), the focus of QA in Library and Information Science (LIS) Schools was 76% on resources, 64% on curriculum design , 52% on outcomes, 58% on student evaluation, with three models of quality assurance emerging based on; program orientation, educational process orientation, and learning outcomes orientation according to Knox (2001 as cited by Tamaro, 2005). There is a paradigm shift of measuring learning by laying emphasis on input like contact hours in teaching hours, to output based methodologies using learning outcomes and competencies (Tamaro, 2005). QA serves for evaluation of infrastructure, programs, staff, and internal mechanisms, functions and procedures for overall institutional performance as indicated in Figure 1. Evaluation of educational programs should be based on student learning experience but not institution experience, and QA processes should focus on; the extent of content learning but not its delivery, the relevance and access to resources, lecturer interaction with students but not contact hours, and the extent to which infrastructure serves students (Pond, 2002).

The QA Model based on learning outcomes focusses on performance and perceptions of students' as elements of QA. Some QA models in education based on industrial standards ISO 9000, considers students as a product or consumer ensures that the student is attended and satisfied in achieve quality education and training (Tamaro, 2005). Others QA models based on Total Quality Management (TQM) lay emphasis on self-evaluation and institutional

enhancement. In education, the coverage should be broader and cover more than processes, product or service approaches of the industrial model, according to (Tamarro, 2005).

FIGURE 1.
The structure of Quality Assurance



Author: Tamarro, (2005)

Standard, structure and Models for student's feedback are diverse as applied for various facets of learning environment. Centralized and standardized forms are among the evaluation methods used by seven Southern African Universities, Sosibo (2010). Gottipati, Shankararaman, & Lin, (2018) focused on qualitative feedback from students' comments using text analytics to identify the gaps in teaching, instructor's performance and content of the course. Hamid and Mahmood, (2010) suggested that Teachers need better quality information about student learning, emphasized the timeliness, and the nature of information provided for teachers self-regulation.

In a constructivist model of feedback the learner is a viewer and an active player able to construct own understanding, and act on the feedback received for self-regulation and improvement on subsequent tasks according to Molloy and Boud (2014). The focus on this model is the learner not the instructor, but feedback is more effective when it focuses on behavior rather than the person, when it is descriptive rather than evaluative, and when continuous and considered as a process Batista, 1976, Hunsacker, 1983 (as cited by Brinko, 1993). Students are most suitable to provide the status of the learning environment (Walber, 1976), as they act as mediators thereby improve their learning environment. This should be supported by technologies for them to focus on learning (Angelo, 2018), especially when they

are supposed to freely give feedback on various aspects of learning without restriction of geospatial location and time.

In TELL Model the features include commenting on observed specific behavior, explanation of impact of observed actions, feedback on corrective or enhancement measures, and commenting on the impact on of failure to take action, (Hamid & Mahmood, 2010). Although the first role of students' feedback is to support the individual lecturer or instructor, the overall objective is enhancement of the quality of the learning environment for greater results and for IQA in technical colleges. Therefore, students should be allowed to give comments freely and feedback should not be restricted to selected areas but should cover all the facets of learning environment. The data collected from the participants of this study will help in classifying the main facets under which all other aspects of the environment will be captured.

The key areas for IQA models in European Network for Quality Assurance in Higher Education (ENQA), (2009 as cited by Kelli et al, 2020) include programs reviews, students assessments, assurance of quality with regard to teachers, resources for learning and support mechanisms for students, information systems etc. Ayodele (2007 as cited by, Ofojebe and Ezugoh, 2010), proposed a model that include the quality of teaching personnel, availability of quality instructional teaching materials, quality institutional infrastructure, and learning delivery as factors in Quality assurance.

Predictive models are used to make inferences on future events like measurement of acquisition of skills, the effect instructional styles or performance of a trainer according to Srungaram, (2017). This is unlike explanatory models which gives explanations of outcome. The predictive model in this study gives the rating of services and a pattern can be projected to show the trade based on daily evaluations occurrences or instances of comments. This should indicate consistency on individual variables and the overall performance, and indicate stagnation or direction on either side.

K-means clustering and LSA methods (Praveen and Prasanna, 2021) were used in prediction of students performance according to Sorour, Mine, Godan and Mine (201) , using students' comments at the end of the lesson to understand their attitudes. Satisfaction of students is valued as integral measure of quality in experiences in learning according to Parahoo et.al (2015, as cited by Kuo, Walker, Bell and Schroder, 2013). This was opined by Parahoo et al (2015) on development of a model to measure students' satisfaction by the preceding contributory factors in online learning environment. This model environment had different factors since there is no face-to- face iteration with the lecturer or trainer and the

virtual space has its own unique factors like connectivity, interactivity, synchronicity, audibility, devices and presentation software etc.

There is need for greater involvement of students in the management of quality assurance in tertiary colleges since they are providing valuable feedback on instructional effectiveness and courses, entrenchment of the institution's quality culture, as well as inculcating sense of responsibility into students (Okpa, Odigwe, Emeribe and Obule, 2020), leading to improvement of their learning experiences and overall development of the educational process (Archibong, 2013).

2.3 Empirical Review

Students' Feedback is difficult but important in the improvement of their environment according to Higher Education Funding Council for England, 2011(as cited by Bashir et al, 2016). There is a positive impact from students' feedback as they participate in modifying the learning processes through monitoring, according to Eraut (2006, as cited by Bashir et al, 2016). There is a need to adopt Modern technologies, (Bashil et al, 2016) in capture and processing students' feedback to match the learning environment replete with digital technologies employed in daily communication and transactions. Various Surveys reveal that feedback has an impact on the improvement of the learning process as is the case of UK (2011), and Australia addressed by Ferguson, (2011, as cited by Bashil et al, 2016).

Kenya is one of the leading countries in the adoption of mobile services according to the communication commission of Kenya, CAK (2016, as cited by Ojino, 2018), and has the potential of being a global leader in ICT and ahead of other countries in the African continent in use of the internet. A rapid increase in subscriptions in mobile telephony was reported in early 2012 with 30.4 million, Ngugi (2014). The penetration of mobile telephony and internet technology, according to Kemp (2021) has attracted 22.86 million internet users (43% of the population of Kenya population), with 97 % is accessing the internet via mobile phone and average time of usage of internet of 4 hours 36 minutes. This depicts a population adopting these two technologies in every facet of their lives. In this regard, the students at universities and tertiary colleges use these technologies since most of them fall within the age bracket of 18 to 34 years that are easily reachable online according to Kemp (2021) on the Kenyan Data report. Most students, 83.6% have messaging applications installed on their phones and are adapting to communicating through these applications (Ojino, 2018) with mobile applications occupying 90% of the time spent, Smartinsights (2016, as cited by, Ojino, 2018). Most of the web traffic originates from mobile phones, Wearesocal (2017, as cited by Ojino, 2018).

When analysis is done at the end of a module or semester, several aspects will have been left unaddressed for a considerable time e.g. a month or three months. These will still impact on the learning processes and student’s learning experience even if they will be addressed. Another aspect that was not captured by Shankararaman et al, (2017) is the trainer’s perspective of use of feedback and how modern data capture methods like mobile application can be used for comments capture for students who are digital savvy and accustomed to applications like WhatsApp and Facebook. The college students engagement on mobile platforms in Laikipia campus was WhatsApp users (93.7%), Facebook (67.4%), among other social media platforms, spending at least two hours on daily basis (James et al,2021). There are changes in skills and expectations for the modern student, with the preference of use of mobile telephony for communication and cyberspace for seeking information rather than the traditional method like the use of physical library resources as observed by Williams et al, 2011(as cited by Conefrey, 2016). With this growing level of exposure and usage of these technologies, the old forms of relaying data and information in hardcopy form are being rendered irrelevant and inapplicable in various aspects of college activities and procedures.

Huxham et al, (2008), in comparing the Questionnaires, H Form, Rapid feedback, Focus groups and Student diaries as methods or tools for gathering student feedback made the following remarks on performance as indicated in **Table 1**:

TABLE 1
Tools Performance on Students’ Feedback

Feedback Method	Performance
Questionnaire	Convenient for final evaluations but don’t to give information to tutors for instructional delivery
Rapid feedback	Completion time is a few minutes for the student, with consultation time of about an 1 hour for tutors
Focus groups	Time consuming
Student Diaries	Too demanding for students, some failing to record comments while others misunderstood questions
H-Form	Considerable time required to organize and analyze

Author: Huxham et al, (2008)

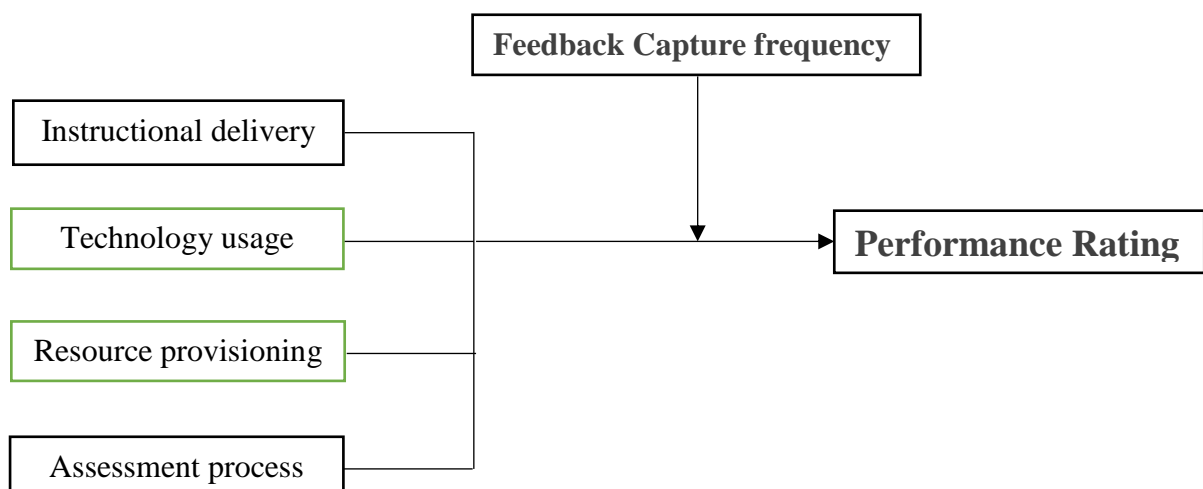
In seeking appropriate technologies for capture, repository and analysis of feedback data, Negara (2019) found that LDA performed 98% better than LSI in Topic Modelling. LDA

was better than Latent Semantic Analysis (LSA) finding topics on written essay according to (Kakkonen, 2008). LDA on twitter data outperformed LSI in accuracy on 1260 tweets with regard to word indexing in Sports topics, scoring 98% better, and is one of the appropriate statistical methods or algorithms for large (Negara, 2019) unstructured datasets in revealing hidden patterns.

2.4 Conceptual Framework

Monitoring the perceptions of stakeholders in a project, industry, organization or market is important in reorganization, policy formulation, designing of strategies as well as the overall management of physical, human or virtual resources. Feedback is the core of this undertaking if it is timely gathered and analyzed. The role of feedback in quality assurance is illustrated in **Figure 1 and 2** with regard to sustenance and improvement of level of quality in training delivery. Understanding the relationship of the variables (Rocco & Plakhotnik, 2009) in training delivery is the first step in appreciating the role of consistent feedback for appraisal and sustenance of acceptable levels of quality training in CBET.

FIGURE 2.
Conceptual Framework



Source: Author (2021)

2.4.1 Preselected Variables

IQA environment is established through planning, having college standards, procedures, and evaluation mechanisms for programmes activities, geared to satisfy stakeholders and ensure realization of quality graduates. Most of researchers reviewed concur on the importance of student's participation in IQA processes as opined by Taurina, (2015), in participatory learning according to Shen (2014) and participatory sensing (Burkey, 2006). This also depicts usage

technology for communication as a key feature as advocated by Rubang (2020) in the use of internet for virtual repository and Sahu (2008) in application of text mining technique for evaluating dashboards. In addition, the use of cyberspace as students' preferred method for seeking information was noted by Williams et al, 2011(as cited by Conefrey, 2016). There is a high degree of convergence on continuous monitoring of the quality of the education services (Jingura & Kamusoko, 2019), (Kelum et al, 2020) and use of continuous feedback according to Batista (1976, as cited by Brinko, 1993). Key areas of Internal quality assurance include review of programs, assessment of students, quality assurance of teachers, learning resources and student support, information systems according KENQA (2009 as cited by Kelim et al, 2020). Lecturer performance will not be included as a variable since it is embedded in other factors like instructional effectiveness.

IQA environment is established through planning, having college standards, procedures, and evaluation mechanisms for programmes activities, geared to satisfy stakeholders and ensure realization of quality graduates. With the environment of IQA in place the correct system for monitoring is formulated to promote consistency and enhancement. The major aspects in the implementations are communication and reinforcement of core values or design variables, to guide and create shared perspective amongst stakeholders (Simon, 1994, as cited by, Daromes and Suwandi, 2015), satisfy stakeholders and ensure realization of quality graduates (Tamaro, 2005). Appropriate teaching, learning and assessment requires creation and maintenance of a responsive, communicative and useful environment to achieve students' satisfaction (Douglas, McClelland, and Davies (2007). A combination of industrial approach and Total quality Management approach (TQM) to capture the needs of most important stakeholders (student) while capturing the most impactful functions (Tamaro, 2005) and processes in learning institution will provide a basis for IQA rating based on students' feedback.

Independent variables preselected before conducting the research survey include Instructional effectiveness (or delivery), Assessment process, Resource provisioning and Technology usage. These variables encompass and embed other aspects of learning student support, lecturer performance, which are evaluated under Instructional effectiveness, communication and resource provisioning. The overall effect results from these variables is a measure of success of training programs implementation. Increase in frequency of feedback will affect the overall performance rating especially when most comments leans on positive or negative side. These makes Feedback Capture frequency the moderating variable. The dependent variable is performance rating which is collective rating based on the students'

positivity rating on the selected variables (**Table 2.**). These variables reflect the major and most important processes that can impact on CBET approach and capture the interest of the digital savvy student.

The results of this system can be used for a college self-evaluation and enhancement through comparison of performance rating based on processes with the overall learning outcomes from students to facilitate decision making. This depicts the level of performance for decision making or prompt further investigation into any other possible or emerging factor on learning outcomes. Continuous capture of the data strengthens the monitoring and effectiveness over the period of colleges learning session, providing daily and continuous focus. The online and interview research captured the trainers’ perspective and observation with a view to classification of the variables, for refinement, affirmation, or confirmation

2.5 Operationalization of variables

TABLE 2
Preselected variables

Variable	Sub-variable	Indicator (Symptom)	Value (Numeric)
Feedback	Instructional delivery	Polarity Score	0.30843
	Technology usage	Polarity Score	0.37127
	Resource provisioning	Polarity Score	0.15770
	Assessment process	Polarity Score	0.27950
	Feedback Capture frequency	Polarity Score	0.37348
Performance	Service Rating	Rating (level)	0.27923

Source: Author (2021)

2.6 Summary

In the Education sector, the environment of both the learner and trainer should be maintained by Student’s feedback in the response to the existing conditions of the learning environment. When such data is appropriately captured, stored and analyzed in technical colleges, managers in learning institutions and government agencies can make better, relevant, up-to-date and informed decisions.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

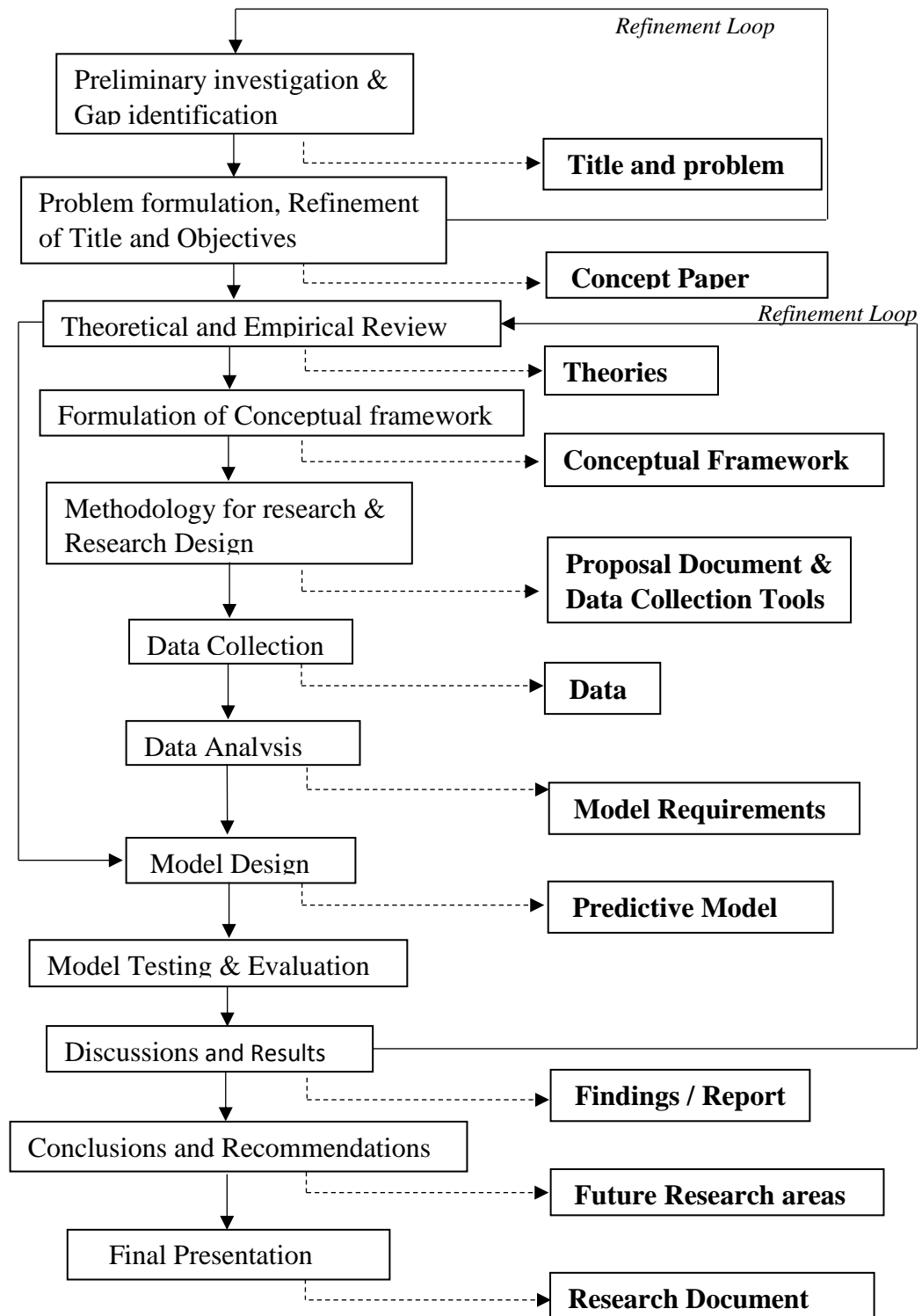
This is an exposition of the process of the research, comprising of methods and procedures in seeking the role of students' feedback, the data capture, storage, and analysis technologies employed in technical colleges in Kiambu County, in the development of the Predictive model for Internal quality Assurance. It consists of the research instrument, population of the study, method of Sampling, Instrument reliability and validity, the collection procedure for data, data processing, analysis, tasks timeline, and the cost of research. Qualitative and descriptive data will help in determination of the facets of students' feedback that account for change in performance rating of a college. This method is instrumental in getting the perspectives of respondents, the key phrases or themes that classify the respondents attitudes or experiences in order to discover their motives behind a decision according to Vinten, (1994, as cited by Jadhav and Khanna, 2016).

