

**USE OF DATA MINING TOOLS IN DETERMINING PATRONS' INFORMATION
NEEDS FOR COLLECTION DEVELOPMENT IN SELECTED ACADEMIC
LIBRARIES IN KENYA.**

KINYUA MARGARET WAMUYU (BSc)

E65/OL/CTY/32645/2016

**A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF
SCIENCE (LIBRARY AND INFORMATION SCIENCE) IN THE SCHOOL OF
PURE AND APPLIED SCIENCES OF KENYATTA UNIVERSITY.**

NOVEMBER, 2025

DECLARATION

I attest that this research project is my own and has not been presented for academic credit elsewhere. Any information, text, images, or tables that were not created by the author were properly credited and cited using the most recent APA style.

Signature _____

Date: _____

Kinyua Margaret Wamuyu

E65/OL/CTY/32645/2016

Department of Computing and Information Science

Kenyatta University

Supervisor's Declaration

I certify that the candidate worked on this project under my supervision as the university's supervisor.

Signature _____

Date: _____

Dr. Rose Njoroge

Lecturer, Department of Computing and Information Science

Kenyatta University

DEDICATION

To my son Geoffrey, my parents, and my siblings for their constant support and encouragement throughout the course of my studies.

ACKNOWLEDGEMENTS

I give thanks to the Almighty God, to whom I owe my life, strength, and health. Because of Him, I am able and have the means to do this research. My supervisor, Dr. Rose Njoroge, thank you for being available and helping me through every step of this research project. Your patience and guidance made me pull through. My appreciation to all lecturers in the Department of Computing and Information Science for the roles they played in aiding my study and for their assistance throughout my Master's Program.

To my parents, I appreciate your unwavering support, mentally, spiritually and financially. May the good Lord bless you abundantly.

Finally, I extend my gratitude to my colleagues and fellow students for the roles they each played during this project.

TABLE OF CONTENTS

DECLARATION.....	ii
DEDICATION.....	iii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS AND ACRONYMS	xi
ABSTRACT.....	xii
CHAPTER ONE	1
INTRODUCTION.....	1
1.1 Background	1
1.2 Statement of the problem	6
1.3 Purpose of the study	7
1.4 Objectives of the study	7
1.5 Research questions	8
1.6 Significance of the study	8
1.7 Limitation of the study	8
1.8 Assumptions of the study	9
1.9 Theoretical Framework	10
1.10 Conceptual framework	11
1.11 Operational definition of terms	13
CHAPTER TWO	14
LITERATURE REVIEW.....	14
2.1 Introduction	14
2.2 Data mining tools	14

2.2.1 Data mining in libraries	15
2.3 Awareness of data mining tools	16
2.4 Use of data mining tools in the collection development process	17
2.5 Current collection development process in academic libraries	19
2.6 Summary of literature and research gap.....	20
CHAPTER THREE.....	22
RESEARCH DESIGN AND METHODOLOGY	22
3.1 Introduction	22
3.2 Research design and methodology.....	22
3.3 Location of the study.....	22
3.4 Target population	24
3.5 Sampling technique and sample size.....	25
3.5.1 Sampling technique	25
3.5.2 Sample size.....	26
3.6 Research instruments.....	28
3.6.1 Questionnaires	28
3.6.2 Interview schedule.....	29
3.6.3 Observation checklist	29
3.7 Pretesting.....	29
3.8 Validity and reliability	30
3.9 Data collection.....	31
3.10 Data analysis and Presentation.....	31
3.11 Logistical and ethical considerations	32
CHAPTER FOUR.....	34
PRESENTATION OF FINDINGS, INTERPRETATION AND DISCUSSION	34
4.1 Introduction	34

4.2 General Information	34
4.3 Response Rate of Participants	35
4.4 Data mining tools available for use in collection development	36
4.4.1 Channels used by librarians to notify their users of new information resources.....	37
4.5 Librarians' awareness of existence and use of data mining tools	39
4.5.1 Librarians' awareness of the platforms used by their libraries to gather data.....	40
4.6 Post-Graduate students' contribution in the collection development process by use of data mining tools.	41
4.6.1 Submission of requests, complaints or suggestions by postgraduate students	42
4.6.3 Postgraduate students' involvement in the collection development process.....	43
4.6.4 Post Graduate Students' opinion on desire to be involved by their academic libraries in the collection development process.....	44
4.7 Current acquisition processes for development of a Patron Driven Acquisition	45
4.7.1 Factors considered during selection of information resources	46
4.7.2 Submission of suggestions for information resources to be acquired	48
4.7.2 Did the library acquire the information resources requested.....	49
4.7.3 User-centeredness of the library collection	50
CHAPTER FIVE	52
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	52
5.1 Introduction	52
5.2 Summary	52
5.3 Conclusion.....	54
5.4 Recommendations	55
5.5 Suggestion for further research	56
REFERENCES.....	57

APPENDICES	64
Appendix I: Informed Consent.....	64
Appendix II: Introductory Letter.....	66
Appendix III: Interview Schedule for University Librarians	67
Appendix IV: Questionnaire for Post Graduate Students	69
Appendix V: Questionnaire for Acquisition Librarians.....	72
Appendix VI: Observation Checklist	75
Appendix VII: Research Permit	76

LIST OF TABLES

Table 3.1: Target population distribution	25
Table 3.2: Sample size distribution.....	28
Table 4.1: Response rate	35
Table 4.2: Available data mining tools	36
Table 4.3: Channels used by librarians to pass information	38
Table 4.4: Requests submitted by Post Graduate students	43

LIST OF FIGURES

Figure 1.1: Conceptual framework	11
Figure 4.1: Librarians' awareness of data mining tools' availability	40
Figure 4.2: Submission of requests, complaints or suggestions by Post Graduate students....	42
Figure 4.3: Postgraduate students' involvement in collection development	44
Figure 4.4: Post-Graduate students' opinions on involvement in collection development	45
Figure 4.5: Factors considered during information resources selection process	47
Figure 4.6: Librarians' response on Post Graduate students' submission of suggestions to their university libraries	49
Figure 4.7: University libraries reaction towards information resources requested	50
Figure 4.8: User-centeredness of the current library collection.....	51

LIST OF ABBREVIATIONS AND ACRONYMS

CPU	-	Cost Per Use
E-resources	-	Electronic Resources
JKUAT	-	Jomo Kenyatta University of Science and Technology
KU	-	Kenyatta University
PDA	-	Patron Driven Acquisitions
RSS feeds	-	Really Simple Syndication; a web feed that allows users access to updates on new online content in their websites of interest without having to visit each individual website to check for updates
SPSS	-	Statistical package for social sciences

ABSTRACT

Data mining is defined as the process of identifying and extracting useful information from a large pool of raw data. In context of this study, data mining tools refer to the techniques, strategies, channels used to collect raw data into the data warehouse. Though data mining tools have been in use for a considerably long time, it has been emphasized and largely used in commercial settings: manufacturing, banks, the health sector and the customer service industries in order to gather feedback from clients in bid to offer satisfactory services to their clients. The studies reviewed lacked information on how raw data was obtained and from whom. The studies reviewed also had a glaring deficiency on the use of data mining tools by academic libraries for developing a patron driven collection. This study therefore sought to look into the application of data mining tools in the collection development process and focused on the aspect of data mining at the point of collecting raw data, rather than analyzing the data to form patterns. The purpose of this study was to evaluate how academic libraries embrace and utilize data mining tools in order to build a collection based on the information needs of its users. The research was based on the following objectives: to establish various data mining tools available in libraries, to assess librarians' awareness of the existence of data mining tools, to determine contribution of postgraduate students in the collection development process by use of data mining tools and to assess the current acquisition processes in use. The study focused on two academic libraries; Kenyatta University and Jomo Kenyatta University of Agriculture and Technology. Data was collected from the university librarian, acquisition librarians and postgraduate students from each university. The university librarians and acquisition librarians were purposively selected and Slovic's formula was used to derive the sample population for Post Graduate students. Data was collected using interviews for the university librarians and questionnaires for acquisition librarians and Post Graduate students. The data collected was analysed using the Statistical Package for Social Sciences, and presented in form graphs, tables, and charts. The study derived the following findings: that data mining tools were available in the libraries, librarians were aware of the existence of various data mining tools, the Post Graduate students were not directly involved in the collection development process and that the libraries relied heavily on curriculum requirements and suggestions from the faculty during selection and acquisition of information resources. The study recommended that libraries should sensitize their staff on the use of data mining tools for collection development. The study further recommended that librarians should capitalize on their patrons' input where data mining tools play a significant role in order to achieve a patron driven collection.