3.2 Research Design

This is an extensive plan (Sileyew, 2019) representing the various aspects of the study from problem formulation, topic refinement, literature review, data acquisition, data analysis, interpretation, model design and evaluation as well as presentation of findings. This elaborate plan is used to understand the appropriate form of students' feedback, and how it can be harnessed to positively impact on quality from the perspective of trainers and managers. This requires qualitative, descriptive and naturalistic data (Nassaji, 2015) from the learning environment in the area under study.

The phenomenon of use of students' feedback in evaluation of education programs and appraisal for trainers has been exercised in learning institutions in kenyan technical colleges. The trainers are at the center of this phenomenon as subjects, participants and promoters of the phenomenon. This study was instituted in order to understand how the training colleges ensures that the quality of learning is continuously improved and maintained, students feedback mechanisms, the trainers perceptions, opinions and observations on involvement of students' in quality assurance, the deployment of modern technologies in IQA, and the relevance of students' feedback in promoting CBET. The approach employed is both a descriptive research (Sileyew, 2019) and ethnographic design in order to gather and profile the phenomenon of

FIGURE 3
Milestones and Deliverables



Source: Author (2021)

interest from the opinions and observation trainers as the participants. Dulock, RN and DNS, (1993) refers to this design as a systematic, factual, accurate account to state the characteristics of the phenomenon. The descriptive aspect of the research design was to enable identification

of aspects of learning that the trainer and the managers in technical colleges consider important and relevant for evaluation through use of feedback from students. An exploration to understand the existing students' feedback mechanism, the role of feedback, and the way to harness and process students' feedback for internal quality assurance in technical colleges was undertaken. This descriptive research design was employed to acquire raw data from trainers. According Nassaji (2015), in making distinction between descriptive research and qualitative research, both acquire naturalistic data but are different since former specifically addresses the phenomenon and its characteristics. This was approach were considered appropriate and instrumental to obtain the feedback and quality assurance characteristics as expressed by trainers, forming the basis of the variables for the predictive model. Description of the feedback capture mechanisms, relevant technologies for analysis and repository were sought with the participants expected to vividly give the status and requirements for appropriate implementation. While the open-ended questions gathered descriptions, the close-ended questions were used to affirm the opinion and to rate the performance of the model by comparison of the result of both linear and Likert scale questions with positivity scores. Linear and Likert scale-based scales were used to decide the positivity on aspects of the study like technology, assessment process and Feedback Capture Frequency which are direct input to be compared, to augment and support the textual data in semantic analysis. This gave the opportunity of comparison and evaluation of whether the overall knowledge of the respondents is adequate or not, and their overall attitude is positive or negative (Hussain, 2013).

The ethnographic approach was to get the participants' sentiments on the feedback mechanisms which was relevant both for realization of independent variables and for the model testing and validation. The open-ended or textual data responses from the semi-structures questionnaire were meant for sentimental analysis to simulate students' feedback comments analyzed by VADER lexicon and rule-based algorithm. In the Google form the paragraph element provided input of the textual data which are the columns in CSV and pickle (PKL) files and selected for polarity scores to simulate student's feedback in model validation and testing. The Linear and Likert scale options were also used to test performance, and for comparison of VADER and TextBlob tools (Praveen and Prasanna, 2021) performance. The textual data for the independent variables were pooled together to assess the text corpus and was compared with the mean of the polarity scores of independent variables with each processed separately. This was to confirm the VADER and the overall model performance for small and large textual data (Hutto and Gibert 2014).

Data collection was carried out to gather experiences, and perceptions, (Tenny, Brannan, Brannan & Sharts-Hopko, 2017) of trainers, IQA officers and managers on internal quality assurance, students' feedback and the infrastructure requirements for utilization of the feedback. The stipulated milestones and deliverables were realized as proposed and indicated in **Figure 3**.

3.2.1 Research strategy

The research strategy was online survey and interviews. An Online survey was conducted via WhatsApp and Email, using mobile phones (Sue and Ritter, 2012). The Online survey method reduced data entry errors by automatic capture and some respondents preferred questionnaires allow answers at one own convenience. This also reduced the cost and time of data collection and facilitated storage and retrieval as opined by Yun and Trumbo, (2000 as cited by Marra and Bogue, 2006). Other respondents preferred filling a downloaded document and either send or the researcher to collect in hand copy. Where the WhatsApp link didn't work in some phones hand copy questionnaire were delivered and received as the researcher continued with the interviews in the fifteen Colleges. This arrangement of various methods for data collection together with TVET and NACCOSTI letters facilitated data collection. The snowballing sampling was utilized to get contacts of some respondents. Although a few participants responded via WhatsApp, a number of contacts were realized through snowballing.

The interview research approach opened a discourse and negotiation between the researcher and respondents allowing elicitation of information, collection and construction knowledge (Kvale, 2007) on students' feedback and IQA. This was partly an ethnographic research where the researcher solicited and elicited the knowledge of the experts in the field of education and the domain of learning (Hopf, 2004). The Selection of participants for interview research was through random sampling of fifteen (15) colleges and each provided one (1) trainer and one (1) participant at supervisory or managerial level. The Head of departments and their deputies, Internal Quality Assurance Officer (IQAO), Monitoring and evaluation officers, and other senior managers were categorized as supervisors for the purpose of these research to ensure the perspective of management was captured. This was necessary to provide the relevant experience and expertise in the study according to Cameron, (2005, as cited by Longhurst, 2003). Standard questions were formulated and derived from the studies from other researchers to capture points of agreement and points of departure or inadequacies as need arose before conducting structured interview (Dunn, 2005, as cited by, Longhurst, 2003). The semi-structured questions generated were based on the preselected variables in the reviewed

literature, leveraging prior knowledge from previous work of researchers to address the students' feedback and IQA problem. The questions in the online questionnaire were part of the semi-structured interview and additional questions were fielded for confirmatory purposes.

A Good Sample Design ensures that the sample realized is representative, lead to a small sampling error, reflect on cost implication, provide good control of systematic bias and lead to results of that can be applied generally with reasonable level of confidence, (Rai and Thapa, 2015). Purposive sampling, one of the most important non-probability sampling methods used to identify the primary participants in qualitative research according to Welman and Kruger (1999, as cited by Groenewald, 2004).

Analysis of data is meant to give a concise description of the phenomenon under study from the participant's perspective. Python libraries and Machine learning algorithms are suitable for analysis data captured by online survey and interviews. VADER is suitable for Natural language processing on textual information and was suitable for this study since it requires limited data for training, and performs better in most instances than other analysis tools like Linguistic Inquiry and Word Count (LIWC), Support Vector Machines (SVMs) according to a comparative study by Hutto and Gibert, (2014).

Most of researchers reviewed in this study conquer on the importance of student's participation in IQA processes as opined by Taurina, (2015), in participatory learning according to Shen (2014) and participatory sensing (Burkey, 2006). This also depicts technology usage as a key feature. Rubang (2020) advocated the use of internet for virtual repository and Sahu (2008) underscored text mining technique for evaluating dashboards. In addition, the use of cyberspace as students' preferred method for seeking information was noted by Williams et al, 2011(as cited by Conefrey, 2016). There is a high degree of convergence on continuous monitoring of the quality of the education services (Jingura & Kamusoko, 2019), (Kelum et al, 2020) and use of continuous feedback according to Batista (1976, as cited by Brinko, 1993). Key areas of Internal quality assurance include review of programs, assessment of students, quality assurance of teachers, learning resources and student support, and information systems according KENQA (2009 as cited by Kelim et al, 2020). Lecturer performance will not be included as a variable since it is embedded in other factors like instructional effectiveness.

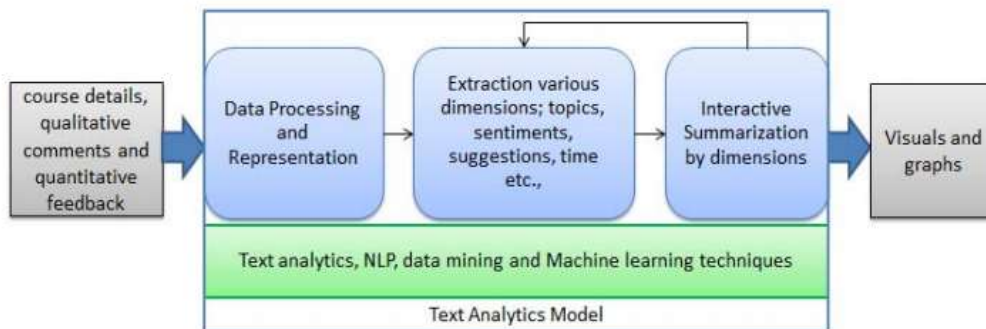
Independent variables preselected before conducting the research survey include Instructional effectiveness (or delivery), Assessment process, Resource provisioning and Technology usage. These variables encompass and embed other aspects of learning like programs, student support, lecturer performance, which is evaluated under Instructional

effectiveness and resource provisioning. The dependent variable is performance rating which is collective rating based on the students' positivity rating on the selected variables. These variables reflect the major and most important processes that can impact on CBET approach and capture the interest of the digital savvy student. The online and interview research captured the trainers' perspective and observation with a view to classification of the variables, for refinement, affirmation, or confirmation.

This methodology enabled description of events occurring in the learner's environment (LeCompte, 1999) for the development of the predictive model for IQA. The realization of a robust and reliable IQA system will then depend on organization readiness, appropriate leadership, and implementation of this technological architecture, (Daguang et. al, 2017) that deploys a mobile application for feedback capture, application of VADER for Natural language processing and other NLTK packages for Visualization, as well HDFS for repository.

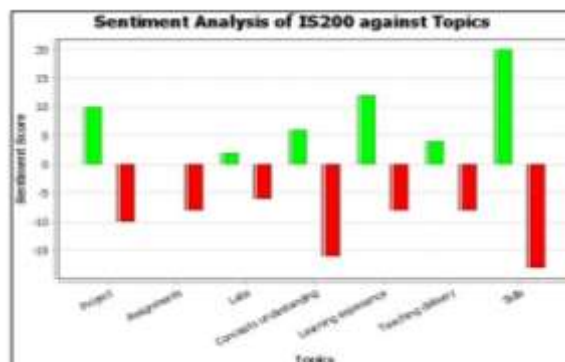
3.2.1 Data models and Pipeline

FIGURE 4
Text Analytic Model



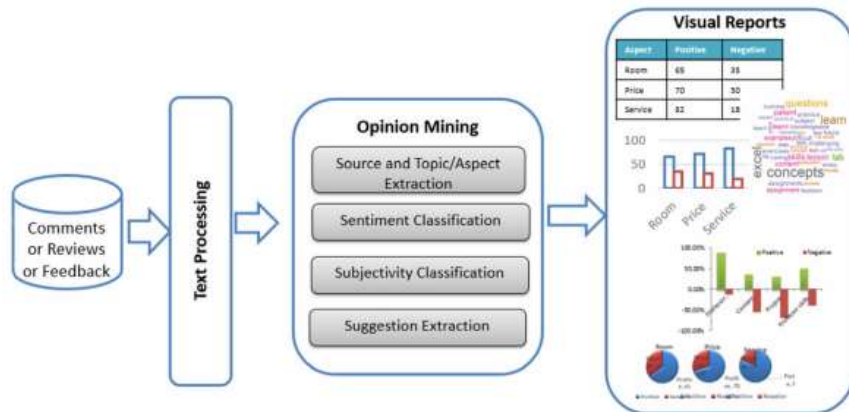
Source: Shankararaman et al, (2017)

FIGURE 5
Topics and sentiments



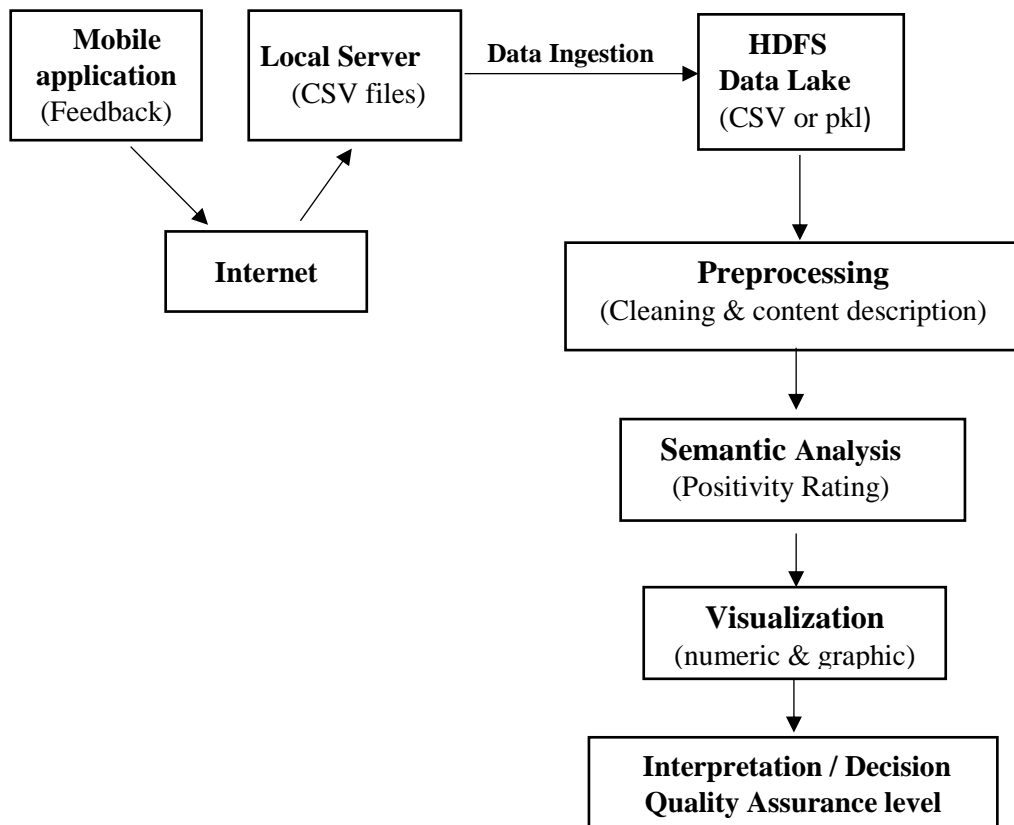
Source: Shankararaman et al, (2017)

FIGURE 6
Opinion Mining Architecture



Source: Gottipati et al (2018)

FIGURE 7
Students' Feedback Analytic Pipeline



Source: Author (2021)

3.3 Study Population

The research population for this study comprises managers, internal quality assurance officers, or trainers in their Colleges in Kiambu County in Kenya. These are active members in technical colleges with knowledge of the teaching-learning process.

3.4 Sample and Sampling Procedure

The techniques employed were Semi-structured interviews, online Survey, purposeful sampling and simple random sampling. Purposeful sampling technique was employed to access the information-rich cases, for the best utilization of available resources, as well as leveraging on participants' knowledge and experience, Patton, (2002 as cited by, Etikan 2016). This also exploited the participant's ability to communicate observations and sentiments with clarity (Etikan 2016).

Online and interview survey were conducted and the determination of the size of the sample was done to satisfy the criteria of getting cases with in-depth knowledge or characteristics required according to Patton, (2002 as cited by, Etikan 2016), reduction of sample bias and data saturation (Aguboshim, 2021).

Applying the formula $n = N / (1 + N(e)^2)$ according to (Yamane 1967, as cited by Israel, 1992) at A 95% confidence level and $P = .5$, with n as the sample size, N population size, and e as the level of precision, for 157 subjects would yield about 113 participants for one participant per college. In qualitative research data saturation is deemed to be reached in the absence of new impact, understanding, information, themes incidental to new data input from participants, and hence there is no need for more input (Aguboshim, 2021). In this qualitative research, the sampling was concerned on the quality of data that emanated from participants rather than the amount of data. The Yamane formula and other formulae like Cochran formula $n_0 = (z^2 * p(1-p)) / e^2$ are appropriate for quantitative research target high samples without close focus on quality, data saturation, time and cost constraints that attendant to this study and level. The sample was determined by Random purposive sampling (Omona, 2016) and snowballing to realize the participants for interviews and the online questionnaire. This study lay focus and emphasis on descriptions in detail that explains the experience of the participants and their expression of their learning environment using these qualitative sampling methods.

The sample size determination in this study was meant to justifying the sufficiency of the dataset for the research problem. Marshall (2013) echoed that in qualitative studies, there is a direct relationship between the number of interviews carried out and the cultural factors and therefore recommended fifteen to thirty interviews for a single case studies. This pool hence comprised people with experience on IQA and students feedback as the phenomenon under investigation (Sibona and Walczak, 2012) and was used in recruitment of participants. Although random sampling is usually applied in quantitative research, random sampling and purposive sampling methods were employed to realize a good sample. Fifteen (15) colleges were selected through random sampling from a list of 157 TVET colleges, each college

providing two participants for the interview survey. A list of 157 technical institutions colleges were sourced from TVETA website and saved in a comma separated file (CSV). These were processed by a python sample function in-built in random module to get 15 of 157 colleges.

The formula applied was: *college_list = [1, 2, 3.....157]* for the population of colleges and calculation for sampled colleges is; *sample_list= random.sample (college_list, 15)*. This identified 15 colleges each providing two participants, an IQA or a supervisor and a Trainer to make 30 participants for the interview survey.

The participants for the online survey were through purposive sampling and snowballing from the WhatsApp group of the internal quality assurance TVET workshop held from 14th to 18th of June 2021 at Kiambu Institute of science and technology and Snowball sampling (Jenkinson, (1997), to realize a population that 30 trainers or supervisors. These are qualified trainers under TVET in Kiambu County, competent to address the phenomenon of students’ feedback and internal quality assurance since they are conscious of the CBET requirements and interact with modern students. In order to increase the response rates for online surveys, solicitation was done as the researcher continued with the Interviews and where necessary visited the institution to collect a downloaded and filled form or to encourage responses online (Dillman, 2000, as cited by Deutskens, et al, 2004), for the participants obtained through snowball sampling.

Although the researchers’ target of total of 60 participants for both online and interview surveys, about 46 respondents were archived which is 76% response rate.

TABLE 3
Sample Size

	Staff	Population
Interviews	trainers	15
	Supervisors	15
Online questionnaires	Supervisors & Trainers	30
Total	Total	60

Source: Author (2021)

3.5 Research Instrument

The online questionnaire in google form was communicated via the mobile phone to obtain data through the written responses from participants. Semi-structured Interviews were captured through a similar questionnaire but on an excel worksheet and hardcopy questionnaire. A

provision was made for print out and filling hard copy where the link could not open in some phones. The purpose of these was to augment and ensure sufficient number of respondents for both the online and interviews questionnaire.

3.6 Instruments' Validity and Reliability

This study is qualitative research and captured only the responses of participants who are members of their technical colleges in Kamba County. In qualitative research, the level and measure of harmony between the description of the phenomenon and what actually happens in the world, achieved by continuous refining and sampling is the validity of the research, (McMillan & Schumacher (2006, as cited by Punjab, 2008). Ethnography is a qualitative research where the researcher interacts with the subjects in their own environment to fully understand the phenomenon under study.

According to Thayer (2010), reliability requires an explicit declaration of the role and status of the Researcher. In this regard, as a Trainer at Kenya Institute of Surveying and Mapping and registered under TVET, the opinion of the researcher was not captured hence did not influence the data processing and outcome for the study. The phone contacts of the subjects are captured for verification as well. Therefore, the measurement is the actual opinion and the belief of trainers on students' feedback and its role, the actual existing infrastructure, the data requirement and attendant system suitable for modern application.

The authenticity of online data capture is verifiable through the mobile phone numbers of participants. The Participants interviewed were also requested to give their phone numbers to reflect participation even though the data was not captured in an excel worksheet. Full documentation including the chain of evidence was done by storage of the online (mobile phone) and the laptop /interview questionnaire to give room for review or confirmation of responses by independent observers (Thayer, 2010). The study is therefore focused on a specialist group in the natural environment where the trainers are in their domain (technical colleges), thereby addressing the issues of representativeness, levels of control, and trustworthiness (Hewson et al., 2016).

3.7 Data Collection Procedure

This research relied on primary data from a sample of participants in technical colleges in Kiambu County. Identification of private and public technical colleges in Kiambu provided the basis of Interviews. The survey had two implementations, offline and online. The online survey via a mobile phone provided ease of access to the trainers' population across the Kiambu County without incurring heavy costs as indicated in the research budget (See **Appendix 6**) and within a short time (Hewson, 2015) as provided in the research schedule (See **Appendix**

5), This survey data was collected and saved in comma-separated format (CSV) to facilitate processing through VADER and other python libraries.

The online survey questionnaire was in the form of open-ended and close-ended questions. The semi-structured interviews were held by fielding the same questions with additional questions for confirmatory purposes.

A pilot study served to assess the appropriateness of data collection tools before the main study. According to Lackey and Wingate, (1998) a pilot study or survey facilitates the evaluation of time to complete the study, the ability to measure and for the clarity of the instruments.

3.7.1 Legal and Ethical Consideration

The approval of research proposal and an introductory letter from KCA University facilitated the issuance of the research permit or license (See **Appendix 3**) from National Commission for Science, technology and innovation (NACCOSTI) to conduct research in Kenya. In addition, a Letter approval from TVETA (See **Appendix 4**) facilitated access to technical colleges in Kamba County. Conformance to the stipulated objectives of the study and observance of policies in Institutions of higher learning like low similarity index percentage, were undertaken in the review of the work of other researchers and in conducting the survey research in the Institutions of learning.

3.8 Data Processing and Analysis

The first step was archiving raw copies of data files as a backup to allow preprocessing and analysis. Preprocessing involved understanding the collected data in terms of the number of records, identification of content by columns and rows. Culling the data was done to remove redundant or irrelevant data and segmenting the data into meaningful and useful units (Bruscia, 2005). One Row with null value was deleted to prevent undue skew of the results. Processing was be done by columns and the result be displayed using VADER module and other python libraries to facilitate interpretation and presentation. The documents for the online questionnaire were aggregated by using the generated CSV document from google form for structure. The Questions in the google form served as column labels which were later renamed to short, cleared labels for ease of understandability and processing (See **Appendix 2**).

CHAPTER FOUR

DATA ANALYSIS FINDINGS AND CONCLUSIONS

4.1 Introduction

This comprises a report of the findings and interpretation of this study. The response from the target participants is exposed and their background presented. The analysis carried out is presented for the variables under the study. This includes the independent variables, the moderating and the dependent variable.

4.2 Response Rate

The population of the colleges targeted was one hundred and fifty seven (157) and a sample size of 15 colleges selected using purposive sampling to get a detailed and rich information from people participating and experiencing the phenomena as (2002 as cited by, Etikan 2016). Each college provided two respondents for the interview both from the management and also from ordinary trainers or instructors to get the overall perspective of the phenomenon of student's feedback and its relevance in the provision of quality learning.

A pilot study or survey was undertaken involving 7 technical college trainers and two participant Interviews to identify any flaws in the questions based on the responses. These identified some flaws in framing and redundancy in questions. One flaw was lack of appropriate label for one multiple choice question identified in the online questionnaire. The pilot interviews discovered more inadequacies because of follow up questions resulting to identification of ambiguity and redundancies. These facilitated the realization of a restructured and rationalized final research questionnaire attached (See **Appendix 2**).

In the main study some colleges were hesitant to allow their trainers to provide the requisite data until assurance was clearly made that their identity and that of the college would not be captured. Two colleges declined access of information from their trainers citing their organizational policies even after confirmation of researcher's identity and full assurance by letter from NACCOSTI (See **Appendix 3**) and approval and facilitation letter from TVET (See **Appendix 4**). A total of 27 questionnaires from interviews 13 technical institutions duly filled were realized. This gap was filled by snow balling (Jenkinson, (1997) by additional two colleges to realize a total of 30 questionnaires filled by way of interviews. Some online participants requested softcopy questionnaire citing the link not opening the google form. This were provided for as the researcher identified their institution and proximity to the other institutions earlier identified for interviews. This yielded 16 questionnaires of which 10 were

from google forms and 6 from contacts through purposive sampling as well as snowballing sampling. This represents a response rate of 77 % as shown in **Table 4.1**.

TABLE 4
Response Rate

Response category	Target respondents	Responses
Interview questionnaires	30	30
Online & Softcopy questionnaires	30	16
Total	60	46
	100%	77%

Source: Author (2021)

According to Mugenda and Mugenda (2009, as cited by Wambui & Gachuho) a response rate of 50% is adequate for analysis and reporting; a rate of 60% is good and a response rate of 70% and over is excellent. Thus, following Mugenda and Mugenda’s assertion, the established response rate is satisfactory to make study inferences.

The delivery of the semi-structured questionnaire was both in softcopy from the laptop and hardcopy as well where time restriction required fast data entry. The Gmail account was instrumental in accessing the Google forms for the online questionnaire. The delivery of the online Questionnaire was hence by the Google form and WhatsApp applications. Follow-up communication was done through WhatsApp via mobile phone for confirmation of access to the questionnaire. Where the link failed due either due to either missing characters or phone inadequacy, a softcopy Questionnaire was sent for download and completion to ensure that the potential and identified respondents presents the opinions, observations and experiences on students feedback and the quality of learning in technical colleges. This was geared to achieving the appropriate response rate for the adequacy of the research.

Descriptive summaries were realized from the open and closed ended questions by both the interviews and the online questionnaires to harness the knowledge from participants of the survey and also to provide a ground for further research (McIntosh and Morse 2015).

4.3 Description of Participants

The Participants although bearing varied designations served as trainers in their institutions. The researcher classified the tutors, trainers and lecturers as trainers while the rest serving supervisory roles hence qualified as supervisors for the purpose of these study.

TABLE 5
Survey Respondents from Kiambu County

Participants Title	Number	Category
Tutor	10	Trainer
Deputy HOD	6	Supervisor
Head of department (HOD)	3	Supervisor
Lecturer	2	Trainer
Dean of students	2	Supervisor
Monitoring &_evaluation co-coordinator	2	Supervisor
Accountant	1	Supervisor
Manager	1	Supervisor
Trainer	18	Trainer
Customer service officer	1	Supervisor

Source: Author (2021)

Out of the 46 respondents, thirty-four (34) of them have participated in workshops or engaged in actual training. This depicts 74 % exposure to programmers in competence-based training and are hence reliable to give their opinion on students' feedback requirements in Technical Vocational training colleges in Kiambu County. Only 14 participants accounting for 26 % of Training staff are without sufficient exposure in CBET mode of training programmes.

4.4 Research Findings

This is a description of the outcome of the research after review of literature and undertaking research survey and analysis.

4.4.1 Objective One results

The first objective was to identify the quality assurance features under which students' feedback will be classified. The processing was carried out using python libraries and VADER (lexicon and rule based) algorithm as indicated in the code (See **Appendix 1**). The results are illustrated for the variables in the **Table 7**.

TABLE 6
Positivity Rating

	Negative score	Neutral Score	Positive score	Compound Score
Variable				
Instructional delivery	0.01660	0.30843	0.54693	0.30843
Technology usage.	0.00740	0.37127	0.49500	0.37127
Resource provisioning	0.01110	0.81184	0.15530	0.15770
Assessment process	0.00576	0.60590	0.38830	0.27950
Mean of the Independent Variables				0.26496
Continuous Feedback preference	0.01295	0.56040	0.42663	0.37348
Performance(service) Rating	0.02126	0.62920	0.34947	0.29141

Source: Author (2021)

In **Table 7**, the compound rating of positive score 0.308432 for Instructional delivery indicates support and importance of use of feedback instructional delivery. Other variables with good support from participant are technology usage with a positive score of 0.37127, assessment process with a score of 0.37127, Continuous Feedback preference with positive 0.37127 and performance rating with a positive score of 0.29141. Resource provisioning has a low positive score of 0.1577 but is supported by the fact that the respondents consider students being aware of the resources with 67% as indicated in **TABLE 10**. This is an indication that student's comments can be variable in monitoring the status, availability or adequacy of the learning resources that include library books, equipment, devices, computers etc.

TABLE 7
Website and Mobile application usage

	Highly efficient	Efficient	Moderately efficient	slightly efficient	Inefficient
Item					
Website Rating	39.13	41.30435	19.565	0	0
Mobile Rating	36.9565	34.78260	13.04348	10.86957	4.34783

Source: Author (2021)

A college Website and mobile applications are considered efficient at 80% and 72% respectively of the respondents, **TABLE 8**. 43.5 % respondents require feedback capture done often while 50% supports being carried out always. This emphasizes the need to deploy mobile technology for continuous feedback independent of special location.

Internet for normal learning as indicated in **TABLE 9**, got a combined 93 % underscoring the use of technology to support all learning activities including students’ feedback. Instructional delivery 80% and above in importance in contribution to the quality of learning hence a key variable for the students feedback model.86 % agree that there is need to deploy of modern technologies for capture, analyses and storage of assessment data. This together with Rating of a positive score on usage of Technology *depicts the* Technology usage as a key variable in a student’s feedback model

TABLE 8

Internet usage and Instructional delivery rating

	Very important	Important	Moderately important	slightly Important	Unimportant
Item					
Internet in normal learning	67.39131	26.08696	4.348	2.1739	0
Instructional delivery contribution to quality	39.13044	41.30435	19.565217	0	0

Source: Author (2021)

TABLE 9

Resource Awareness and Technology Requirement

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Item					
Resource Awareness	15.21739	52.17391	26.086957	2.1739	0
Technology requirement	35.55556	48.88889	8.888889	4.444444	2.222222

Source: Author (2021)

The frequency of feedback is crucial in monitoring IQA as depicted by 50% respondents routing for capture of feedback ‘Always’ 43.5% ‘Often’ while 6.5% ‘Sometimes’.

Therefore a combination of the respondents for ‘Always’ and ‘Often’ (93.5%) implies continuous feedback is better and preferred than a college end-of-term feedback as indicated in **Table 11**. These affirms and confirms the good support for Continuous Feedback preference in **Table 7**. In totality, the best mode of capture of feedback is continuous without time or spatial restrictions.

TABLE 10

Feedback capture frequency

	Always	Often	Sometime	Rarely	Never
Mode of feedback capture	50.00000	43.47826	6.521739	0	0

Source: Author (2021)

4.4.2 Objective Two Results

Spark distributed data processing engine was deployed for data ingestion of local files to Data Lake in HDFS as indicated in appendix 1. The similitude of a students’ feedback file was archived with a file path 'hdfs://localhost:9000/datalake/archived_student_feedback.csv' (See **Appendix 1**). Hadoop daemons or processes were launched using start-dfs.cmd and start-yarn.cmd to view the saved file the address and port used is http://localhost:9870/.

The data from the Data Lake were saved to a local file in the directory path C:\ResearchData\ feedback_questionnaire.csv. Extraction, Loading Transformation and (ETL) was undertaken before processing the feedback file using the selected columns in **table 12**. The Selected columns are ID', 'Use_of_techn',' Inst_deliverly', 'resource_req', 'Assess_process', 'Feedback_freq

TABLE 11

Model Data Files

Data Lake file	'hdfs://localhost:9000/datalake/archived_student_feedback.csv'					
Local File	' C:/ResearchData/feedback_questionnaire.csv '					
CSV columns	ID	Use_of_techn	Inst_deliverly	resource_req	Assess_process	Feedback_freq

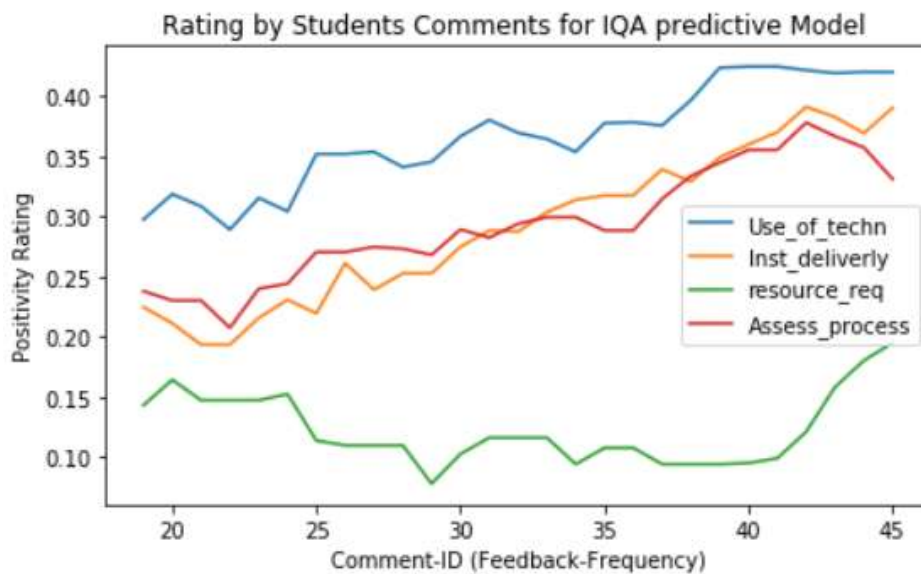
Source: Author (2021)

A code designed to predict the positivity rating was realized and tested on the research data as illustrated in Figure 9. Analysis of the research data revealed that four variables would serve as independent variables as depicted by the graphic results from python libraries and VADER

(lexicon and rule-based algorithm). All the variables obtained a positive score from the survey respondents. The Mean was calculated to get the positivity rating of 0.2792255 from the four variables as a measure service rating college performance. Data ingestion was demonstrated by moving file into the Hadoop Data Lake using python libraries in testing as demonstrated in the Code (See appendix 1). A Mobile application will receive data input and send data into a comma separate file (CSV) in Hadoop file system (HDFS).