CHAPTER ONE

INTRODUCTION

1.1 Background

The world is currently in the knowledge economy where large amounts of information are produced and collected daily (Ojanpera, Graham, Staumann, Sabbata & Zook, 2017). With the adoption and use of advanced technology, a lot of this information is captured and stored in various storage devices such as organization servers. It is therefore not possible for organizations to manually go through all the information they collect in a day to retrieve data that is relevant to the smooth running of their business. In bid to access and retrieve information relevant to them, they put in place various measures. Traditionally data analysis was done manually which was a cumbersome exercise and very time consuming. Technology's widespread use and adoption has resulted in the automation of what was formerly known as data analysis; this practice is now known as data mining. Information is extracted from pre-existing data and patterns through a process known as "data mining." Other terms similar to data mining are knowledge extraction; knowledge mining from data etc. (Renaud, Britton & Wang, 2015).

(Silwattananusarn & Tuamsuk, 2012) define data mining as the process of extracting meaningful knowledge from massive datasets utilising data mining tools. (Khemakhem & Boujelbene, 2018) further describe data mining as a set of strategies and algorithms for extracting information and knowledge from huge databases in order to describe the current behaviour and/or forecast the future behaviour of data. From the studies referenced, data mining is interpreted as use of tools, techniques or strategies to filter out meaningful information from a pool of raw unorganized data; alternatively referred to as a data warehouse. After data mining is conducted successfully,

organizations are able to derive meaningful statistics or information which makes a large contribution towards offering satisfactory services to their clients.

In their studies, (Warokka, 2020; Zong, Xia & Zang, 2021) outlined the following as steps of data mining process: State the problem, collect the relevant data, pre-process the data, extract patterns, interpret the patterns and draw conclusions. However, the studies fail to discuss how the raw data was collected, from whom and how it was captured into the pool or data warehouse. This study therefore focuses on the activities, tools and strategies put in place by selected academic libraries to collect raw data from their clients. This study will discuss in detail the various categories of key stakeholders, their contributions towards data mining and the various tools used. For the purpose of this study, the researcher chose to use the term 'data mining tools' to represent all activities, strategies and/or techniques used to collect raw data. (Scout, 2022) discusses some of the most commonly used tools in collecting raw data. They include: emails, social media channels such as (Facebook, LinkedIn, Instagram), usage statistics, websites, live chats, phone calls, and in-person consultations.

In the pharmaceutical industry, the flow of information is very high. These firms are embracing technology enabled products and services. The use of data mining tools helps them discover patterns from their data repository thus facilitating them in improving the quality of the drugs they produce. In a study conducted in India by (Sohrabi, 2019) the researcher found out that identification of the types of drugs regularly requested by clients is very vital to the pharmaceutical industry as it forms a basis for decision making on the types and quantity of drugs to be produced. Information on how patients interact and react to various drugs is also very crucial to this industry as it guides them on how to issue prescriptions.

According to (Khemakhem & Boujelbene, 2018), the banking industry also relies on data mining tools to predict credit risk by evaluating the financial health of individuals or companies, helps in market segmentation, customer service, catching fraud, figuring out credit scores, and benchmarking companies and predict future client trends, which is used for strategic decision-making. (Warokka, 2020) discusses the role played by data mining in the marketing industry when making financial decisions. Data mining allows for marketers to craft more effective marketing campaigns and strategies based on their clients' behaviour, preferences, and purchasing patterns.

According to (You, Joo & Katsurai, 2022), libraries use data mining tools to anticipate information needs of their patrons by identifying the most referred topics. They then incorporate this information when selecting information resources to acquire and thus contributing towards developing a 'just in time' collection.

Academic libraries in America are implementing patron-driven acquisitions (PDA). It is a tool for collection development that uses patrons' feedback to meet just-in-time information needs (Schroeder, Wright & Murdoch, 2010). PDA largely shifts the decision for material acquisitions from librarians to the information users. Library patrons provide vast information as feedback, complaints, suggestions etc. which is stored in the library's database. Data mining tools enable librarians identify the specific information they need to aid in selecting and acquiring information resources. A study done in University of Illinois found out that the information resources acquired based on students' suggestions were frequently used compared to those selected by librarians (Crawford, 2020; Galbraith, 2020).

In Nigeria, the most commonly adopted collection development framework is the use of curriculum requirements, where the faculty members identify information resources relevant to

their courses and recommend them (Oyeleye, 2015). New knowledge is created every day, new publications and coupled by the fact that academic libraries are often faced with budgetary constraints (F. I. Okogwu, 2018; Rath & Rath, 2014), therefore, these libraries are tasked with the duty of maximizing the limited financial resources and ensuring they achieve patrons' satisfaction.

A study conducted by (Okogwu & Ozioko, 2018) reported that academic libraries do not generate income and therefore relies on the parent institution for funding. Often, the financial resources they are allocated are limited and therefore they have to devise means of maximizing the funds they have. Therefore, it is the responsibility of the acquisitions librarian to design measures to ensure that the final collection will meet the needs of library patrons. Acquiring information resources that have little or no relevance to their users would be a waste of the available funds, as they would be barely utilized. Academic libraries serve patrons with diverse information needs. These needs change rapidly as courses are eliminated or introduced into the institution. Regardless of these changes, libraries have to keep up and provide relevant information needs to every user etc. (Mason, 2016). Another challenge faced by acquisition librarians is the rapid growing information industry. With adoption and use of online and self-publishing, new publications are released every now and then. Some subject areas tend to prefer certain publishers over others. In an effort to deal with these challenges, acquisition librarians ought to put into consideration involving their patrons in the process of selection of information materials. This can be achieved by providing platforms for patrons to interact with librarians through making suggestions of preferred resources, giving feedback, complaints etc.

(S. Diwani & Sam, 2014) conducted a study in Tanzania on the use of data mining in the health sector. The study found out that data mining was used to categorize patients into groups, find patients with recurrent health issues, determine the relationship between an illness and its

symptoms, reduce treatment costs, anticipate medical diagnoses which in turn assist doctors in determining which drugs work best for their patients. The study also pointed out that the health care workers' particularly their willingness to accept data mining tools, can be considered as a major success or hurdle to implementing them.

Academic libraries exist primarily to promote research, teaching, and learning. Libraries therefore have to attend to individual information needs of their clientele to their satisfaction. This is sometimes hindered by the ever-increasing publications of electronic resources which lead to information explosion. When renewing or making new subscriptions, academic libraries put into consideration the financial resources available and coupled by the fact that they are often faced with budgetary constraints; measures to avoid misuse of the limited financial resources available ought to be put in place (Crawford, Condrey & Avery, 2020). Diversity of patrons' information needs requires academic libraries to acquire information resources to satisfy each user group in the academic institution. The resources should have relevant content, desirable formats and their currency should also be considered.

This study was conducted in two academic libraries, hosted in public universities: Kenyatta University and Jomo Kenyatta University of Science and Technology. In Kenyatta University, the research focused on the School of Security, Diplomacy and peace studies. According to the university website, the school was established in the late 2015 and is the latest addition to Kenyatta University's schools. Its focus is offering educational programs to students from law enforcement agencies and departments namely; the directorate of criminal investigations, the Kenya defence forces, national police service and the administration police. While in Jomo Kenyatta University, the research focused on the school of Communication and Development studies, established in mid-2014, according to the school website. This research looked into how the university libraries

are equipped with the relevant information resources to cater for the information needs of the students in the universities' latest established schools. This study aimed to determine if the selected academic libraries had sufficient information resources to suit the demands of graduate students taking classes at those institutions.

1.2 Statement of the problem

Numerous studies have been conducted on the effectiveness of data mining tools in the medical field, industrial sector, banking and customer service industry and few studies have looked into the application of data mining tools in libraries for the purpose of the collection development.

In the pharmaceutical industry, identification of the types of drugs regularly requested by clients is very vital data mining tools form a basis for decision making on the types and quantity of drugs to be produced. Information on how patients interact and react to various drugs is also very crucial to this industry as it guides them on how to issue prescriptions, determine the relationship between an illness and its symptoms, anticipate medical diagnoses which in turn assist doctors in determining which drugs work best for their patients (S. Diwani & Sam, 2014; Sohrabi, 2019).

(Khemakhem & Boujelbene, 2018) reported that the banking industry also relies on data mining tools to predict credit risk by evaluating the financial health of individuals or companies, helps in market segmentation, customer service, catching fraud among others. (Warokka, 2020) discusses the role played by data mining in the marketing industry when making financial decisions; they facilitate crafting of more effective marketing campaigns and strategies based on their clients' behaviour, preferences, and purchasing patterns. these discussed in depth the advantages derived from incorporating data mining tools in the respective industries. However, these studies put a lot of emphasis on the technical application of data mining tools which result to pattern prediction that play a huge role in decision making. The studies failed to show the process followed by the

industries in obtaining raw data and from whom it was collected. The studies reviewed also displayed a glaring deficiency on the use of data mining tools by academic libraries for developing a patron driven collection. This study therefore sought to look into the application of data mining tools in the collection development process and focused on the aspect of data mining tools at the point of collecting raw data, rather than analyzing the data to form patterns.

This study was guided by the following objectives: establishing various data mining tools available for collection development in selected academic libraries in Kenya, assessing librarians' awareness of the existence of data mining tools for collection development in selected academic libraries in Kenya, determining the contribution of postgraduate students in collection development by use of data mining tools in selected academic libraries in Kenya.

1.3 Purpose of the study

The goal of this study was to determine how selected academic libraries employ various data mining tools in determining the information needs of their patrons in order to facilitate development of a patron-driven collection.

1.4 Objectives of the study

1. To establish various data mining tools available for use in collection development in selected academic libraries in Kenya.
2. To assess librarians' awareness of the existence of data mining tools for collection development in selected academic libraries in Kenya.
3. To determine contribution of postgraduate students in the collection development process by use of data mining tools in selected academic libraries in Kenya.
4. To assess the current acquisition processes for collection development in use by selected academic libraries in Kenya.

1.5 Research questions

- i. Which are the various data mining tools available for use in the libraries that facilitate the collection development process in selected academic libraries in Kenya?
- ii. Are librarians aware of the existing data mining tools for collection development in their library in selected academic libraries in Kenya?
- iii. Do postgraduate students make any contribution during the collection development process by use of data mining tools in their academic libraries?
- iv. How do libraries identify and select the information resources to acquire for their libraries?

1.6 Significance of the study

This study provided a baseline to librarians intending to incorporate data mining tools in their acquisition process. In addition, this study demonstrated the need for academic libraries to provide user-friendly platforms where their patrons can post feedback, make enquiries and suggestions on the information resources they would like added in their academic libraries' collection.

1.7 Limitation of the study

This study experienced some limitations. The first limitation was slow response rate from the post graduate students. This was attributed to the fact that this category of respondents was on DSVOL mode of learning, therefore they were rarely present in the libraries in person thus it was a challenge following up on the responses. The researcher mitigated these limitations by letting the respondents know that their responses would contribute greatly to the collection development

process in their respective libraries, which would in turn result to satisfaction of their information needs.

The second limitation was that the respondents were in fear of being victimized based on the responses they gave. To counter this, the researcher notified the respondents that all the responses collected were anonymous and that the research was guided by a code of ethics which the researcher had to abide to.

1.8 Assumptions of the study

The researcher assumed:

- i. All postgraduate students were under Open and Distance and Electronic Learning mode, hence the need for using data mining tools to reach their respective libraries.
- ii. The librarians gave accurate information on the use of data mining tools in determining patron's information needs for collection development.
- iii. The postgraduate students provided truthful information on their involvement during the selection process and their perception towards the existing collection in their libraries.
- iv. The researcher used postgraduate students from two specific schools; that is, School of Security, Diplomacy and peace studies at Kenyatta University and school of Communication and Development studies at Jomo Kenyatta University of Agriculture and Technology, to represent the library users' category. The researcher selected these schools based on the assumption that since they are relatively new compared to the other schools/ faculties, the university libraries were actively working on collection development to meet the information needs of their users under these schools.

- v. The researcher also worked with the assumption that post graduate students are in a better position to identify and articulate their information needs, thus the focus on this category of library users.

1.9 Theoretical Framework

This study adopted the crowd theory in collection development by (Strong & Galbraith, 2018). This theory was developed from the book “The Wisdom of Crowds by James Surowiecki (2004). The theory argues that a crowd of cognitively diverse people make better decisions than one or two experts.

The crowd theory in collection development was developed after conducting research on the collection development process in libraries and the factors considered during the selection process. It suggests that the most pressing issue in libraries is the process of selecting and acquiring information resources. The theory found out the use of patron driven acquisition increased the rate of relevance and usage of information resources. The theory argues that library patrons, as end users of the materials the libraries acquire, can provide invaluable feedback during the selection process. Patrons are from diverse groups thus shedding light on what they consider relevant to them. Implementing crowd feedback highly contributes to developing the best possible collection. Since the theory so convincingly illustrated how "power lies in the crowd," it was appropriate for use in this research. Information users are the end consumers of the collection development process. Patron driven acquisition can only be achieved if information users are involved in the selection and acquisition process. Each user has different requirements in terms of format, subject matter, and currency, so their contributions will go a long way toward ensuring that the collection built meets their requirements.