FIGURE. 8

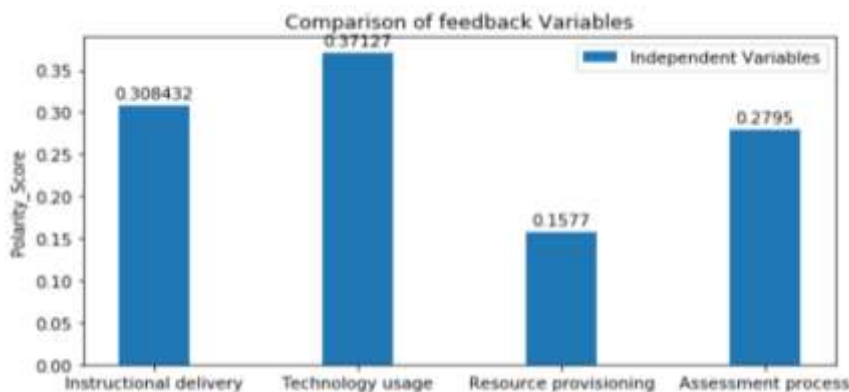
Visualization of independent variables



Source: Author (2021)

FIGURE. 9

Positivity Ratings for Independent Variables



Source: Author (2021)

The feedback frequency is represented by comments increase in the graph of variables since they are you unique and are unit in the x-axis (**figure 8.**) and figure 10 with a count of 46.

4.4.3 Objective Three Results

The mean of the four independent variables calculated using Python sum and mean functions on the results from VADER scores for each independent variable gave a Positive rating score of 0.26496 as indicated in **Figure 9**. This represents respondents support for use of students' feedback on quality assurance. Processing by VADER realized a Performance (service) Rating of a positive score of 0.29141 from respondents as the dependent variable. Therefore there is a strong relationship between the mean of the independent variables and the feedback rating on dependent variable or overall service delivery of a college, with both recording positive support values hence underscoring the importance of students feedback for IQA.

TABLE 12
VADER and TextBlob Scores by column content

Item	All Variables	Technology usage	Instructional delivery	Resource provisioning	Assessment process
TextBlob scores	0.36611	0.38789	0.41344	0.34967	0.28793
VADER scores	0.99970	0.37127	0.30843	0.15773	0.27951

Source: Author (2021)

TABLE 13
Scores on Linear scale Importance measurement

Item	Very important	important	Moderately important	Slightly Important	Unimportant
TextBlob scores	0.5200	0.4000	0.4000	0.4000	0.400
VADER scores	05600	0.20230	0.20230	0.1298	-0.3182

Source: Author (2021)

TABLE 14
Scores on Linear scale Efficiency

Item	Highly efficient	efficient	Moderately efficient	Slightly efficient	Inefficient
TextBlob scores	0.1600	0.0000	0.0000	0.1667	0.0000
VADER scores	0.5400	0.4215	0.4215	0.3626	0.0000

Source: Author (2021)

TABLE 15**Scores on Likert scale Agree measurement**

Item	Strongly Agree	agree	neutral	disagree	Strongly disagree
TextBlob scores	0.4300	0.0000	0.0000	0.0000	0.43333
VADER scores	0.7300	0.3612	0.0000	-0.3818	-0.1280

Source: Author (2021)

The combined text from independent variables produces results that differed greatly between the TextBlob and the VADER tools with the VADER scoring 0.9997 while TextBlob score was 0.36. This affirmed the study by Hutto and Gibert, (2014) that VADER is good for small text sentimental analysis. Although the text generated from student's comments can be large, the individual comments are small based on the devices and the interface and the purpose for which they are intended. The processing was done in rows for accuracy and to facilitate comparison within a single comment by ID as supported by parallel graph for visualization. After processing by rows, summation and averaging was done on the cumulative polarity from each comment. The cumulative polarity on each row gave good the results (**figure 8** and **9**) giving the appropriate capture and analysis for IQA based on students feedback and deploying VADER.

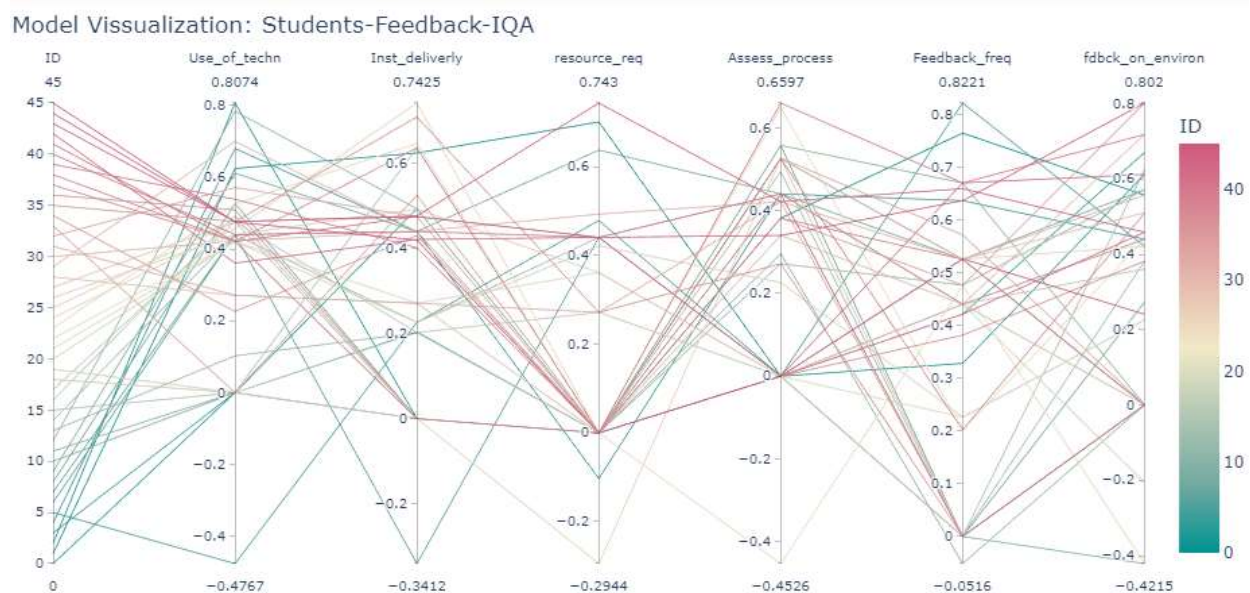
TextBlob gave lesser quality in positivity results on known labels used in the questionnaire without clear distinction of progressively increase on decrease in the responses linear and Likert scales as indicated in **Table 12, 13** and **14**. On the other hand VADER performed better showing progressing gradation on the labels based on the real life value attached to them. A major error in Textblob is rating "Strongly agree" the same with 'Strongly disagree' since they should be exactly opposite in value and side of polarity (**Table 14**). The formulation of the model to analyses rows independently then get cumulative results enhances its performance based on the efficiency of VADER textual data as well as the data capture mechanisms for the model.

The visualization was done by comment (**figure 9** and **10**) as captured by comment identification number. This facilitated comparison of rating on one variable by a students in relation to the other comments on other variables on the same unique identity. This is to promote deciphering patterns of opinions and perceptions, levels of satisfaction on various aspects of learning as well as the variables greatly contributing either positively or negatively to performance. This forms the basis of decision making to boost or maintain the good

performance in various aspects of a technical college and the overall service provision to the learners.

Adoption of modern technologies for participatory sensing through proposed interconnection of technologies to address IQA, student’s feedback, assessment facilitation and instructional delivery monitoring. The technologies include mobile application for data capture, Lexicon and rule-based algorithms for semantic analysis and Hadoop Distributed File system (HDFS) as data repository and python libraries like Spark for data ingestion and retrieval. Propagation of the phenomenon of participatory sensing ((Bur key, 2006) is enhanced by this data pipeline and methodology for measuring Internal Quality Assurance by exploiting both the efficiency of modern technologies, and the students digital skills and capabilities.

FIGURE. 10
Visualization of Model Output



Source: Author (2021)

4.5 Discussions of Results

The predictive model in this study differs from the one proposed by Praveen and Prasanna, (2021) that was on prediction of student’s performance using K-means clustering and LSA methods. The one by Sorour, et al was on student’s performance through comments at the end of each lesson to understand their attitudes. Parahoo et al (2015) proposed model had similarity with the one in this study by measuring students’ satisfaction but differs by focusing online learning environment which has dissimilar unique factors contributing to students learning experience. The difference includes factors in the online space like connectivity, interactivity, synchronicity, audibility, devices and presentation software etc. Predictive models proposed by

Srungaram, (2017) focused on the approaches in teaching and the trainers performance as opposed to all the factors integral to sustenance and improvement of quality as advocated in this study.

The objective of students' feedback model is to equip the managers in promoting effectiveness in teaching and to facilitate decision making according to Marsh and Dunkin (1992, as cited by Brennan (2003)). The use of mobile telephony and cyberspace is becoming more and more entrenched in the culture of students and trainers. This is becoming the method of communication and transactions. Modern technologies like tablets should now be adopted for capture of skills in assessments to meet the CBET requirements and needs. Zinzou1 and Rubang (2020) emphasized on exploiting the internet and adoption of modern techniques in application of knowledge for quality assurance in education.

The Data from feedback variables are presented using parallel coordinate in **figure 10**. for interrogation interactively through brushing to overcome occlusion and over-plotting hence view values of interest accounting to changes in IQA index or service rating. Through selection, highlighting and filtering or masking, the relationship of the variables is interrogated. Data analytical algorithms like VADER and other libraries in Python together with repository systems like HDFS are required for processing and for storage of large volumes video and audio data. The VADER lexicon and rule-based algorithm was used evaluation of the model using the study data since it has proved to be versatile according to a comparative study by Hutto and Gibert, (2014).

The quality of education can greatly be enhanced through Continuous monitoring, right policies and implementation of good polices (Jingura & Kamusoko, 2019). The capture of students' feedback based on these findings should be without constraint and restrictions to capture the right data from the student as the main stakeholder in education. All the variables captured were found to be important by both positivity ratings harnessed from opinions of participants in the open-ended questions and the degree of agreement as captured in the closed-ended questions. This model reduces complexities by having fewer variables for effective rating of performance by sentimental positivity (**Table 12**) and where there are some similarities with other models by Shankaraman et al, (2017) in Figure 4 and 5, parallel coordinates is used to allow as filtering effectively mask variables and providing a view desired variables **Figure 10**.

The Students' Feedback Analytic Pipeline (**Figure 1**) in this study is detailed and updated to reflect the application of modern technologies like model telephony, the internet and HDFS, for efficiency and ease of data capture, storage and analysis.

4.6 Summary

The use of technology for capture of student's feedback will improve the process of interaction of the students with the college processes and personnel to create a better and modern participatory environment. The monitoring and evaluation of college processes and facilities will help in keeping learning environment updated and promote satisfaction of the trainee.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This is a summary of the findings of the study based on the main objectives and guided by the theoretical framework. It covers the study objectives, the conclusions derived from the theoretical review and survey research as well as the recommendations on research related to the current study.

5.2 Conclusions

Student's feedback should be harnessed using the four main processes of a technical college. The variables are Instructional delivery, resource provisioning, and assessment process and technology usage. A Mobile application can capture and send data to a central repository system for storage and analysis. The processing by an algorithm like VADER presents a simplified way of analyzing opinion or sentiments from textual data.

5.3 Contribution to the study

The predictive model provides a simple and fairly accurate method of data capture facilitated by the internet, messaging application and the mobile telephony. This provides an easy and simple mechanism for feedback while according the digital savvy students infrastructure commensurate with expectations and technological skills. Few independent variables have been realized that include assessment process, resource provisioning, instructional delivery and technology usage. The feedback frequency is the moderating variable that can influence the dependent variable or output for the model

The proposed data pipeline and IQA model will facilitate sensing the extent to which the activities, events and processes set-out in Kiambu CBET based learning ecosystem promote the trainee in acquisition of Skills and requisites knowledge in each trade area. The phenomenon of participatory sensing is extended in this study where the student, the mobile phone, an application embedded in the phone and the internet are combined to form a sensing mechanism in learning.

Analysis of feedback through VADER in the data pipeline will facilitate easy, fast and reliable results of the opinion of students forming the basis of decision-making for assurance of internal quality of a technical college. The HDFS in the data pipeline will support storage and retrieval even for large institutions and serve as an efficient repository with provision for concurrent processing. Although Tammaro, 2005) addresses a paradigm shift of measuring learning by laying emphasis on input like contact hours in teaching hours to output based methodologies using learning outcomes and competencies. The approach promoted in this

study is to monitor the status of the learning environment for continuous improvement. The shift is from measuring learning by input or output to monitoring the environment of learning for continuous improvement based on feedback from students. The model focuses on Students' learning experience rather than an institutions' experience. This addresses how the student participate, responds, experience, or feel while operating under set policies, through processes and procedures, in events, and through timelines as set in college schedules and sessions.

This study therefore helps focus on pedagogical issues like weak QA mechanisms and insufficient skilled trainers as reported by Akala et al (2018), support full implementation CBET, facilitate Timely and updated information on the trend in the environment, reduce complexities in feedback capture and IQA mechanisms, and empower the digital savvy student as main stakeholder in the learning ecosystem.

The overall effect of adoption of the predictive model is enhancement of Business Intelligence (BI) in the context of learning institutions by effectively promoting students learning experience, improve the learning environment, enhance decision making and giving a technical colleges a competitive advantage (Negash 2004)

5.4 Recommendations for future research

In processing the combined text corpus from the independent variables VADER was outperformed by TextBlob python library. TextBlob had a score of 0.36611, which was close the result (0.26496 score) calculated from the mean of independent variables while VADER produced almost a perfect positive (0.99997 polarity score) which was not feasible from all the variables. This meant there was almost complete concurrence on every independent variable under the study by all the participants which was serving as college performance based on student's feedback. The choice of processing variables independently was fruitful and suggest that the data capture messaging application should restrict the input to a about few lines for accuracy required to be realized. The dependent variable was therefore a result of cumulative summation of processed rows of each individual independent variable and then the results is averaged based on the four variables to get the performance index or multiplied by 100 to get percentage performance.

Internal Quality assurance System based on students feedback will require a mobile application for data capture. This can have the form of Online-Chat application or other tools like Google-form with provisions to select the topic area on which to make comment.

A machine learning algorithm can be developed to ensure grammatical correctness in typing to ensure the words are within the lexicon of a specified language adopted for a college. This will promote sentimental analysis using lexicon, rule based and machine learning

algorithms. One way could be by a dictionary to complete entry of words based on a lexicon or give suggestions of possible words to avoid a situation where a respondent writes the word GOO instead of GOOD as experienced by the researcher during data exploration. Such a word will have a strong neutral value while the intended word had a high positivity score in sentimental analysis based on lexicon.

An optional feature for the mobile application would be to have provision for linear scale for rating with visual objects (e.g.wedge) to indicate the strength of options alongside descriptions for accurate and easy data entry. The mobile application is necessary for continuous feedback capture and relaying the data online to a file in a data lake like HDFS.

In addition, further research should be carried out for an appropriate mechanism for trainers to respond to the management for their views to be considered in implementation of quality assurance.