1.10 Conceptual framework

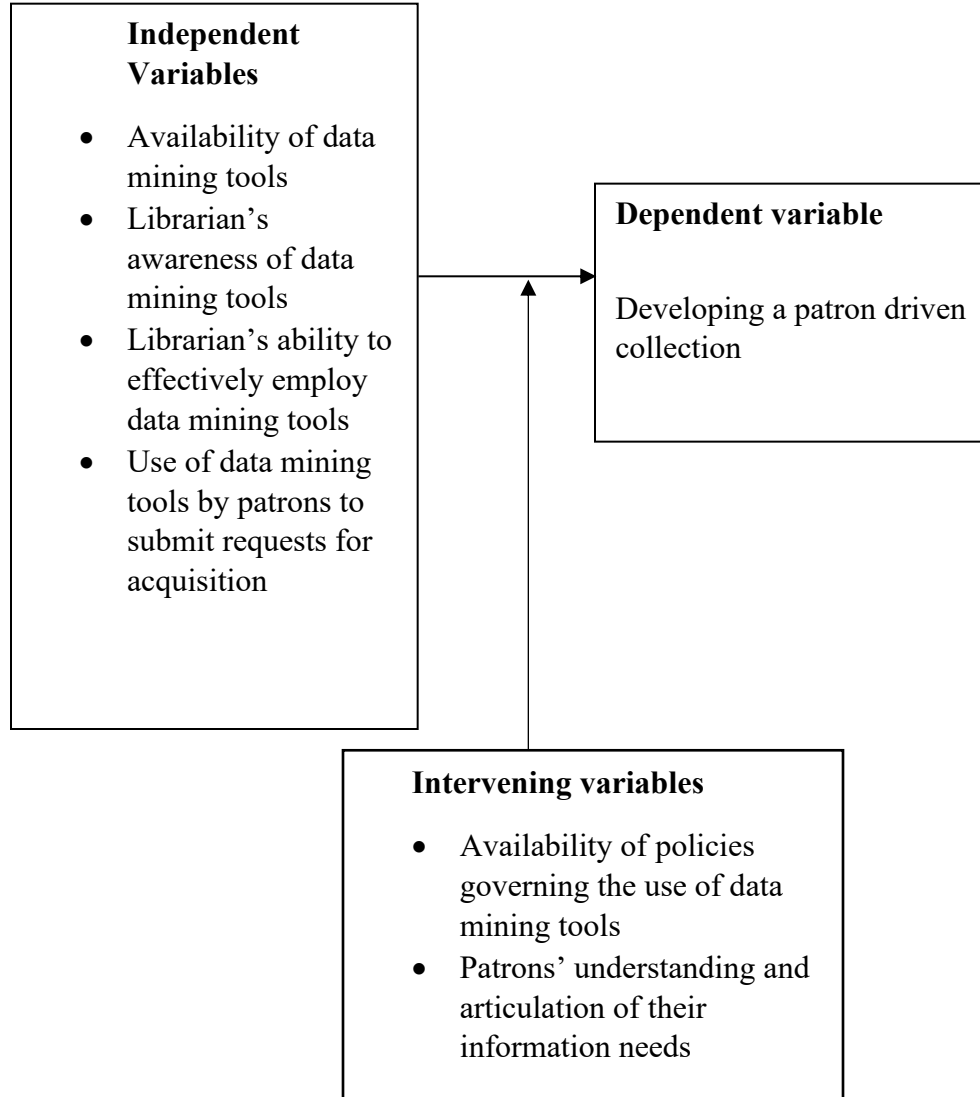


Figure 1.1: Conceptual framework

Source: Researcher (2022)

The key role of academic libraries is research, teaching and learning which means that their success is measured by how well they are able to satisfy their diverse patrons: students, researchers, the faculty (Crawford, 2020; Galbraith, 2020). From the above framework, the satisfaction of patrons' information needs was influenced by the kind of collection developed. Identifying the specific information needs of library users is crucial in developing a collection that is tailored to their preferences. Data mining tools can be used to facilitate identification of these needs. However, for this to be successful, patrons will need to be aware of, and able to articulate their information needs and use data mining tools to make specific requests for the information resources they need. Librarians too need to be aware of data mining tools and how to use them effectively.

The 'Crowd theory' by (Strong & Galbraith, 2018) explicitly discusses how the power belongs to the masses. It outlines the importance of taking into consideration the opinions and voices of the involved persons. Per the above framework, the ultimate goal of collection development is building a patron driven acquisition (PDA). However, in order to achieve a PDA, the libraries should consider the contributions/ voices of the end users of the collection being developed. This can be achieved by: determining the availability of data mining tools, librarian's awareness of data mining practices, librarian's ability to effectively employ data mining tools, use of data mining tools by patrons to submit requests for acquisition, policies governing the use of data mining tools and patrons' understanding and articulation of their information needs.

1.11 Operational definition of terms

This section identifies specific terms and gives their meaning as used in the study.

Acquisition librarian- In context of this, this are Heads of various sections in the library. They have been referred to as acquisition librarians since they are best positioned to identify the various resources most needed by library patrons.

Bibliomining- The process of data mining carried out in libraries.

Data mining tools- In context of this study, these are the techniques, strategies, channels used to collect raw data into the data warehouse.

Data mining- the process of identifying and extracting useful information from a large pool of raw data.

Data warehouse- a central repository that holds raw data collected by organizations from their stake holders, for in instance; suggestions, complaints, enquiries.

Just on time information resources- information resources acquired solely as a result of patrons' suggestions.

Just-in-case information resources -information resources acquired using a library's acquisition policy where ideally the librarians select the resources

Patron Driven Acquisition (PDA)- information resources collection built with the patron's information needs in mind; also referred to as user-centred collection.

Patron -an individual who goes to the library to utilize its information resources

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter presents peer reviewed literature on the various types of data mining tools, librarians' awareness and knowledge of data mining tools, data mining tools available on library websites, user perceptions of library collections and factors that determine library acquisitions. It also discusses the research gap.

2.2 Data mining tools

The amount of data collected every day in organizations grows at an enormous rate and it is randomly stored in databases. This means that it is common to find totally unrelated data in one database which makes it hectic for organizations to act upon their clients' requirements. This has necessitated the adoption and use of data mining tools. The primary goal of data mining is to discover and identify new patterns within a large pool of data by extracting new and relevant information from the pool. Patterns identified are used to make predictions about clients' future behavior e.g., they help determine the likelihood of possible repeated purchases or lack thereof. (Chang & Chen, 2006) identified seven crucial steps for data mining in the library: Lay out data mining goals by deciding and outlining the main objectives to be achieved, identify relevant data worth considering and with helpful attributes, filter out irrelevant and erroneous data, data transformation- it brings together related data that could have been stored in different locations or formats, data storage- once all related data is collected and brought together, it is stored in a central location, data mining- based on the goals set, relevant tools and tools are put to work in order to extract meaningful information from the existing raw data and finally evaluation of mining results and taking relevant actions.

Various parts of the economy have embraced the use of data mining to facilitate smooth running of their operations and meet their clients' needs satisfactorily. In the banking sector for instance, customers tend to leave for another bank if their needs are not met as expected. Each client expects to get individualized attention and services from their bank (Khemakhem & Boujelbene, 2018b). Losing customers is a huge blow to banks and they might take a while to recover losses incurred. Data mining comes in handy when the bank needs to review feedback given by their clients on their services. The information extracted predicts customers' future patterns and gives the bank an upper hand in its efforts to retain clients. This can be achieved by offering incentives that are tailored to meet individual needs of clients. Data mining also helps banks detect possible avenues of fraud. In some instances, customers' transaction history and demographics are a red flag and could possibly facilitate fraud. Through data mining, financial institutions are able to identify suspicious activities e.g., doubtful clients' transactions. The process identifies transactions that are not standard repeat transactions or those that do not fit into a particular category (Chitra & Subashini, 2013). Through data mining banks are also at a better place in their risk management. Information retrieved shows loan repayment trends of their existing customers and therefore they can avoid incurring losses by issuing loans to clients who default or make delayed payments. Analysts also use information collected through data mining to review their marketing trends, identify what worked or failed earlier and devise new and improved marketing strategies in order to capture more customers.

2.2.1 Data mining in libraries

In libraries, data mining is referred to as bibliomining. It is used to track behaviour changes, trends and patterns of library transactions (Renaud, 2015). Academic libraries are required to continuously prove their effectiveness in facilitating the parent institution to achieve its primary goals in offering

quality education. They are expected to meet information needs of the ever-growing user population. The curriculum is often revised resulting to emergence of new information resources while others are considered irrelevant. These libraries ought to keep in touch with the changes and take necessary actions. In order to maximize on the limited financial resources available, academic libraries ought to be very specific and intentional during selection and acquisition of information resources.

A study conducted by (Renaud, 2015) reported that book usage analysis through the circulation history indicated how often patrons borrowed print information resources. Books borrowed frequently were considered relevant and of high importance to patrons compared to those that remained on the shelves. The study also revealed that graduate students preferred previously done thesis in their subject areas of interest compared to books. Librarians were also able to identify which databases were frequently visited. This helped in determining which resources and subject areas were on high demand. The search history of e-resources and circulation history of print resources were an indication of relevance of the existing collection in meeting patrons' information needs. This information can be used to predict access and use of information resources in future and thus very helpful during the selection and acquisition process.

2.3 Awareness of data mining tools

In order to utilize data mining tools and achieve maximum success, it is imperative that librarians are aware of the existence of these tools and how to put them to use. (Diwani & Sam, 2017) conducted a study on awareness and readiness to use data mining tools. The study reported that the respondents were neither aware of data mining tools existence nor how to use them. According to the study, the respondents were not aware of these tools since they had not been educated or

trained about them. However, they were ready and willing to adopt the use of data mining tools if given the chance to.

The fifth law of libraries by Ranganathan states that 'libraries are growing organisms'. In order to identify the information needs of the growing user population, libraries ought to identify their users' information needs and put them into consideration during collection development. A study by (Sengupta, 2017) revealed that despite the advantages derived from use of data mining tools in achieving client satisfaction in other industries, libraries were yet to embrace and apply data mining in their operations such as the selection and acquisition process. The study attributed this to the fact that data mining is a relatively complex process which can only be applied successfully when librarians are aware of the existence of data mining tools. In addition, they should be equipped with the relevant skills to apply these tools.

(Ajibade & Muchaonyerwa, 2022) put more emphasis on the importance of data mining in developing a patron driven acquisition. The study further states that in order to employ data mining tools effectively, librarians ought to be aware of the existence of the data mining tools at their disposal and have the necessary skills to employ them.

Equipped with information from these studies, this research sought to establish the librarians' awareness of data mining tools and their knowledge to employ the various tools effectively.

2.4 Use of data mining tools in the collection development process

Academic libraries have continued to play a vital role in institutions and they largely contribute towards their parent institutions' main aim of teaching, learning and research. Having observed the role played by academic libraries, (Das,2020) described libraries as the 'heart of the institutions'. In the knowledge economy, knowledge is power. It is a tool that comes in handy when dealing with the day-to-day occurrences in life. Libraries are home to various forms of

information resources. Their presence is inevitable in every institution. In academic institutions, they play a crucial role as an essential tool for learners and researchers. Libraries instill reading culture among learners, researchers, lecturers and equip them with learning skills, develop their creativity and innovation which is a huge contribution towards them being responsible individuals. (Nzivo, 2012) reported that rather than solely focusing on content delivery, libraries should develop a personal connection with their clients in bid to create a service delivery relationship which in turn creates trust between patrons and the library staff. This can be achieved by providing exceptional and personalized information services.

As an effort to meet their patrons' diverse information needs, academic libraries have kept in touch with emerging trends in information resources e.g., use of e-resources. They have also redesigned their facilities and services in order to accommodate all types of patrons. In order to ensure user satisfaction, academic libraries ought to capture changes in their patrons' behaviour and perception of the existing collection. Capturing these changes is essential in facilitating planning library services and facilities. How patrons perceive the library is dependent on whether the library meets their information needs satisfactorily. Satisfying user information needs is the key objective of academic libraries. (Kassim, 2009) pointed out that user satisfaction can be measured in three dimensions; through library services, infrastructure and the information resources available.

(Ajibade & Muchaonyerwa, 2022) note that data mining tools such as twitter feeds could help libraries identify their client base, categorize them in their respective subjects and preferred format of information resources. This information could be derived from the users' engagement with various content. Libraries could also evaluate which fields of knowledge get the most traffic and the specific demographic of users showing interest in those fields. A librarian could also create a

tweeter feed or Facebook page and interact with library users in real time to gather data that would help develop a patron driven acquisition.

2.5 Current collection development process in academic libraries

Traditionally, the responsibility of selection and acquisition of information resources has largely lied with the librarians and faculty members (Siddiqui, 2003). Acquisition librarians liaise with subject experts to get a recommended reading list. They also consult selection sources e.g. publishers' brochures, catalogues to identify and select information resources they deem relevant to their users.

A study by (Kulp & Rupp-Serrano, 2007) reported that there were three major approaches to selection of information resources in academic libraries . First, through representative groups which mostly comprise of departmental heads from multiple disciplines, second, using subject groups; these are experts from various disciplines in the institution and third individuals; heads of libraries who make decisions on information resources to acquire without approval from a larger group. According to the study, majority of the libraries combined two approaches in the selection and acquisition process: use of representative groups and individuals.

(Dahl, 2012) states that the value of a library is no longer assessed by how much a it holds rather by the relevance of its collection to the intended users. With this in mind, and coupled by the fact that academic libraries are often faced with financial constraints, there is need to develop a 'just in time' collection. This is a collection that is relevant to users' information needs, timely, current and in the preferred format. To achieve this, libraries have slowly embraced the concept of Patron Driven Acquisition (PDA), where information users are given the opportunity to voice their requirements and suggest relevant information resources. The study also revealed that the circulation rate for the information resources acquired through PDA was higher than those

acquired without users' involvement. From this study, users reported that they found the information resources acquired after their involvement more satisfying compared to previous collections.

2.6 Summary of literature and research gap

A study by (Warokka, 2020) reported that data mining is a key component in organizations and it has contributed to the success of service delivery and helped raise clients' satisfaction levels.

In the health sector, data mining tools come in handy in identifying patients with recurring health conditions, anticipate diagnoses and treatment options, (S. Diwani & Sam, 2014). Data mining tools are used in the marketing industry to allow for marketers to craft more effective marketing campaigns and strategies based on their clients' behavior, preferences, and purchasing patterns (Warokka, 2020).

According to (Khemakhem & Boujelbene, 2018), the banking industry also relies on data mining tools to predict credit risk by evaluating the financial health of individuals or companies, they are important for market segmentation, customer service, catching fraud, figuring out credit scores, and benchmarking companies. It extracts information from large databases using a set of methods and algorithms to describe current behaviour and predict future client trends, which is used for strategic decision-making.

A study by (Zong, Xia & Zang, 2021) outlined the following as steps of data mining process: State the problem, collect the relevant data, pre-process the data, extract patterns, interpret the patterns and finally draw conclusions. The studies reviewed majorly focused on one aspect of data mining tools; the pattern extraction and did not go into details of how raw data was collected.

From the studies reviewed, much emphasis has been placed on the effectiveness of data mining tools in the industrial sector, banking and customer service industry and only a few studies have looked into the use of data mining tools in the acquisitions process in libraries. This study is therefore aimed at exploring how the raw data is collected into the data warehouse and explore the awareness and use of data mining tools in collection development.

CHAPTER THREE

RESEARCH DESIGN AND METHODOLOGY

3.1 Introduction

This chapter describes the research design, where the study was conducted, the target population, sampling procedures, pretesting, sample size, reliability and validity of the data collection tools, research techniques employed, how data will be collected and analysed, logical and ethical considerations.