REFERENCES

- Aguboshim, F.C. (2021) Adequacy of sample size in a qualitative case study and the dilemma of data saturation: A narrative review. *World Journal of Advanced Research and Reviews*, 2021, 10(03), 180–187. DOI: <https://doi.org/10.30574/wjarr.2021.10.3.0277>
- Akala, Winston J and Changilwa, Peter (2018). Status of Technical and Vocational Education and Training (TVET) in Post-secondary Education in Kenya. *Journal of Popular Education in Africa*. 2(7), 15-25.
- Amollo, B. A. (2011). Digitization for libraries in Kenya. *International Conference on African Digital Libraries and Archives (ICADLA-2)*, University of Witwatersrand, Johannesburg, South Africa.
- Angelo, C. D. (2018). The Impact of Technology: Student Engagement and Success. *Technology and the Curriculum: Summer 2018*. <https://techandcurriculum.pressbooks.com/chapter/engagement-and-success/>
- Atmowardoyo, H.(2018) Research Methods in TEFL Studies: Descriptive Research, Case Study, Error Analysis, and R & D. *Journal of Language Teaching and Research*, 9(1), 197-204, DOI: <http://dx.doi.org/10.17507/jltr.0901.25>
- Bakarman, A. A. (2005) Attitude, Skill, and Knowledge: (ASK) a New Model for Design Education. *Proceedings of the Canadian Engineering*
- Bashir, A. M., Kabir, R., Rahman, I. (2016). The Value and Effectiveness of Feedback in Improving Students' Learning and Professionalizing Teaching in Higher Education. *Journal of Education and Practice*, 7(16)
- Brennan, J., Brighton, R., Moon, N., Richardson, J., Rindl, J. & Williams, R. (2003) Collecting and using student feedback on quality and standards of learning and teaching in higher education. A report to the Higher Education Funding Council for England
- Brinko, K. T. (1993). The Practice of Giving Feedback to Improve Teaching: What Is Effective?: *The Journal of Higher Education*, 64(5,) 574-593: Taylor & Francis. <https://doi.org/10.2307/2959994>
- Bruscia, K. E. (2005). Data analysis in qualitative research. *Music therapy research*
- Burkey, J. Estrin, D., & Oarker, A. (2006). Participatory sensing
Callegaro, M. Manfreda K. & Vehovar, V. (2015). *Web survey methodology 2015* - books.google.com
- Conefrey, T. (2016). Technology in the College Classroom: Crisis and Opportunity. *Educational Technology*, 56(4), pp. 37-40. <https://www.jstor.org/stable/44430476>

- Cote, F. (2016). Evaluating Attitudes: Tools for more objective assessments. *Pedagogie collegiale*. 29(3). file:///E:/evaluatin%20attirtudes.pdf
- Daguang, W., Zuoxu, X., Fan, W., & Yanjie, Q. (2017). *Enhancing Teaching and Learning through Internal Quality Assurance*. Paris: International Institute for Educational Planning.
- Daromes, F. E. & Suwandi, N. G. (2015). Indonesia Embedding core value into the internal quality assurance systems in higher education. *2nd Global Conference on Business and Social Science-2015, GCBSS-2015, 17-18 September 2015*
- Deutskens, E., Ruyter, K. D., Wetzel's, M., & Oosterveld, P. (2004). Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study. *Marketing Letters (15)1*, 21–36, 2004
- Dulock, H. L., RN, & DNS (1993) Research Design: Descriptive Research. *Journal of oncology Nursing*, 10(4), pp.154-157
- Douglas, J., McClelland, R. & Davies (2007). The development of a conceptual model of student satisfaction with their experience in higher education. *Student satisfaction with HE*. www.emeraldinsight.com/0968-4883.htm
- Dunrong, B. & Fan, M. (2009). On Student Evaluation of Teaching and Improvement of the Teaching. Quality Assurance System at Higher Education Institutions. *Chinese Education and Society*, vol. 42, no. 2, March_April 2009, pp-115. DOI 10.2753/CED1061-1932420212
- Dringus, P. L., & Ellis, T. (2005). *Using data mining as a strategy for assessing asynchronous discussion forums*. *Computers & Education*, 45 (2005), 141–160
- Ellis, R. (1995) *Quality Assurance for University Teaching*. The Society for Research into Higher Education 8c Open University Press, Buckingham MK18 1XW. USA
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*. 5(1)1-4. Doi: 10.11648/j.ajtas.20160501.11
- Feeney, E. J. (2007). Quality Feedback: The Essential Ingredient for Teacher Success. *The Clearing House*. 80(4). pp 191-197. Taylor & Francis. <https://www.jstor.org/stable/30192148>
- Ferej, A. (2020). National TVET Standards. Kenya Report 2020.
- Gilbert, P. (2021). Kenya Internet usage lags mobile penetration <https://dataportal.com/reports/digital-2021-kenya>
- Gottipati, S., Shankararaman, V. & Lin, R. J. (2018). Text analytics approach to extract course improvement suggestions from students' feedback. *Research and Practice in Technology Enhanced Learning*. <https://doi.org/10.1186/s41039-018-0073-0>

- Groenewald, T. (2004). A Phenomenological Research Design Illustrated International. *Journal of Qualitative Methods* 2004, 3(1)
<https://doi.org/10.1177/160940690400300104>
- Hamid, Y. & Mahmood, S. (2010). Understanding constructive feedback: A commitment between teachers and students for academic and professional development. *J Pak Med Assoc*, 2010 - *academia.edu*
- Handbook of Online Research Methods (2nd Ed.). London: Sage, 57–75.
- Harvey, L. (2003). Student feedback. *Quality in Higher Education*, 9(1)
- Hussein, N. (2015). Qualitative and descriptive research: Data type versus data analysis. *Language Teaching Research*. 19(2) 129–132
- Hewson, Claire (2017). *Research design and tools for internet research*. Sage Handbook of Online Research Methods (2nd Ed.). London: Sage, (57–75).
- Hopf, C (2004) Qualitative interviews: An overview. A companion to qualitative research, 2004 - books.google.com
- Hussain, Z. (2013). Methods For Decision making In Survey Questionnaires Based On Likert Scale. *Journal of Asian Scientific Research*, 3(1), 35-38
- Hutto, C.J. & Gibert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*
- Huxham, M., Laybourn, P., Cairncross, S., Brown, N., Goldfinch, J., & Earl, S. (2008). Collecting student feedback: a comparison of questionnaire and other methods. *Assessment & Evaluation in Higher Education* 2008, 1--12
- Jadhav, V., & Khanna, M. (2016). Factors Influencing Online Buying Behavior of College Students: A Qualitative Analysis. *The Qualitative Report*, 21(1), 1-15.
<http://nsuworks.nova.edu/tqr/vol21/iss1/>
- James, N., Janerose, M., Isaac, K., & Wanjohi (2021). Media Usage and Students' Psychosocial Wellbeing in Laikipia University, Kenya. *American Journal of Sciences and Engineering Research E-ISSN - 2348 -703X*, 4(3), 2021
- Jeffrey, A. B., Burke, D. E., Mark. H., Andrew. P., Ramanathan, N., Reddy, S., & Srivastav, M. B. (2006). Participatory sensing. Center for Embedded Networked Sensing (CENS), University of California, Los Angeles.
<https://escholarship.org/uc/item/19h777qd>
- J Sci Educ Technol (2014) 23:98-107 DOI 10.1007/s 10956-013-9453-9
- enkinson, C. (1997) *Assessment and evaluation of health and medicare. A methods text*
 Buckingham, Philadelphia's: Open University press

- Jingura, R.M. & Kamusoko, R. (2019). 'A competency framework for internal quality assurance in higher education', *Int. J. Management in Education*, 13(2), 119–132
- Israel G.D. (1992) Determining sample size. 1992- researchgate.net
- Kakkonen, T., Myller, N., Sutinen, E., & Timonen, J. (2008). Comparison of Dimension Reduction Methods for Automated Essay Grading. *Educational Technology & Society*, 11(3), 275-288
- Kelum A. A. G., Roshan, R. G. G. Pradeep, Najdanovic-Visak, R. G. G. and Gunawardhana, N (2020). Academic Standards and Quality Assurance: The Impact of COVID-19 on University Degree Programs. *Sustainability* 2020, 12, 10032; doi: 10.3390/su122310032
- Kemp, S. (2021-February). Digital 2020 Kenya: <https://datareportal.com/reports/digital-2021-Kenya>
- Kenya Vision 2030. A Globally Competitive and Prosperous Kenya Retrieved: <https://www.opendata.go.ke/>
- Kis, V. (2005). Quality Assurance in Tertiary Education: Current Practices in OECD Countries and a Literature Review on Potential Effects. *Tertiary review* www.oecd.org/edu/tertiary/review.
- Kvale, S. (2007). *Doing Interviews*. 2012 - books.google.com
- Laws of Kenya (2013) Technical and Vocational Education and Training No. 29 <http://kenyalaw.org:8181/exist/kenyalex/sublegview.xql?subleg=No.%2029%20of%202013>
- Lackey, N. R. & Wingate, A. L. (1998) *The Pilot Study: One Key to Research Success*.
- P. J. Brink, M. J. Wood *Advanced design in nursing* (2 ed), P. J. Brink, M. J. Wood. London, New Delhi, Sage Publications
- LeCompte, M.D., & Schensul, J. J. (1999). *Designing and conducting ethnographic research. The Ethnographers Toolkit 1*
- Longhurst, R. (2003) *Semi-structured interviews and focus groups. Key methods in geography*, 2003 - books.google.com
- Makori, E. O., Mauti, B., Norah, O. (2011). Digitization for libraries in Kenya. International Conference on African Digital Libraries and Archives (ICADLA-2), University of Witwatersrand, Johannesburg, *South Africa. Library Philosophy and Practice (ejournal)*. 1379. <http://digitalcommons.unl.edu/libphilprac/1379>
- Marra, R.M. & Bogue, B. (2006). A Critical Assessment of Online Survey Tools. *Proceedings of the 2006 WEPAN Conference, WEPAN-Women in Engineering Programs and Advocates Network*.

- Marshall, B., Cardon, P., Poddar, A. & Fontenot, R. (2013). Does Sample Size Matter in Qualitative Research? A Review of Qualitative Interviews in is Research. *Journal of Computer Information Systems*, 54(1), 11-22, DOI: 10.1080/08874417.2013.11645667
- Massey, D., Johnston, Byrne, H. A., & Osborne, M. (2019). The digital age: A scoping review of nursing student's perceptions of the use of online discussion boards, *Nurse Education Today*, <https://doi.org/10.1016/j.nedt.2019.06.013>
- McIntosh, M.J & Morse, J. M. (2015). Situating and constructing diversity in semi-structured interviews. *Global qualitative nursing research*.
- Ministry of education (2019). Competence based education policy framework. <https://www.education.go.ke/index.php/downloads/file/615-competency-based-education-and-training-competency-based-training-framework>
- Ministry of education (2021). *Technical and vocational training Authority* <https://www.tveta.go.ke/vision-mission/>
- Molloy, E. & Boud, D. (2014). Feedback models for learning, teaching and performance. *Handbook of Research on Educational Communications and Technology (4th Ed.)*. Springer: New York. DOI 10.1007/978-1-4614-3185-5_33,
- National TVET Standards (2020). Kenya Report <https://www.tveta.go.ke/#>
- Negara, E. S., Triadi, D., Andryani, R (2019). Topic Modelling Twitter Data with Latent Dirichlet Allocation Method. *International Conference on Electrical Engineering and Computer Science*. DOI: 10.1109/ICECOS47637.2019.8984523
- Negash, S.(2004). Business Intelligence Communications of the Association for Information Systems, 3, 177-195
- Nair, C. S. (2011). Students' Feedback: an imperative to enhance Quality of engineering education. *International Journal of Quality Assurance in Engineering and Technology Education*, 1(1), 58-65
- Ngugi, K. (2014). Determinants influencing the growth of mobile telephony in Kenya: A case of Safaricom LTD. *International Journal of Social Sciences and Entrepreneurship*, 1 (10), 218-230.
- Ofojebe, W. N. and Ezugoh, T. C. (2010). Teachers' Motivation and its Influence on Quality Assurance in the Nigerian Educational System,(Pp. 398-417), *An International Multi-Disciplinary Journal , Ethiopia Vol. 4 (2)*
- Ojino, R., Mich, L. (2018). Mobile Applications in University Education: the case of Kenya. *Journal of e-Learning and Knowledge Society*, 14, (1), 111-125. ISSN: 1826-6223, e-ISSN: 1971-8829 DOI: 10.20368/1971-8829/1369

- Okpa, O. E. 1, Odigwe1, F. N., Emeribe, K. V. & Obule, E. V. (2020). Students' Participation in Quality Assurance Management and Tertiary Institutions Effectiveness in Nigeria. doi:10.5539/ies. v13n12p41.
- Omona, J.(2013). Sampling in Qualitative Research: Improving the Quality of Research Outcomes in Higher Education. *Makerere Journal of Higher Education ISSN: 1816-6822; 4(2) (2013) 169 – 185*. DOI: <http://dx.doi.org/10.4314/majohe.v4i2.4>
- Parahoo, S.K., Santally, M. I., Rajabalee ,Y., Harve, H.L(2015)Designing a predictive model of student satisfaction in online learning *.Journal of Marketing for Higher Education* DOI: 10.1080/08841241.2015.1083511. <http://dx.doi.org/10.1080/08841241.2015.1083511>
- Pond, W. K. (2002). Twenty-first century education and training Implications for quality assurance. *Internet and Higher Education 4 (2002) 185–192*
- Praveen, G. J. & Prasanna., K.H (2021) Sentiment Analysis:Textblob For Decision Making. *International Journal of Scientific Research & Engineering Trends, 7(2)*
- Punjab, M. B., Afzal M. T., & Azeem, M. (2008). Reliability and Validity of Qualitative and Operational Research Paradigm. *Pak.j.stat.oper.res. 4(1), 35-45*
- Rahi, S. (2017). Research Design and Methods: A Systematic Review of Research Paradigms, Sampling Issues and Instruments Development. *International Journal of Economics & Management Sciences*. DOI: 10.4172/2162-6359.1000403
- Rai N., & Thapa, B. (2015). A study on purposive sampling method in research. *Academia, accelerating the worlds' research*. 2015 - academia.edu
- Reeves, S., Kuper, A., Hodges, B. D. (2008). Qualitative Research: Qualitative Research Methodologies: Ethnography. *British Medical Journal, 337(7668), 512-514* <http://www.jstor.org/stable/20510696>
- Reigeluth, M.C. et al (2008). Roles for Technology in the Information-Age Paradigm of Education: Learning Management Systems. *Educational Technology. 48(6), 32-39*. <https://www.jstor.org/stable/44429625>
- Reigeluth, M. C. (2014). The Learner-Centered Paradigm of Education: Roles for Technology. *Educational Technology, 54(3), 18- 21*: <https://www.jstor.org/stable/44430267>
- Rocco, T. S, & Plakhotnik, M. S. (2009). Literature Reviews, Conceptual Frameworks, and Theoretical Frameworks: Terms, Functions, and Distinctions. *Human Resource Development Review, 8(1) DOI: 10.1177/1534484309332617*
- Sahu, C. (2008). An evaluation of selected pedagogical attributes of online discussion boards. *In Hello! Where are you in the landscape of educational technology? Proceedings ascilite Melbourne 2008*. <http://www.ascilite.org.au/conferences/melbourne08/procs/sahu.pdf>

- Shah, M., Cheng, M. & Fitzgerald, R. (2016). Closing the loop on student feedback: the case of Australian and Scottish Universities. *Higher Education*. 74(1), 115-129, Springer. <https://www.jstor.org/stable/26448748>
- Shah, M. Cheng, M. & Fitzgerald, R. (2017). Closing the loop on student feedback: the case of Australian and Scottish universities. *Higher Education*, 74(1) pp. 115-129, Springer. <https://www.jstor.org/stable/26448748>
- Shankararaman, V., Gottipati, S., & Gan, S. (2017). A conceptual framework for analyzing students' feedback. *FIE 2017: Proceedings of 47th Annual Frontiers in Education Conference, Indianapolis, Indiana, October 18-21, 1-8*. https://ink.library.smu.edu.sg/sis_research/3836
- Shen, J., Wu, D., Achhpiliya, V., Bieber, M., & Roxanne Hiltz (2004). Participatory Learning Approach: A Research Agenda. <http://is.njit.edu/pla/> Technical Report Version 1.0 - 6/10/04
- Sibona, C. & Walczak, S. (2012) Purposive Sampling on Twitter: A Case. 2012 25th Hawii International Conference on System sciences
- Sileyew, K.J. (2019). Text Mining - Analysis, Programming and Application. Research design and Methodology, (pp. 27). DOI: 10.5772/intechopen.85731
- Sorour,S.E.,Mine, T.,Godan, K., Hirokawa, S.(2014).A predictive model to evaluate students performance
- Sosibo, L. (2010). The views of academics on the use of student feedback for curriculum improvement. *Journal of Education*, 49, 2010 - researchgate.net
- Srungaram, V. (2017): *Predictive Modelling in Teaching and Learning*. In Christopher, C. Thompson (Eds) Handbook of Learning Analytics SOCIETY for LEARNING,1, DOI: 10.18608/hla17
- Sue, V. & Ritter L (2012) Planning the online survey. *Conducting online surveys*, 2012 - methods.sagepub.com
- Tai, J. & Ajjawi, R (2016) Undertaking and reporting qualitative research. *The Clinical Teacher's Toolbox*. <https://doi.org/10.1111/tct.12552>.
- Tammaro, A. M. (2005). Report on quality assurance models in LIS programs. *Education and Training Section*
- Taurina, Z. (2015). Motivation and Learning Outcomes: Significant Factors in Internal Study Quality Assurance System. *International Journal for Cross-Disciplinary Subjects in Education (IJCDSE), Special Issue, 5(4), 2015*
- Tenny S· Brannan G., Brannan J. & Sharts-Hopko N. (2017), Qualitative Study StatPearls Publishing, Treasure Island (FL)
- Thyer, B. A. (2010). *The handbook of social work research methods* (2nd Ed.)

DOI: <https://dx.doi.org/10.4135/9781544364902>

Treagust, D.F. (2014). Measuring Students' attitude and behavior in Technology-Rich class in Biology Classrooms. *Journal of Science and Technology*. 23 (1), pp 98-107

United Nations (2015). Sustainable Development Goals
<https://www.un.org/sustainabledevelopment/education/>

United Nations (2015). Sustainable Development Goals
<https://www.un.org/sustainabledevelopment/infrastructure-industrialization/>

VM Sue, LA Ritter - Conducting online surveys, 2012 - methods.sagepub.com

Wambui, M. K. & Gichuho, C.M. (2013). Determining Influence of Management Commitment on Development of Emotional Intelligence in Public Service - A Survey of Ministries in Kenya. *International Journal of Science and Research (IJSR)*, 2(5), India Online ISSN: 2319-7064

Wamuyu, P. K., & Maharaj, M. (2011). "Factors influencing successful use of mobile technologies to facilitate E-commerce in small enterprises: The case of Kenya," *The African Journal of Information Systems*: 3(2), Article 2:
<http://digitalcommons.kennesaw.edu/ajis/vol3/iss2/2>

Wazir, Z. K., Yang X, M., Mohammed Y. A., & Quratulain, A. Mobile Phone Sensing Systems: A Survey.

Williams, Ruth and Brennan, John (2003). Collecting and using student feedback on quality and standards of learning and teaching in Higher Education. HEFCE, Bristol, UK.

Yulherniwati, Y. & Ikhsan, A. (2020). *Assessment of Institution Readiness in Adopting Technology: A Study on Vocational Education Internal Quality Assurance System*.