3.2 Research design and methodology

This study employed a descriptive research to describe the distinctive characteristics of the target population. This design allowed the researcher to collect data on many variables as well as information on the use of data mining tools in selection and acquisition of information resources as well as assess information users' satisfaction with their library's information resources. To explore various data mining tools available for collection development in selected academic libraries in Kenya, to assess librarians' awareness of the existence of data mining tools for collection development in selected academic libraries in Kenya, to determine the use of data mining tools by postgraduate students to request for information resources in selected academic libraries in Kenya and to establish the current acquisition processes in use by selected academic libraries in Kenya. Interviews were used to get information from university librarians, questionnaires; both structured and unstructured were used to get information from acquisition librarians and postgraduate students and an observation checklist was used to identify the data mining tools available in the selected academic libraries.

3.3 Location of the study

According to Commission of University Education's website, there are 63 Chartered universities in Kenya (35 public and 28 private), offering a wide variety of academic programs under different modes of learning. This study focused on two of the Public universities; Kenyatta University and Jomo Kenyatta University of Science and Technology. These universities were selected because they are among the oldest chartered universities. They also have a considerably high number of students under Digital School of Virtual and Open Learning (DSVOL).

Kenyatta University (KU) is a public university accredited by the Commission of University Education. According to the university website, the institution offers diverse courses in various levels; certificate, diploma, undergraduate and postgraduate. It also offers different modes of learning; fulltime, evening, institution based and online learning through Digital School of Virtual and Open Learning (DSVOL). To quench the thirst of knowledge of their clients: both students and staff members, the institution has an academic library famously referred to as The Post-modern library (PML). The library has a variety of information sources to meet the needs of a client base that is always growing. Post-Modern library has information resources in several formats; print resources, electronic resources, multimedia, and braille among others. This is in a bid to ensure that each user category is catered for.

Jomo Kenyatta University of Science and Technology (JKUAT) is a public university accredited by the Commission of University Education. It has a strong research interest in biotechnology and engineering. The university offers diploma, undergraduate and postgraduate programs. The university library has both printed and electronic information sources to meet the needs of all its users. The universities serve students, university staff members, and researchers from all over the world who have various information demands; therefore, they must be precise when identifying

the information needs of their patrons in order to maximize the limited funds available and guarantee user satisfaction.

3.4 Target population

This is the population on which the study intends to conduct research and develop findings (Barnsbee, 2018). The researcher collected data from the library management (university librarian), acquisition librarians and postgraduate students.

Kenyatta University library is divided into nine sections: circulation services, education subject, social sciences, humanities, science and technology, research and special collections, ICT, special needs section and technical services. Jomo Kenyatta University library has seven sections: technical services, reference services, circulation services, digital services, library systems, special collection and special needs.

The researcher targeted 2 university librarians; 1 from each academic library. University librarians were purposely targeted in order to represent the library's body of management.

In Kenyatta University, the research focused on the School of Security, Diplomacy and peace studies. According to the university website, the school was established in the late 2015 and is the latest addition to Kenyatta University's schools. Its focus is offering educational programs to students from law enforcement agencies and departments namely; the Directorate of Criminal Investigations, the Kenya Defence Forces, National Police Service and the Administration Police. While in Jomo Kenyatta University, the research focused on the school of Communication and Development studies, established in mid-2014, according to the school website.

The researcher focused on the relatively new schools since the other areas of study have been in existence for a longer period, thus their collection of information resources was built over time.

The study also focused on students DSVOL mode since they rarely visit the library in person thus the need to build the relevant collection.

Kenyatta University had 125 active postgraduate students in the School of Security, Diplomacy and Peace studies, while Jomo Kenyatta University had 200 postgraduate students in the school of Communication and Development studies.

Table 3.1 shows distribution of the target population, from which the sample population was drawn.

Table 3.1: Target population distribution

Institution	University Librarians	Acquisition Librarians	Postgraduate Students
Kenyatta University	1	8	125
Jomo Kenyatta University of Science and Technology	1	7	200
Total	2	15	325

3.5 Sampling technique and sample size

This section discusses in detail how the sample size was obtained.

3.5.1 Sampling technique

According to (Taherdoost, 2017) a good sample should represent the entire population it was selected from. The researcher used both purposive and simple random sampling techniques. (Sharma, 2017) defines purposive sampling as a subjective and judgmental sampling technique. It is completely dependent on the researcher's judgment in light of the nature of the study and target population. Purposive sampling was used in this case because the researcher needed data from

specific categories of respondents; the university librarians, acquisition librarians, and postgraduate students in specific schools.

Purposive sampling was used to single out one category of student; Post Graduate students. However, in order to achieve a manageable sample size from the target population, the researcher used simple random sampling. In this technique, all respondents have similar characteristics, each of them has an equal chance of selection and are selected independently from other respondents (Sharma, 2017). The author further reported that, to ensure that the respondents were randomly selected and given equal chances, responses are accepted on a first come first served basis until. The researcher adopted this method. All responses were accepted until the intended sample was achieved.

3.5.2 Sample size

A total sample population of 245 respondents was used in this research. This sample consisted of 228 postgraduate students and 17 librarians; both the university librarians and acquisition librarians from Kenyatta University and Jomo Kenyatta University of Agriculture and Technology.

The sample size was determined using the Slovin's formula (2016) where:

To determine the sample size, this study used the Slovin's formula where:

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

n- Sample size

N- Total population

e- Margin of error.

The researcher used a confidence level of 95%. According to the Slovin's formula the margin of error is calculated by subtracting the level of confidence from 1. In this case, the margin of error was 1-95% (0.95) to get 0.05.

In KU, the sample size was:

$$n = 125 / (1 + 125 * 0.05^2)$$

$$n = 125 / 1 + 0.3125$$

$$n = 125 / 1.3125$$

$$n = 95$$

Sample size = 95 students

In JKUAT, the sample size was:

$$n = 200 / (1 + 200 * 0.05^2)$$

$$n = 200 / 1 + 0.5$$

$$n = 200 / 1.5$$

Sample size = 133

The total sample size for postgraduate students was 228.

Table 3.2: Sample size distribution

Category of respondents	Sample size
Post-graduate students	228
University librarians	2
Acquisition librarians	15
TOTAL	245

3.6 Research instruments

The techniques used by researchers to collect data on their area of interest are termed research instruments (Jabbar, 2017). To collect data, the researcher used questionnaires, interview schedule and observation checklist.

3.6.1 Questionnaires

Based on the goals of the study, questionnaires were designed. When collecting data from participants, the researcher asked a mix of structured and unstructured questions. One of the Ministry of Health's community guidelines on COVID 19 was to limit exposure. In adherence to this guideline, the researcher cut down interactions, physical contact and limited travel as much as possible. To achieve this, the questionnaires were distributed virtually through Goggle Forms. The researcher collaborated with the schools' departments to facilitate this. Questionnaires were distributed to the acquisition librarians and postgraduate students via email.

The researcher opted to use questionnaires since they ensure respondents' anonymity was maintained and was possible to reach a large group of participants simultaneously irrespective of their geographical location. In order to maintain confidentiality and anonymity of respondents, the researcher deactivated collection of respondents' emails when submitting responses.

In order to ensure each respondent had an equal chance at submitting their response, the researcher deactivated acceptance of responses once the intended sample size was achieved.

3.6.2 Interview schedule

Interview schedule is a guide or plan for the study being conducted. It comprises of a set of pre-coded questions which make it easier for the researcher and helps keep the interviewer and respondent on track (Taherdoost, 2017). This study's interview schedule comprised both structured and open-ended questions. Thus, the researcher was able to obtain qualitative data.

(Alamri, 2019) reported that interviews provide a relatively higher chance of response rate; the researcher is able to collect more detailed data as well as interrogate sensitive matters. In addition to these merits, the researcher opted to use interview schedule for university librarians since they were only two respondents thus they were manageable.

3.6.3 Observation checklist

(Smit & Onwuegbuzie, 2018) define observation as one of the foundational data collection methods where the researcher is actively involved in the data collection process. The researcher collects data using their own senses: looking and listening in a systematic way.

This study used an observation checklist to identify the various data mining tools available in the selected academic libraries' websites. The researcher used data derived to make comparison between the observation made and information provided by the respondents.

3.7 Pretesting

According to (Jabbar, 2017), a pretesting is the process of trying out the efficiency of the research techniques on a population with equivalent characteristics to the target population. A pre-test study was conducted at Laikipia University library, a public university accredited by the Commission for University Education in 2013. A pilot study is conducted with the aim of anticipating possible challenges that could be encountered during the actual study and come up with solutions for each hurdle (Simkus, 2023). The researcher tested the questions readability,

ability of respondents to respond within the allocated time and identify any other barriers that would hinder successful data collection. Efficiency was measured by how well the respondents understood the tools and how relevant their responses were to the study. 25 questionnaires were sent to Laikipia University students, which represented 10% of the sample population. Based on the responses received, the researcher re-phrased questions that were not clearly understood and removed those that gave repetitive responses.

3.8 Validity and reliability

According to (Ishtiaq Ahmed & Sundas Ishtiaq, 2021), validity is determined by the data collections tools' accuracy in measuring the variable in question. The study further reported that having clearly defined objectives, data collection procedure and the target population largely contribute to the validity of data collection tools.

Reliability on the other hand is defined as the ability of a data collection tool in to achieving consistency of results (Jabbar, 2017). For a tool to be termed as reliable, it should produce the same results if the same study is to be conducted at a different time, under the same conditions.

To ensure that the data collections tools maintained validity and reliability, the researcher ensured that the selected data collection tools were customized to address study objectives and the research questions that the study sought to answer. The researcher also conducted an expert in the subject when designing the data collection tools. According to (Mugenda, 2003), a subject expert is often consulted to ascertain the validity of research tools. The data collections were then reviewed and revised by supervisors. Modifications were made until they were deemed fit for data collection.

3.9 Data collection

Data was collected using questionnaires and semi-structured interview schedule. In order to observe the Ministry of Health directives on COVID-19, the researcher designed questionnaires using Google forms, to be filled and submitted online. The researcher collaborated with the schools' departments and distributed these questionnaires to the target population through email. Once the respondents filled the questionnaires, they submitted them to the researchers' email ready for analysis.

The researcher allowed submission of responses by the postgraduate students and acquisition librarians for a period of six weeks and deactivated online submission when this period lapsed. The researcher settled on 6 weeks since that was the duration left to the end of the semester. Data from the university librarians was collected using in-person interviews. The researcher navigated the libraries' websites noting down the data mining tools available, which were checked against the observation checklist.

The researcher received 219 responses from postgraduate students, 13 responses from acquisition librarians and 2 from the university librarians.

3.10 Data analysis and Presentation

The term "data analysis" refers to the process of using rational and analytical techniques to evaluate research data (Ngulube, 2015). This study collected both qualitative and quantitative data. The information gathered in this study was used to develop themes based on the study objectives, which were then analysed and presented in charts and graphs. Qualitative data was analysed using Microsoft Excel while quantitative was labelled and examined through Statistical Package for Social Sciences.

3.11 Logistical and ethical considerations

Ethics in research are a guideline on how researchers should behave and conduct themselves. It guides them on how to observe copyright laws, maintain confidentiality, privacy and anonymity of respondents (Jabbar, 2017). In this study, the researcher ensured participants' privacy and anonymity was upheld all through. The questionnaires were distributed to the respondents in form of a link once they clicked on it; it opened the Google form questionnaire. After filling they clicked on the submit button, immediately the response was collected in the researcher's email. When designing the questionnaire, the researcher customized it in way that the respondents' email addresses were not collected. This way, none of the respondents' personal information was collected. This ensured that responses remained anonymous and the participants' confidentiality was upheld.

The researcher sought authorization from relevant authorities before embarking on data collection. Approval was sought from the researcher's department, Kenyatta University graduate school, Kenyatta University Ethics review committee, institutions where the study was carried out and National Commission for Sciences, Technology, and Innovation (NACOSTI). To observe the laws of copyright, the researcher ensured proper citation and acknowledgement of information sources used in the study.

The researcher also explained to the participants the study's purpose and why their contribution was important to the study, assured them that they would remain anonymous and responses submitted were confidential; and most importantly allowed them the freedom to decide whether to participate or not without coercion. When responding, the Google Forms were set in a way that it

was not compulsory for the respondents to answer. This way, participants were at liberty to choose which questions to respond to; this enhanced freedom of participation.

CHAPTER FOUR

PRESENTATION OF FINDINGS, INTERPRETATION AND DISCUSSION

4.1 Introduction

This chapter outlines the data presentation, analysis, results and discussions of findings. The information presented on this chapter is in line with the relevant literature reviewed and the study objectives which were to:

- i. To establish various data mining tools available for collection development in selected academic libraries in Kenya.
- ii. To assess librarians' awareness of the existence of data mining tools for collection development in selected academic libraries in Kenya.
- iii. To determine contribution of postgraduate students in the collection development process by use of data mining tools in selected academic libraries in Kenya.
- iv. To assess the current acquisition processes for collection development in use by selected academic libraries in Kenya.

4.2 General Information

The study targeted to collect data from two public university libraries: Kenyatta University and Jomo Kenyatta University of Science and Technology. The respondents were 2 University Librarians, 15 Acquisition Librarians and 228 Post Graduate students. An interview schedule was used to collect data from the 2 University Librarians while questionnaires were administered to the 15 Acquisition Librarians and 228 Postgraduate students. Out of the administered questionnaires, responses were collected from 219 Postgraduate students and 13 from Acquisition Librarians, 13. Interviews were conducted with the 2 University Librarians.

The analysed results were presented in charts, graphs and tables. Qualitative data was analysed using Microsoft Excel while quantitative was analysed with the help of Statistical Package for Social Sciences.

4.3 Response Rate of Participants

The study targeted 228 Post Graduate students, 15 acquisition librarians and 2 university librarians giving a total of 245 respondents. Out of this population, the researcher collected responses from 219 postgraduate students and 13 acquisition librarians and 2 university librarians. The response rate for Post Graduate students in this study was 96% representing 219 respondents out of the targeted 228 postgraduate students while that of the acquisition librarians was 81% representing 13 out of the targeted 15, both university librarians were interviewed which gave 100% response rate as illustrated on Table 4.1 below.

Table 4.1: Response rate

Category	Number of respondents targeted	Responses received	Percentage
Postgraduate students	228	219	96
Acquisition Librarians	16	13	81
University librarians	2	2	100

Source Field data (2022)

(Morton, Bandara, Robinson & Carr, 2012) noted that 70% response rate is considered sufficient in survey research. This study had an overall sample population of 245 respondents. The researcher was able to collect responses from 234 respondents, translating to 95%. With a non-response rate of 5% the target population adequately represented the study population. According to (Taherdoost, 2021) non-responsiveness could be attributed to several potential reasons: some

individuals are not open to submitting their personal views despite assuring them that confidentiality and anonymity will be observed, others find that filling the questionnaires take much of their time which they could be handling other issues, others only participate in research when there are rewards in return.

4.4 Data mining tools available for use in collection development

The first objective of the study was to find out the techniques used to gather users' information needs in the academic libraries. This was achieved by determining the various platforms that postgraduate students use to submit their suggestions and recommendations for information resources that their libraries should acquire.

The researcher used an observation checklist to identify the various data mining tools available in the selected academic libraries. The researcher identified the following data mining tools: Library website, Facebook pages, Email, phone numbers, library live chat (Ask a Librarian) and library help desk.

The researcher also collected this data from acquisition librarians and university librarians.