APPENDICES

APPENDIX 1

A) Sample selection code

#1 .Importing the necessary libraries

```
#header in csv file
import re
import pandas as pd
import csv
import numpy as np
# display options to `max_colwidth` for better text readability:
pd.set_option ('display.max_colwidth', None)
```

2. Loading data source

```
# file path or directory
data_file ='C:/ResearchData/ k_tvet_institutions.csv'
# use engine='python to handle with error of encoding
#UnicodeDecodeError: 'utf-8' codec
# can't decode byte 0x92 in position 2: invalid start byte
df_kcolleges = pd.read_csv(data_file,engine='python')
```

3. Random sample selection

```
sample_colleges=df_kcolleges.sample(n=15)
sample_colleges.to_csv('C:/ResearchData/k_15_col_sample.CSV')
sample_colleges
```

B) DATA INGESTION INTO HDFS AND ANALYSIS

Step 1: LAUNCH THE HADOOP DAEMONS OR PROCESSES

```
#start-dfs.cmd
#start-yarn.cmd
#then run the files
```

Step 2 : import required libraries

```
import pandas as pd
import os
import re
import csv
```

```

import numpy as np
import pyspark
from pyspark.sql import SparkSession
from pyspark import SparkContext, SparkConf
from pyspark.ml.feature import VectorAssembler
from pyspark import SQLContext
%matplotlib inline
print (os.environ['SPARK_HOME'])
print (os.environ['HADOOP_HOME'])
#Step 3 : Creating a Spark session
# you can also pass spark home path to init() method like below
# findspark.init("/path/to/spark")
spark = SparkSession.builder.appName('PySpark DataFrame From External
Files').getOrCreate()
# i : building a SparkConf object that contains info about application
#conf=SparkConf().setAppName('appName').setMaster('local')
conf=SparkConf().setAppName('appName').setMaster('local')
sc=SparkContext.getOrCreate(conf=conf)
# iii) using spark context to create a new session
spark=SparkSession(sc)
spark=SparkSession.builder.appName("spark basic operations").getOrCreate()
sqlContext=SQLContext(sc)
# Step 4 : Extracting data from local directory file
# User and hadoop directory
# create a dataframe from csv in spark
path='c://ResearchData/student_feedback.csv' # path to hadoop directory
spk_data_obj_csv=spark.read.csv(path,inferSchema=True,header=True) #read files in the
directory
dfr=spark.read.option("header",True).csv(path)
dfr=sqlContext.read.csv(path,header=True)
spark=SparkSession(sc)
dfr.printSchema()
# Step 5: archiving to HDFS
file_to_hdfs='hdfs://localhost:9000/datalake/archived_student_feedback.csv'

```

```

pdf = dfr.toPandas()
spk_data_obj_csv.write.csv(file_to_hdfs)
spk_data_obj_csv

1. Importing the necessary libraries
#header in csv file

import re
import pandas as pd
import csv
import numpy as np
import pickle

# display options to `max_colwidth` for better text readability:
pd.set_option('display.max_colwidth', None)

#2. Loading data source
# file path or directory
file_path='C:/ResearchData/students_feedback.csv'
df_feedback = pd.read_csv(file_path,engine='python')
df_feedback.head(3)
df_feedback.describe()

# Get the Column names i.e the questions as column labels
# i) making a list of column_name from questions
column_names=list_questions = list(df_feedback.columns.values.tolist())
# ii) Create an array of column names
columnsNamesArr =df_feedback.columns.values
columnsNamesArr

# iii) Display the indexes of columns by removing header labels
df_feedback = pd.read_csv(file_path,header=None,engine='python')
#use engine='python to deal with error of encoding
#UnicodeDecodeError: 'utf-8' codec
# can't decode byte 0x92 in position 2: invalid start byte
df_feedback.head(2)

#Select and Rename columns
#select necessary columns as depicted by the index
rename_dfr = pd.read_csv(file_path, usecols=[2,3,4,5,6,7,8,9,10,11,12,13,14,15,17])
# View the necessary data columns

```

```

print(df_feedback.shape[1])      # Show dimension of the columns
print (range(df_feedback .shape[1])) # Show range of the columns
#renaming the columns
rename_dfr.columns =
['trainer_pos','tvet_participation','website_rating','mobile_app_rating','technology_usage','inst
_delivery_qlty', 'Instruct_derivery_eval',
'internet_norm_learn','resource_aware','resource_req', 'Assess_process_tech',
'Assess_learn_qlty','fdbck_capture_freq','contin_vs_periodic','effect_fdbk_environ']
#Display columns labels and data
#rename_dfr
# Exploration of data values
#Handling null values
#pd.isna(df_colleges)
is_NaN = rename_dfr.isnull()
row_has_NaN = is_NaN.any(axis=1)
rows_with_NaN = rename_dfr[row_has_NaN]
print(rows_with_NaN)
# Data cleaning
# view the structure of data
#remove null values
clean_data=rename_dfr.dropna()
clean_data.head(2)
# Save cleaned and renamed data to pkl or csv files
#serialize Python object structures by Pickle i.e convert an object in the memory to a byte
stream,
# for storage as a binary file on disk
#pickle module keeps track of the objects it has already serialized,
#so that later references to the same object won't be serialized again,
#thus allowing for faster execution time.
#save file with new columns using pickle
# file path or directory
file_dir='C:/ResearchData/feedback_questionnaire.csv'
rename_dfr = pd.read_csv(file_dir,engine='python')
rename_dfr.to_pickle('C:/ResearchData/feedback_questionnaire2.pkl')

```

```

#rename_dfr.to_csv('C:/ResearchData/feedback_questionnaire.csv')
rename_dfr
# read file to pandas dataframe
pkl_df=pd.read_pickle('C:/ResearchData/feedback_questionnaire2.pkl')
pkl_df
Get the VADER sentiments
from nltk.sentiment.vader import SentimentIntensityAnalyzer
# using the SentimentIntensityAnalyzer function to evaluate and determine the sentiment
analyzer = SentimentIntensityAnalyzer()
# scores are three different categories i.e. negative, neutral, and positive.
# and the compound attribute i.e a summarized result of all the three categories score.
#The value of compound ranges from -1(most extreme negative) and +1 (most extreme
positive).
#stud_comments['neg'] = stud_comments['feedback'].apply(lambda
x:sia.polarity_scores(x)['neg'])
#stud_comments['neu'] = stud_comments['feedback'].apply(lambda
x:sia.polarity_scores(x)['neu'])
#stud_comments['pos'] = stud_comments['feedback'].apply(lambda
x:sia.polarity_scores(x)['pos'])
#stud_comments['compound'] = stud_comments['feedback'].apply(lambda
x:sia.polarity_scores(x)['compound'])
def get_vader_sentiment(analyzer, opinion):
    opinion = opinion.replace('#,") # include hashtag text
    vader_scores = analyzer.polarity_scores(opinion)
    compound_score = vader_scores['compound']
    vader_sentiment = None
    # using thresholds from VADER developers/researchers
    if (compound_score >= 0.05):
        vader_sentiment = 'positive'
    elif (compound_score < 0.05 and compound_score > -0.05):
        vader_sentiment = 'neutral'
    elif (compound_score <= -0.05):
        vader_sentiment = 'negative'
    return vader_sentiment

```

```

# a) technology_usage Positivity
# i) Positivity for 'technology_usage'
pkl_df['vader_sentiment'] = pkl_df.apply(lambda row:
get_vader_sentiment(analyzer, row['technology_usage']), axis=1)
pkl_df[['technology_usage','vader_sentiment']]
# ii) technology_usage sentimental scores
from nltk.sentiment.vader import SentimentIntensityAnalyzer
# using the SentimentIntensityAnalyzer function to evaluate and determine the sentiment
s_analyzer = SentimentIntensityAnalyzer()
# scores are three different categories i.e. negative, neutral, and positive.
# and the compound attribute i.e a summarized result of all the three categories score.
#The value of compound ranges from -1(most extreme negative) and +1 (most extreme
positive).
pkl_df['neg'] = pkl_df['technology_usage'].apply(lambda
x:s_analyzer.polarity_scores(x)['neg'])
pkl_df['neu'] = pkl_df['technology_usage'].apply(lambda
x:s_analyzer.polarity_scores(x)['neu'])
pkl_df['pos'] = pkl_df['technology_usage'].apply(lambda
x:s_analyzer.polarity_scores(x)['pos'])
pkl_df['compound'] = pkl_df['technology_usage'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
# overall sentiment score of texts by average of compound attribute
vader_neg_tech_usage = pkl_df['neg'].mean()
vader_neu_tech_usage= pkl_df['neu'].mean()
vader_pos_tech_usage = pkl_df['pos'] .mean()
vader_comp_tech_usage = pkl_df['compound'].mean()
#Results of Sentiment Analysis
print("The negative polarity of technology_usage variable is:",vader_neg_tech_usage)
print("The neutral polarity of technology_usage variable is:",vader_comp_tech_usage )
print("The positivs polarity of technology_usage variable is:",vader_neu_tech_usage )
print("The average polarity of technology_usage variable is:",vader_comp_tech_usage )
pkl_df.head()
# b) Instruct_derivery_evaluation : Positivity
# i) Positivity rating for 'Instruct_derivery_eval'

```

```

pkl_df['vader_sentiment'] = pkl_df.apply(lambda row:
get_vader_sentiment(analyzer, row['Instruct_derivery_eval']), axis=1)
pkl_df[['Instruct_derivery_eval','vader_sentiment']]
# ii) Instruct_derivery_evaluation sentimental scores
from nltk.sentiment.vader import SentimentIntensityAnalyzer
# using the SentimentIntensityAnalyzer function to evaluate and determine the sentiment
s_analyzer = SentimentIntensityAnalyzer()
# scores are three different categories i.e. negative, neutral, and positive.
# and the compound attribute i.e a summarized result of all the three categories score.
#The value of compound ranges from -1(most extreme negative) and +1 (most extreme
positive).
pkl_df['neg'] = pkl_df['Instruct_derivery_eval'].apply(lambda
x:s_analyzer.polarity_scores(x)['neg'])
pkl_df['neu'] = pkl_df['Instruct_derivery_eval'].apply(lambda
x:s_analyzer.polarity_scores(x)['neu'])
pkl_df['pos'] = pkl_df['Instruct_derivery_eval'].apply(lambda
x:s_analyzer.polarity_scores(x)['pos'])
pkl_df['compound'] = pkl_df['Instruct_derivery_eval'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
# overall sentiment score of texts by average of compound attribute
vader_neg_Instr_deriv = pkl_df['neg'].mean()
vader_neu_Instr_deriv = pkl_df['neu'].mean()
vader_pos_Instr_deriv = pkl_df['pos'] .mean()
vader_comp_Instr_deriv = pkl_df['compound'].mean()
#Results of Sentiment Analysis
print("The negative polarity of Instruct_derivery_eval variable is:",vader_neg_Instr_deriv )
print("The neutral polarity of Instruct_derivery_eval variable is:",vader_comp_Instr_deriv )
print("The positivs polarity of Instruct_derivery_eval variable is:",vader_neu_Instr_deriv )
print("The average polarity of Instruct_derivery_eval variable is:",vader_comp_Instr_deriv )
pkl_df.head()
# c) resource_provisioning : Positivity
# i) Positivity rating for 'resource_provisioning'
pkl_df['vader_sentiment'] = pkl_df.apply(lambda row:
get_vader_sentiment(analyzer, row['resource_req']), axis=1)

```

```

pkl_df.head(3)
pkl_df['vader_sentiment']
pkl_df[['resource_req','vader_sentiment']]
# ii) resource_provisioning sentimental scores
from nltk.sentiment.vader import SentimentIntensityAnalyzer
# using the SentimentIntensityAnalyzer function to evaluate and determine the sentiment
s_analyzer = SentimentIntensityAnalyzer()
# scores are three different categories i.e. negative, neutral, and positive.
# and the compound attribute i.e a summarized result of all the three categories score.
#The value of compound ranges from -1(most extreme negative) and +1 (most extreme
positive).
pkl_df['neg'] = pkl_df['resource_req'].apply(lambda x:s_analyzer.polarity_scores(x)['neg'])
pkl_df['neu'] = pkl_df['resource_req'].apply(lambda x:s_analyzer.polarity_scores(x)['neu'])
pkl_df['pos'] = pkl_df['resource_req'].apply(lambda x:s_analyzer.polarity_scores(x)['pos'])
pkl_df['compound'] = pkl_df['resource_req'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
# overall sentiment score of texts by average of compound attribute
vader_neg_resource_req = pkl_df['neg'].mean()
vader_neu_resource_req = pkl_df['neu'].mean()
vader_pos_resource_req = pkl_df['pos'] .mean()
vader_comp_resource_req = pkl_df['compound'].mean()
#Results of Sentiment Analysis
print("The negative polarity of resource_provisioning variable is:",vader_neg_resource_req )
print("The neutral polarity of resource_provisioning variable is:",vader_neu_resource_req)
print("The positivs polarity of resource_provisioning variable is:",vader_pos_resource_req )
print("The average polarity of resource_provisioningt variable is:",vader_comp_resource_req
)
pkl_df.head()
# d) asesment process : Positivity
# i) Positivity rating for 'Assessment process'
pkl_df['vader_sentiment'] = pkl_df.apply(lambda row:
get_vader_sentiment(analyzer, row['Assess_learn_qlty']), axis=1)
pkl_df.head(3)
pkl_df['vader_sentiment']

```

```

pkl_df[['Assess_learn_qlty','vader_sentiment']]
# ii) Assessment process sentimental scores
from nltk.sentiment.vader import SentimentIntensityAnalyzer
# using the SentimentIntensityAnalyzer function to evaluate and determine the sentiment
s_analyzer = SentimentIntensityAnalyzer()
# scores are three different categories i.e. negative, neutral, and positive.
# and the compound attribute i.e a summarized result of all the three categories score.
#The value of compound ranges from -1(most extreme negative) and +1 (most extreme
positive).
pkl_df['neg'] = pkl_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['neg'])
pkl_df['neu'] = pkl_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['neu'])
pkl_df['pos'] = pkl_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['pos'])
pkl_df['compound'] = pkl_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
# overall sentiment score of texts by average of compound attribute
vader_neg_Assess_learn_qlty= pkl_df['neg'].mean()
vader_neu_Assess_learn_qlty= pkl_df['neu'].mean()
vader_pos_Assess_learn_qlty = pkl_df['pos'] .mean()
vader_comp_Assess_learn_qlty = pkl_df['compound'].mean()

#Results of Sentiment Analysis
print("The negative polarity of Assessment process variable
is:",vader_neg_Assess_learn_qlty )
print("The neutral polarity of Assessment process variable
is:",vader_neu_Assess_learn_qlty)
print("The positivs polarity of Assessment process variable
is:",vader_pos_Assess_learn_qlty )
print("The average polarity of Assessment process variable
is:",vader_comp_Assess_learn_qlty)
pkl_df.head()

```

e) Feedback frequency :sentimental positivity

```
# i) Positivity rating for mediator variable 'Assessment frequency'
pkl_df['vader_sentiment'] = pkl_df.apply(lambda row:
get_vader_sentiment(analyzer, row['contin_vs_periodic']), axis=1)
pkl_df.head(3)
pkl_df['vader_sentiment']
pkl_df[['contin_vs_periodic','vader_sentiment']]
# ii) feedback frequency sentimental scores
from nltk.sentiment.vader import SentimentIntensityAnalyzer
# using the SentimentIntensityAnalyzer function to evaluate and determine the sentiment
s_analyzer = SentimentIntensityAnalyzer()
# scores are three different categories i.e. negative, neutral, and positive.
# and the compound attribute i.e a summarized result of all the three categories score.
#The value of compound ranges from -1(most extreme negative) and +1 (most extreme
positive).
pkl_df['neg'] = pkl_df['contin_vs_periodic'].apply(lambda
x:s_analyzer.polarity_scores(x)['neg'])
pkl_df['neu'] = pkl_df['contin_vs_periodic'].apply(lambda
x:s_analyzer.polarity_scores(x)['neu'])
pkl_df['pos'] = pkl_df['contin_vs_periodic'].apply(lambda
x:s_analyzer.polarity_scores(x)['pos'])
pkl_df['compound'] = pkl_df['contin_vs_periodic'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
# overall sentiment score of texts by average of compound attribute
vader_neg_contin_vs_periodic= pkl_df['neg'].mean()
vader_neu_contin_vs_periodic= pkl_df['neu'].mean()
vader_pos_contin_vs_periodic = pkl_df['pos'] .mean()
vader_comp_contin_vs_periodic = pkl_df['compound'].mean()
#Results of Sentiment Analysis
print("The negative polarity of Continuous feedback rated against periodic
is:",vader_neg_contin_vs_periodic )
print("The neutral polarity of Continuous feedback rated against
periodics:",vader_neu_contin_vs_periodic)
```

```

print("The positive polarity of Continuous feedback rated against
periodic", vader_pos_contin_vs_periodic )
print("The average polarity of Continuous feedback rated against
periodic:", vader_comp_contin_vs_periodic)
pkl_df.head()

```

f) feedback on Overall environment :sentimental positivity

i) Positivity rating for dependent variable 'overall environment'

```

pkl_df['vader_sentiment'] = pkl_df.apply(lambda row:
get_vader_sentiment(analyzer, row['effect_fdbk_enviro']), axis=1)
pkl_df.head(3)

```

```

pkl_df['vader_sentiment']

```

```

pkl_df[['effect_fdbk_enviro','vader_sentiment']]

```

ii) feedback on Overall environmental scores

```

from nltk.sentiment.vader import SentimentIntensityAnalyzer

```

using the SentimentIntensityAnalyzer function to evaluate and determine the sentiment

```

s_analyzer = SentimentIntensityAnalyzer()

```

scores are three different categories i.e. negative, neutral, and positive.

and the compound attribute i.e a summarized result of all the three categories score.

#The value of compound ranges from -1(most extreme negative) and +1 (most extreme positive).

```

pkl_df['neg'] = pkl_df['effect_fdbk_enviro'].apply(lambda
x:s_analyzer.polarity_scores(x)['neg'])

```

```

pkl_df['neu'] = pkl_df['effect_fdbk_enviro'].apply(lambda
x:s_analyzer.polarity_scores(x)['neu'])

```

```

pkl_df['pos'] = pkl_df['effect_fdbk_enviro'].apply(lambda
x:s_analyzer.polarity_scores(x)['pos'])

```

```

pkl_df['compound'] = pkl_df['effect_fdbk_enviro'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])

```

overall sentiment score of texts by average of compound attribute

```

vader_neg_effect_fdbk_enviro= pkl_df['neg'].mean()

```

```

vader_neu_effect_fdbk_enviro= pkl_df['neu'].mean()

```

```

vader_pos_effect_fdbk_enviro = pkl_df['pos'] .mean()

```

```

vader_comp_effect_fdbk_enviro = pkl_df['compound'].mean()

```

```

#Results of Sentiment Analysis
print("The negative polarity of feednack on Overall environmental variable
is:",vader_neg_effect_fdbk_environ)
print("The neutral polarity of feednack on Overall environmental variable
is:",vader_neu_effect_fdbk_environ)
print("The positivs polarity of feednack on Overall environmental variable
is:",vader_pos_effect_fdbk_environ )
print("The average polarity of feednack on Overall environmental variable
is:",vader_comp_effect_fdbk_environ)
pkl_df.head()

```

Visualization of positivity rating

```

# The Dependent variables
vader_comp_tech_usage=0.37127
vader_comp_resource_req=0.1577
vader_comp_Assess_learn_qlty=0.27950
vader_comp_Instr_deriv=0.308432
import matplotlib.pyplot as plt
import numpy
#labels = ['NLTK Vader','TextBlob','Pattern']
labels = ['Instructional delivery','Technology usage','Resource provisioning','Assessment
process']
sentiment_count =
[vader_comp_Instr_deriv,vader_comp_tech_usage,vader_comp_resource_req,vader_comp_A
ssess_learn_qlty]
x = np.arange(len(labels)) # the label locations
width = 0.35 # the width of the bars
#In the below code we have specified the labels for both the axes and also the data
represented by each axis. Along with this, autolabel function will help in displaying the count
of each bar over the top.
#The tight_layout automatically adjusts subplot params so that the subplot(s) fits into the
figure area.

```

#As it can be inferred from the visualization, NLTK Vader has given higher scores as compared to the other two libraries. Both TextBlob and Pattern libraries have almost identical results.

#1.Instructional delivery

#2.Technology usage.

#3.Resource provisioning

#4.Assessment process

#Feedback Capture frequency

```
#fig = plt.figure()
```

```
#fig.set_figwidth(20)
```

```
#f.set_figheight(1)
```

```
fig, ax = plt.subplots()
```

```
#rects = ax.bar(x, sentiment_count, width, label='Dependent Variables')
```

```
bars = ax.bar(x, sentiment_count, width, label='Independent Variables')
```

```
ax.set_ylabel('Polarity_Score')
```

```
#plotting the rating
```

```
fig.set_figwidth(15)
```

```
ax.set_ylabel('Polarity_Score')
```

```
ax.set_title('Comparison of feedback Variables')
```

```
ax.set_xticks(x)
```

```
ax.set_xticklabels(labels)
```

```
ax.legend()
```

```
def autolabel(bars):
```

```
    """Attach a text label above each bar in *rects*, displaying its height."""
```

```
    for scorebar in bars:
```

```
        height = scorebar.get_height()
```

```
        ax.annotate('{}' .format(height),
```

```
                xy=(scorebar.get_x() + scorebar.get_width() / 2, height),
```

```
                xytext=(0, 3), # 3 points vertical offset
```

```
                textcoords="offset points",
```

```
                ha='center', va='bottom')
```

```
autolabel(bars)
```

```
# fitting subplot parameters into the figure area by tight_layout.
```

```

#fig.tight_layout()
#plt.show()

# g) Calculating the mean for independent variables
# all independent variables
indepentent_var= [vader_comp_tech_usage, vader_comp_Instr_deriv, vader_comp_res_prov,
vader_comp_assess_proc]
#calculating mean
sum_rating = sum(indepentent_var)
combined_rating_variables=sum_rating/4
print(combined_rating_variables)

#3 Descriptive statistics of respondents
# read file to pandas dataframe
pkl_df=pd.read_pickle('C:/ResearchData/feedback_questionnaire.pkl')
pkl_df.head()
pkl_df=pd.read_csv('C:/ResearchData/feedback_questionnaire.csv',engine='python')
pkl_df.head()
# i) Exposure to TVET programmes
print("Number of Respondents by participation in TVET programmes",sum(pkl_df
['tvvet_participation'] == 'YES'))
print("Respondents without sufficient exposure in TVET programmes ",sum(pkl_df
['tvvet_participation'] == 'NO'))
# Summarry 'tvvet_participation'
frequency = pkl_df ['tvvet_participation'].value_counts(normalize =True)
frequency =frequency * 100
frequency.round()
# ii) Descriptive statistics of trainers by column 'trainer_pos'
Freq_trainer_pos = pkl_df ['trainer_pos'].value_counts(normalize =True)
Freq_trainer_pos =frequency * 100
Freq_trainer_pos.round()
print("Partiscipants Description",f Freq_trainer_pos.round() )

```

4. Website Rating

```
print("Number of Participants for website_rating:very important",sum(pk1_df['inst_delivery_qlty'] == 1))
print("Number of Participants for website_rating:important",sum(pk1_df ['inst_delivery_qlty']
== 2))
print("Number of Participants for website_rating:Moderatly
important",sum(pk1_df['inst_delivery_qlty'] == 3))
print("Number of Participants for website_rating:slightly
important",sum(pk1_df['inst_delivery_qlty'] == 4))
print("Number of Participants for website_rating:umimpoertant",sum(pk1_df
['inst_delivery_qlty'] == 5))
print("\nRating % website Rating for feedback")
frequency = pk1_df ['inst_delivery_qlty'].value_counts(normalize =True)
mobile_app_rating =frequency * 100
mobile_app_rating.round()
mobile_app_rating
```

5 Mobile application Rating

```
print("Number of Participants for mobile_rating:very important",sum(pk1_df['mobile_app_rating'] == 1))
print("Number of Participants for mobile_rating:important",sum(pk1_df ['mobile_app_rating']
== 2))
print("Number of Participants for mobile_rating:Moderatly
important",sum(pk1_df['mobile_app_rating'] == 3))
print("Number of Participants for mobile_rating:slightly
important",sum(pk1_df['mobile_app_rating'] == 4))
print("Number of Participants for mobile_rating:slightly
important",sum(pk1_df['mobile_app_rating'] == 5))
print("\nRating % Mobile application Rating for feedback")
frequency = pk1_df ['mobile_app_rating'].value_counts(normalize =True)
mobile_app_rating =frequency * 100
mobile_app_rating.round()
mobile_app_rating
```

6. internet in normal learning Rating

```
print("Number of Participants for internet in normal learning Rating: Very
Important",sum(pk1_df['internet_norm_learn'] == 1))
print("Number of Participants for internet in normal learning
Rating:Important",sum(pk1_df['internet_norm_learn'] == 2))
print("Number of Participants for internet in normal learning Rating:Moderately
Importantt",sum(pk1_df['internet_norm_learn'] == 3))
print("Number of Participants for internet in normal learning Rating Rating:slightly
Importantt",sum(pk1_df['internet_norm_learn'] == 4))
print("Number of Participants for internet in normal learning Rating
Rating:unImportant",sum(pk1_df['internet_norm_learn'] == 5))
print("\nRating % internet in normal learning")
frequency = pk1_df ['internet_norm_learn'].value_counts(normalize =True)
inst_delivery_qlty_rating =frequency * 100
inst_delivery_qlty_rating
```

7. Instructional delivery on quality Rating

```
print("Number of Participants for 'inst_delivery_qlty
efficient",sum(pk1_df['inst_delivery_qlty'] == 1))
print("Number of Participants for Mobile application
Rating:Efficient",sum(pk1_df['inst_delivery_qlty'] == 2))
print("Number of Participants for Mobile application Rating:Moderately
efficient",sum(pk1_df['inst_delivery_qlty'] == 3))
print("Number of Participants for Mobile application Rating:slightly
efficient",sum(pk1_df['inst_delivery_qlty'] == 4))
print("Number of Participants for Mobile application
Rating:inefficient",sum(pk1_df['inst_delivery_qlty'] == 5))
print("\nRating % inst_delivery_qlty Rating for feedback")
frequency = pk1_df ['inst_delivery_qlty'].value_counts(normalize =True)
inst_delivery_qlty_rating =frequency * 100
inst_delivery_qlty_rating
```

8. students resource awareness

```
print("Number of Respondents on resource awareness: Strongly Agree",sum(pk1_df
['resource_aware'] == 'Strongly Agree'))
print("Number of Respondents on resource awareness: Agree",sum(pk1_df ['resource_aware']
== 'Agree'))
print("Number of Respondents on resource awareness: Neutral",sum(pk1_df ['resource_aware']
== 'Neutral'))
print("Number of Respondents on resource awareness :Disagree",sum(pk1_df
['resource_aware'] == 'Disagree'))
print("Number of Respondents on resource awareness :Strongly Disagree",sum(pk1_df
['resource_aware'] == 'Strongly Disagree'))
print("\nRating % internet resource awareness")
frequency = pk1_df['resource_aware'].value_counts(normalize =True)
resource_aware_rating =frequency * 100
resource_aware_rating
```

9. Technology requirement on assessment process

```
print("Number of Respondents on technology requirement on assessment process: Strongly
Agree",sum(pk1_df ['Assess_process_tech'] == 'Strongly Agree'))
print("Number of Respondents on technology requirement on assessment process ;
Agree",sum(pk1_df ['Assess_process_tech'] == 'Strongly Agree'))
print("Number of Respondents on technology requirement on assessment process:
Neutral",sum(pk1_df ['Assess_process_tech'] == 'Neutral'))
print("Number of Respondents on technology requirement on assessment process
:Disagree",sum(pk1_df ['Assess_process_tech'] == 'Disagree'))
print("Number of Respondents on technology requirement on assessment process :Strongly
Disagree",sum(pk1_df ['Assess_process_tech'] == 'Strongly Disagree'))
print("\nvote % technology requirement on assessment process")
frequency = pk1_df['Assess_process_tech'].value_counts(normalize =True)
Assess_process_tech =frequency * 100
Assess_process_tech
```

10. feedback capture frequency

```
print("Number of Respondents on feedback capture frequency: Always",sum(pk1_df
['fdbck_capture_freq'] == 'Always'))
```

```

print("Number of Respondents on capture frequencys : Often",sum(pk1_df
['fdbck_capture_freq'] == 'Often'))
print("Number of Respondents on capture frequency: Sometimes",sum(pk1_df
['fdbck_capture_freq'] == 'Sometimes'))
print("Number of Respondents on capture frequency :Rarely",sum(pk1_df
['fdbck_capture_freq'] == 'Rarely'))
print("Number of Respondents on capture frequency :Never",sum(pk1_df
['fdbck_capture_freq'] == 'Never'))
print("\nvote % capture frequency")
frequency = pk1_df['fdbck_capture_freq'].value_counts(normalize =True)
fdbck_capture_freq =frequency * 100
fdbck_capture_freq

```

11. Calculating the service rating as a measure of IQA

```

# Getting the average of independent variables
import numpy
labels = ['Instructional delivery','Technology usage','Resource provisioning','Assessment
process']
sentiment_count =
[vader_comp_Instr_deriv,vader_comp_tech_usage,vader_comp_resource_req,vader_comp_A
ssess_learn_qlty]
lst_avg = numpy.average(sentiment_count)
print(lst_avg)

```

12. Feedback- IQA multivariety data from all participants

```

#Parallel coordinates for feedback-IQA multivariate data.
import pandas as pd
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.datasets import load_iris, load_boston, load_wine
from sklearn.preprocessing import MinMaxScaler
import plotly.express as px
import plotly.graph_objects as go

```

```

%matplotlib inline
read file to pandas dataframe
pk1_df=pd.read_pickle('C:/ResearchData/feedback_questionnaire5.pkl')
# name the unnamed index col to : ID
pk1_df.insert(loc=0, column='ID', value=np.arange(len(pk1_df)))
# ii) technology_usage sentimental scores
from nltk.sentiment.vader import SentimentIntensityAnalyzer
# using the SentimentIntensityAnalyzer function to evaluate and determine the sentiment
s_analyzer = SentimentIntensityAnalyzer()
# scores are three different categories i.e. negative, neutral, and positive.
# and the compound attribute i.e a summarized result of all the three categories score.
#The value of compound ranges from -1(most extreme negative) and +1 (most extreme
positive).
pk1_df['neg'] = pk1_df['techn_usage'].apply(lambda x:s_analyzer.polarity_scores(x)['neg'])
pk1_df['neu'] = pk1_df['techn_usage'].apply(lambda x:s_analyzer.polarity_scores(x)['neu'])
pk1_df['pos'] = pk1_df['techn_usage'].apply(lambda x:s_analyzer.polarity_scores(x)['pos'])
#pk1_df['compound_tech'] = pk1_df['technology_usage'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
pk1_df['Use_of_techn'] = pk1_df['techn_usage'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])

# ii) Instruct_derivery_evaluation sentimental scores
pk1_df['neg'] = pk1_df['inst_deliv_eval'].apply(lambda x:s_analyzer.polarity_scores(x)['neg'])
pk1_df['neu'] = pk1_df['inst_deliv_eval'].apply(lambda x:s_analyzer.polarity_scores(x)['neu'])
pk1_df['pos'] = pk1_df['inst_deliv_eval'].apply(lambda x:s_analyzer.polarity_scores(x)['pos'])
#pk1_df['compound_Instruct'] = pk1_df['Instruct_derivery_eval'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
pk1_df['Inst_deliverly'] = pk1_df['inst_deliv_eval'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])

# iii) resource_provisioning sentimental scores
pk1_df['neg'] = pk1_df['resource_req'].apply(lambda x:s_analyzer.polarity_scores(x)['neg'])
pk1_df['neu'] = pk1_df['resource_req'].apply(lambda x:s_analyzer.polarity_scores(x)['neu'])
pk1_df['pos'] = pk1_df['resource_req'].apply(lambda x:s_analyzer.polarity_scores(x)['pos'])

```

```

#pk1_df['compound_resource'] = pk1_df['resource_req'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
pk1_df['resource_req'] = pk1_df['resource_req'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])

# iv)Assessment process sentimental scores
pk1_df['neg'] = pk1_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['neg'])
pk1_df['neu'] = pk1_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['neu'])
pk1_df['pos'] = pk1_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['pos'])
#pk1_df['compound_assess'] = pk1_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
pk1_df['Assess_process'] = pk1_df['Assess_learn_qlty'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])

# v) frequency of feedback (contin_vs_periodic) sentimental) scores
pk1_df['neg'] = pk1_df['contin_vs_period'].apply(lambda
x:s_analyzer.polarity_scores(x)['neg'])
pk1_df['neu'] = pk1_df['contin_vs_period'].apply(lambda
x:s_analyzer.polarity_scores(x)['neu'])
pk1_df['pos'] = pk1_df['contin_vs_period'].apply(lambda
x:s_analyzer.polarity_scores(x)['pos'])
#pk1_df['compound_cont'] = pk1_df['contin_vs_periodic'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
pk1_df['Feedback_freq'] = pk1_df['contin_vs_period'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])

# vi) effect on overall learning enviromnwt sentimental scores
pk1_df['neg'] = pk1_df['ef_fdbk_environ'].apply(lambda
x:s_analyzer.polarity_scores(x)['neg'])

```

```

pkl_df['neu'] = pkl_df['ef_fdbk_environ'].apply(lambda
x:s_analyzer.polarity_scores(x)['neu'])
pkl_df['pos'] = pkl_df['ef_fdbk_environ'].apply(lambda
x:s_analyzer.polarity_scores(x)['pos'])
#pkl_df['compound_effect_environ'] =
pkl_df['compound_Instruct'effect_fdbk_environ'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
pkl_df['fdbck_on_environ'] = pkl_df['ef_fdbk_environ'].apply(lambda
x:s_analyzer.polarity_scores(x)['compound'])
pkl_df=pkl_df[['ID','Use_of_tech'n','Inst_deliverly','resource_req','Assess_process','Feedback_
freq','fdbck_on_environ']]
pkl_df

```

13. Model Visualization for students'Feedback- IQA

```

cols =
['ID','Use_of_tech'n','Inst_deliverly','resource_req','Assess_process','Feedback_freq','fdbck_on
_environ']
fig = px.parallel_coordinates(pkl_df, color="ID", dimensions=cols,
                             color_continuous_scale=px.colors.diverging.Tealrose,
                             title="Students feedback Parallel Coorinate Plot")
# saving this parallel coordinate Visualization as a html file:
fig.write_html("C:/ResearchData/students_feedback_parallel-coordinate-plot-plotly.html")
#pd.plotting.parallel_coordinates(pkl_df,
fig.show()

```

#14. Visualization of comments positivity for variables using Lineplot

```

import matplotlib.pyplot as plt
plt.xlabel("comments ID")
plt.ylabel("positivity Rating")
#pkl_df['mean_rating_fd'].rolling(window =20).mean().plot()
x=pkl_df['ID']
y=pkl_df['Assess_process'].rolling(window =20).mean().plot.line(figsize=(13,7),x ='ID')
y=pkl_df['Inst_deliverly'].rolling(window =20).mean().plot.line(figsize=(13,7),x ='ID')
y=pkl_df['Use_of_tech'n'].rolling(window =20).mean().plot.line(figsize=(13,7),x ='ID')

```

```
y=pkl_df['resource_req'].rolling(window =20).mean().plot.line(figsize=(13,7),x ='ID')
```

```
#plt.plot(x, y)
```

15 Testing and Validation

Comparison of VADER and TextBlob

```
# Get polarity function
```

```
# i)function in using textBlob
```

```
def get_polarity(text):
```

```
    return TextBlob(text).sentiment.polarity
```

```
# ii)function in using VADER
```

```
#Importing analyzer
```

```
from nltk.sentiment.vader import SentimentIntensityAnalyzer
```

```
# evaluate and determine the sentiments
```

```
analyzer = SentimentIntensityAnalyzer()
```

```
def get_vader_sentiment(analyzer, opinion):
```

```
    opinion = opinion.replace('#,") # include hashtag text
```

```
    vader_scores = analyzer.polarity_scores(opinion)
```

```
    compound_score = vader_scores['compound']
```

```
    vader_sentiment = None
```

```
    # using thresholds from VADER developers/researchers
```

```
    if (compound_score >= 0.05):
```

```
        vader_sentiment = 'positive'
```

```
    elif (compound_score < 0.05 and compound_score > -0.05):
```

```
        vader_sentiment = 'neutral'
```

```
    elif (compound_score <= -0.05):
```

```
        vader_sentiment = 'negative'
```

```
    return vader_sentiment
```

```
# iii) feedback polarity scores on combined text from independent variables
```

```
pkl_df["combined_Sentiment"] =
```

```
pkl_df[["techn_usage", "inst_deliv_eval", "resource_req", "Assess_learn_qlty"]].agg('-',join,
```

```
axis=1)
```

```

# darray=pk1_df["combined_Sentiment"].values
#pk1_df['combined_Sentiment'] = pk1_df['combined_Sentiment'].astype('str')
pk1_df['combined_Sentiment'] = pk1_df['combined_Sentiment'].apply(get_polarity)
tblob_polarity=pk1_df['combined_Sentiment'].mean()
x = round(tblob_polarity, 2)
print(x)
# converting column values to list for independent variables'
tech_list= pk1_df["techn_usage"].tolist()
inst_list= pk1_df["inst_deliv_eval"].tolist()
res_list= pk1_df["resource_req"].tolist()
asses_list= pk1_df["Assess_learn_qlty"].tolist()

# v converting list to Text
str_tech=listToString(tech_list)
str_inst=listToString(inst_list)
str_res=listToString(res_list)
str_asses=listToString(asses_list)

# b) TextBlob Polarity scores for Independent variables
#applying the TextBlob function
tech_text_p = TextBlob(str_tech)
inst_text_p = TextBlob(str_inst)
res_text_p = TextBlob(str_res)
asses_text_p = TextBlob(str_asses)
# Getting polarity scores
tech_p=tech_text_p.sentiment
inst_p=inst_text_p.sentiment
res_p=res_text_p.sentiment
asses_p=asses_text_p.sentiment
print("technology\n",tech_p)
print("instruction\n",inst_p)
print("resource\n",res_p)
print("assessment\n",asses_p)

```

```

# c) Service rating using TextBlob
technology_p=0.3878888888888889
instruction_p=0.41343621399176955
resource_p=0.34966666666666666
assessment_p=0.2879338842975207
comb_p=technology_p+instruction_p+resource_p+assessment_p
comb_p_round=round(comb_p,2)
service_rating=IQA=comb_p_round/4
print("service rating index:",service_rating)
# print(str_tech)
# print(inst_list)
# print(str_res)
# print(str_asses)
# iv) Getting polarity for combined text by VADER
combined_var_text_scores = s_analyzer.polarity_scores(combined_text)
print(str(combined_var_text_scores))

```

16. Linear and Likert Scale

```

# Textblob polarity scores Likert and Linear Scales
text_vi = TextBlob("very important")
text_he = TextBlob("highly efficient")
text_sa = TextBlob("strongly Agree")
# Getting polarity for text by VADER
Vad_text_vi_score = s_analyzer.polarity_scores("very important")
print(str(Vad_text_vi_score))
Vad_text_he_score = s_analyzer.polarity_scores("highly efficient")
print(str(Vad_text_he_score))
Vad_text_sa_score = s_analyzer.polarity_scores("strongly Agree")
print(str(Vad_text_sa_score))
print("polarity score by textBlob for 'very important'",text_vi.sentiment)
print("polarity score by textBlob for 'highly efficient'",text_he.sentiment)
print("polarity score by textBlob for 'strongly Agree\n'",text_sa.sentiment)

print("polarity score by VADER for 'very important'",Vad_text_vi_score)

```

```

print("polarity score by VADER for 'highly efficient'",text_he_score)
print("polarity score by VADER for 'strongly Agree'",text_sa_score)
text_im = TextBlob("important")
text_ef = TextBlob("efficient")
text_ag = TextBlob("agree")
# Getting polarity for text by VADER
Vad_text_im_score = s_analyzer.polarity_scores("important")
print(str(Vad_text_im_score))
Vad_text_ef_score = s_analyzer.polarity_scores("efficient")
print(str(Vad_text_ef_score))
Vad_text_ag_score = s_analyzer.polarity_scores("agree")
print(str(Vad_text_ag_score))
print("polarity score by textBlob for 'important'",text_im.sentiment)
print("polarity score by textBlob for 'efficient'",text_ef.sentiment)
print("polarity score by textBlob for 'agree'",text_ag.sentiment)
print("polarity score by VADER for 'very important'",Vad_text_im_score)
print("polarity score by VADER for 'highly efficient'",Vad_text_ef_score)
print("polarity score by VADER for 'strongly Agree'",Vad_text_ag_score)
# Getting polarity for text by VADER

text_mi = TextBlob("moderately important")
text_me = TextBlob("moderately efficient")
text_nu = TextBlob("neutral")

# Getting polarity for text by VADER
Vad_text_mi_score = s_analyzer.polarity_scores("moderately important")
print(str(Vad_text_mi_score))

Vad_text_me_score = s_analyzer.polarity_scores("moderately efficient")
print(str(Vad_text_me_score))
Vad_text_nu_score = s_analyzer.polarity_scores("neutral")
print(str(Vad_text_nu_score))

print("polarity score by textBlob for 'moderately important'",text_mi.sentiment)
print("polarity score by textBlob for 'moderately efficient'",text_me.sentiment)
print("polarity score by textBlob for 'neutral'",text_nu.sentiment)

print("polarity score by VADER for 'moderately important'",Vad_text_mi_score)
print("polarity score by VADER for 'moderately efficient'",Vad_text_me_score)
print("polarity score by VADER for 'neutral'",Vad_text_nu_score)

```

```
# Getting polarity for text by VADER
```

```
text_si = TextBlob("slightly important")  
text_se = TextBlob("slightly efficient")  
text_di = TextBlob("diasgree")
```

```
#getting polarity for column text VADER
```

```
Vad_text_si_score = s_analyzer.polarity_scores("slightly important")  
print(str(Vad_text_si_score))  
Vad_text_se_score = s_analyzer.polarity_scores("slightly efficient")  
print(str(Vad_text_se_score))
```

```
Vad_text_di_score = s_analyzer.polarity_scores("disagree")  
print(str(Vad_text_di_score))  
print("polarity score by textBlob for 'moderately important'",text_si.sentiment)  
print("polarity score by textBlob for 'moderately efficient'",text_se.sentiment)  
print("polarity score by textBlob for 'neutral'",text_di.sentiment)
```

```
print("polarity score by VADER for 'slightly important'",Vad_text_si_score)  
print("polarity score by VADER for 'slightly efficient'",Vad_text_se_score)  
print("polarity score by VADER for 'disagree'",Vad_text_di_score)
```

```
# Getting polarity for text by VADER
```

```
text_ui = TextBlob("unimportant")  
text_ie = TextBlob("inefficient")  
text_sd = TextBlob("strongly diasgree")  
#getting polarity for column text VADER
```

```
Vad_text_ui_score = s_analyzer.polarity_scores("unimportant")  
print(str(Vad_text_un_score))
```

```
Vad_text_in_score = s_analyzer.polarity_scores("inefficient")  
print(str(Vad_text_in_score))
```

```
Vad_text_sd_score = s_analyzer.polarity_scores("strongly disagree")  
print(str(Vad_text_sd_score))
```

```
print("polarity score by textBlob for 'unimportant'",text_ui.sentiment)  
print("polarity score by textBlob for 'inefficient'",text_ie.sentiment)  
print("polarity score by textBlob for 'strongly diasgree'",text_sd.sentiment)
```

```
print("polarity score by VADER for 'unimportant'",Vad_text_un_score)  
print("polarity score by VADER for 'moderately efficient'",Vad_text_in_score)  
print("polarity score by VADER for 'neutral'",Vad_text_sd_score)
```

APPENDIX 2

Research Questionnaire

REQUEST FOR COMPLETION OF QUESTIONNAIRE

This Questionnaire is for the fulfillment of my Dissertation in Master's Degree in Data Analytics (MDA) at KCA University. The research is geared to identifying the requirements for students' feedback in Competence Based Education Training (CBET) in Technical colleges in Kiambu County. This will aid in the development of a data model for capture, storage and analysis for Internal Quality Assurance (IQA).

The information provided will only be used for academic purpose and your Identity will be separated from the information for your security. Please accord me your variable assistance on the actual status in your college in any aspect of training as requested below: (*Select one option or Enter as appropriate*)

Section A: Background Information

1. Which county is the College situated? _____
2. Enter your position in the institution: _____
3. Have you participated in Competence Based Education Training in either workshop or actual training? i) YES.... ii) NO.....

Section B: Study Variables Measurements

4. How do you rate a Website in capturing students' feedback in a college? (*tick only one*)
i) Highly efficient.. ii) Efficient.... iii) Moderately efficient.... iv) Slightly efficient...
v) Inefficient....
5. How do you rate the efficiency of mobile applications in provision of feedback from students? i) Highly efficient..ii) Efficient.. iii) Moderately efficient.. iv) Slightly efficient... v) Inefficient...

What is your opinion on the use of Technology in monitoring implementation of Competence Based Education Training (CBET)? _____

Instructional delivery

7. How do you rate Instructional delivery in the quality of learning? (*tick only one*)
i) Very important... ii) Important...iii) Moderately important...iv) Slightly important...v) Unimportant...
 8. What is your opinion of students' feedback in evaluation of instructional delivery?
-

Resource provisioning

9. How important is the internet as a resources in normal (physical) learning environments?

i) Very important ii) Important. iii) Moderately important iv) Slightly important v) Unimportant.

10. Students in technical colleges are aware of learning resources in their professions or trades?

i) Strongly Disagree... ii) Agree.... iii) Neutral... iv) Disagree... v) Strongly Disagree...

11. What is your opinion on resource requirements for competence based education training (CBET) as compared to knowledge based education training (KBET). _____

Assessment process

12 Assessment process for Competence Based Education Training require modern technologies

for Capture, Analysis and Storage than Knowledge Based Education Training?

i) Strongly Agree... ii) Agree.... iii) Neutral... iv) Disagree... v) Strongly Disagree...

13. What is your opinion on the influence of assessment process on the consistency of quality learning? _____

Feedback Capture Frequency

14. How should Students' feedback be captured in evaluating the training-learning process?

i) Always..... ii) Often..... iii) Sometimes..... iv) Rarely..... v) Never.....

15. What is your opinion on the effectiveness of continuous feedback from students for monitoring learning in comparison with periodic (end-of-term)

feedback? _____

Service / Performance rating

16 What is your opinion of Students' feedback in evaluation of training programs? _____

17 What is your opinion of students' feedback in the overall performance of a college? _____

APPENDIX 3

NACCOSTI Research Permit

REPUBLIC OF KENYA
Ministry of Education, Science and Technology
NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY & INNOVATION

Ref No: **84097**
Date of Issue: **04 October 2021**

RESEARCH LICENSE



This is to Certify that **Mr. Stephen Muga wa Muri** of **KCA University**, has been licensed to conduct research in **Kenya** on the topic: **A PREDICTIVE MODEL FOR QUALITY-ASSURANCE IN NORMAL LEARNING THROUGH STUDENTS FEEDBACK IN KIAMBU TECHNICAL COLLEGES** for the period ending: **04 October 2022**.

License No: **NACOSTIP/21/13485**
Applicant Identification Number: **84097**

Walter Muri
Director General
NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY & INNOVATION


Verification QR Code



NOTE: This is a computer generated License. To verify the authenticity of this document, Scan the QR Code using QR scanner application.

APPENDIX 4

TVET approval and facilitation letter



**TECHNICAL AND VOCATIONAL EDUCATION
AND TRAINING AUTHORITY**

Utalii House, Utalii Street
P. O. BOX 35825 – 00100
NAIROBI

Tel: +254 20 2392140
Email: info@tveta.go.ke
Website: www.tveta.go.ke

When replying please quote:
REF: TVETA/7/4-(38)

8th October, 2021

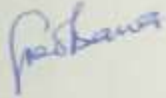
Principals
Technical and Vocational Colleges
Kiambu County

RE: DATA COLLECTION BY STEPHEN MAGU WAWERU

The above named is a Master's degree student at KCA University. He would like to collect data for his project titled **A predictive Model for Quality Assurance in Normal Learning through Students' feedback in Kiambu County Technical Colleges**. The information from this study shall important in improving the quality assurance and standards of training in the TVET institutions.

The purpose of this letter is to request you to allow the student to collect the data for his study.

Kindly provide him the necessary assistance.



Dr. Otta Osawa
FOR: DIRECTOR GENERAL/CEO

APPENDIX 5

Research Schedule

ID	Task Name	Start	Finish	Duration	Timeline (18/5/2021 to 1/10/2021)																																																																																																																																																																	
					18/5/21	19/5/21	20/5/21	21/5/21	22/5/21	23/5/21	24/5/21	25/5/21	26/5/21	27/5/21	28/5/21	29/5/21	30/5/21	31/5/21	1/6/21	2/6/21	3/6/21	4/6/21	5/6/21	6/6/21	7/6/21	8/6/21	9/6/21	10/6/21	11/6/21	12/6/21	13/6/21	14/6/21	15/6/21	16/6/21	17/6/21	18/6/21	19/6/21	20/6/21	21/6/21	22/6/21	23/6/21	24/6/21	25/6/21	26/6/21	27/6/21	28/6/21	29/6/21	30/6/21	1/7/21	2/7/21	3/7/21	4/7/21	5/7/21	6/7/21	7/7/21	8/7/21	9/7/21	10/7/21	11/7/21	12/7/21	13/7/21	14/7/21	15/7/21	16/7/21	17/7/21	18/7/21	19/7/21	20/7/21	21/7/21	22/7/21	23/7/21	24/7/21	25/7/21	26/7/21	27/7/21	28/7/21	29/7/21	30/7/21	31/7/21	1/8/21	2/8/21	3/8/21	4/8/21	5/8/21	6/8/21	7/8/21	8/8/21	9/8/21	10/8/21	11/8/21	12/8/21	13/8/21	14/8/21	15/8/21	16/8/21	17/8/21	18/8/21	19/8/21	20/8/21	21/8/21	22/8/21	23/8/21	24/8/21	25/8/21	26/8/21	27/8/21	28/8/21	29/8/21	30/8/21	31/8/21	1/9/21	2/9/21	3/9/21	4/9/21	5/9/21	6/9/21	7/9/21	8/9/21	9/9/21	10/9/21	11/9/21	12/9/21	13/9/21	14/9/21	15/9/21	16/9/21	17/9/21	18/9/21	19/9/21	20/9/21	21/9/21	22/9/21	23/9/21	24/9/21	25/9/21	26/9/21	27/9/21	28/9/21	29/9/21	30/9/21	1/10/21	2/10/21	3/10/21	4/10/21	5/10/21	6/10/21	7/10/21	8/10/21	9/10/21	10/10/21	11/10/21	12/10/21	13/10/21	14/10/21	15/10/21	16/10/21	17/10/21	18/10/21	19/10/21	20/10/21	21/10/21	22/10/21	23/10/21	24/10/21	25/10/21	26/10/21
1	Title & Problem Definition	20/5/2021	27/5/2021	6.0 d.																																																																																																																																																																		
2	Concept paper preparation	22/5/2021	17/6/2021	19.0 d.																																																																																																																																																																		
3	Formulation of Strategy & Methodology	22/6/2021	27/7/2021	9.0 d.																																																																																																																																																																		
4	Writing Research proposal	7/7/2021	2/9/2021	42.0 d.																																																																																																																																																																		
5	Literature review	22/6/2021	1/9/2021	52.0 d.																																																																																																																																																																		
6	Presentation of Research Proposal	27/8/2021	27/8/2021	1.0 d.																																																																																																																																																																		
7	Data collection	28/8/2021	1/10/2021	25.0 d.																																																																																																																																																																		
8	Data analysis	21/9/2021	24/9/2021	4.0 d.																																																																																																																																																																		
9	Model Design	24/9/2021	29/9/2021	4.0 d.																																																																																																																																																																		
10	Model Testing & Evaluation	27/9/2021	30/9/2021	4.0 d.																																																																																																																																																																		
11	Preparation of dissertation document	28/8/2021	15/10/2021	35.0 d.																																																																																																																																																																		
12	Compilation of Final Research Document	1/10/2021	5/10/2021	3.0 d.																																																																																																																																																																		
13	Final Presentation	5/10/2021	5/10/2021	1.0 d.																																																																																																																																																																		

Source: Author (2021)

APPENDIX 6

Research Budget

Budget item	Number of items	Cost per item(Shs)	Total cost (Shs)
Interviews Travel	10 Trips	1,500	15,000
Computer	1	50,000	50,000
Stationaries	1 realm	500	500
Daily-Data bundles	120	200	24,000
Phone calls	40	100	4,000
			91,500

Source: Author (2021)