The responses were as summarized on table 4.2 below:

Table 4.2: Available data mining tools

Data mining tools	Responses received
Library website	9
Facebook page	13
LinkedIn	0
Email	10
Phone calls	5
Library live chat	14
Library help desk	3

Source: Field data (2022)

From the data collected and analysed on table 4.2 above, the most used tool was the Library live chat at 93%. The Facebook pages followed closely at 87%, then Email and Library website at 67% and 60% respectively.

From the responses collected, only 3 librarians reported to have collected information from their users through the library help desk, while 5 librarians indicated that they had used phone calls as a tool for collecting information needs of their users.

This is in line with (Carlsson & Tornngren, 2020) who reported that library users found it useful having access to large collection of e-resources with efficient search systems, responsive library staff also play a huge role in encouraging library users to submit their requests. The study also reported that library users largely made use of their library e-mail, website and social media accounts to air their opinions. Equipped with their users' information needs, librarians are able to develop their library's collection with the specific subject areas in mind. This goes a long way in developing a just-in-time collection which in turn helps the libraries achieve a high level of user satisfaction.

From the data collected from librarians, it was evident that majority of the data mining tools with high usage levels were technologically based. A considerably low number of librarians, 3 out of 15 (20%) reported to have collected information from their library users over the library help desk. The researcher attributed this data to the fact that since the target population was under the Open Distance mode of learning, the Postgraduate students rarely visited the library in person.

4.4.1 Channels used by librarians to notify their users of new information resources

The study further sought to establish how the librarians pass information to their users once they acquire the requested resources. Library users submit their requests with the hope that their libraries will acquire the information resources they require. It was therefore imperative to find out

the various communication channels used by librarians to reach out to their users and how effective they are. The channels below were indicated in the responses: the library website, email addresses, liaison with departments, library Facebook account and conducting exhibitions.

Table 4.3: Channels used by librarians to pass information

Channel	Responses received
Library website	15
Emails	13
Liaison with departments	7
Facebook	10
Exhibitions	13

Source: Field data (2022)

Out of the submitted responses, 100% of the respondents (15 out of 15) reported that they used the library website to reach out to their users, 87%(13 out of 15) reported that they sent out emails to their users, another 87% (13 out of 15) reported that their libraries organized exhibitions whenever they had new acquisitions, 67% (10 out of 15) reported to have notified their users via the library Facebook page and 7 out of 15 (47%) indicated that their libraries liaised with teaching departments to notify their users, and

(Fallon, 2016) outlines one of the key significance of communication in the library as marketing and promoting library resources and services. For a library developing a patron-driven acquisition, it would be futile to collect data on their users' information needs, acquire the requested resources but fail to notify the users. Results derived from this study demonstrated that libraries consciously made efforts to reach out to their users and conduct current awareness service. It is interesting to

note that less than half of the respondents (46%) stated that their libraries reached out to their users through the respective teaching departments. This could be attributed to the fact that once Post Graduate students embark on their theses, they rarely visit their departments. With this in mind, the libraries opted to engage their users virtually; through their website, social media account (in this case Facebook pages) or send out email and organizing book exhibitions. This could be attributed to the fact that libraries have embraced and incorporated Information technology in their day to day operations. This has in turn led increased use of virtual data mining tools compared to the traditional approach (Faulk & Crist, 2020). The study further reported that library users are more likely to gain information literacy skills through open and direct communication with their librarians.

4.5 Librarians' awareness of existence and use of data mining tools

The second objective for this study was to assess the librarians' awareness of the existence of data mining tools in their libraries. (S. A. Diwani & Sam, 2017) reported that in order to utilize data mining tools and achieve maximum success, it is imperative that librarians are aware of the existence of these data mining tools and how to put them to use. The key function of data mining tools to make recommendations for library resources based on search history and users' profile (Kovacevic, Devedzic & Pocajt, 2010). Effective application of data mining tools therefore requires libraries to gather and store a pool of data in regards to their users' information needs.

According to (Silwattananusarn & Tuamsuk, 2012), the most commonly used data mining tool in libraries is clustering. Librarians are able to get data from their data pool and create sub-groups based on similar characteristics. This will help them identify patterns based on their users' borrowing records, search histories or database searches. In order to effectively apply this tool, the

librarians ought to be aware of the various ways they can collect this data without necessarily consulting their users in real time.

4.5.1 Librarians' awareness of the platforms used by their libraries to gather data

This study sought to establish whether librarians were aware of existence of data mining tools in their libraries and the various tools they were aware of. The responses collected were as illustrated on Figure 4.1 below.

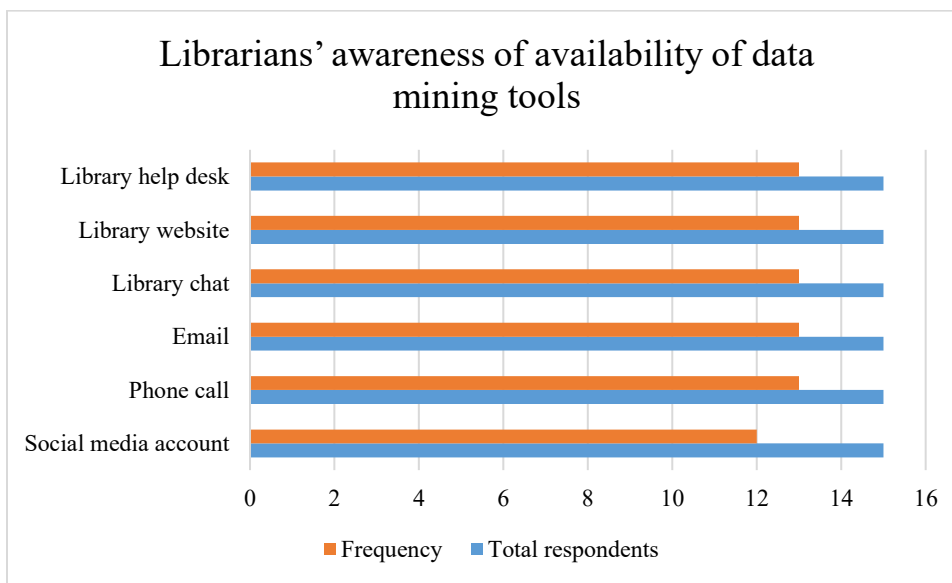


Figure 4.1: Librarians' awareness of data mining tools' availability

Source: Field data (2022)

As illustrated on Figure 4.1 above, librarians were aware of several data collection tools namely; library social media accounts, phone call, email, live chat, website and help desk.

The researcher conducted an observation to establish the specific social media channels used in the libraries. It was noted that the libraries had 2 social media accounts; Facebook and LinkedIn. However, data collected from librarians on data mining tools used for collection development,

none of the librarians indicated they used LinkedIn as a data mining tool. For the purpose of this study, Facebook will be used to represent social media accounts.

According to the data collected, all 13 respondents reported that they were aware of the existence of phone call, email, live chat, library website and help desk as data mining tools, translating to 100% of received responses. 92 % (12 out of 13 respondents) reported to have been aware of the existence of social media, specifically Facebook as a data mining tool in their library.

These are tools purposely set aside for users to interact with librarians in their academic libraries, this way, librarians are able to establish the various categories of users they are serving and also identify their information needs in bid to develop a patron driven acquisition. A study by (Cortés, Gabriel & Martins, 2011) concurred with these findings. This could be to the fact that in the 21st Century, a lot of activities, transactions are happening in the digital space and library services are no exception. Enquiries and requests that necessitated users to visit libraries in person have been made available online thus saving time for the users and increasing efficiency in service delivery.

These findings were in line with the data collected on objective one, on the use of data mining tools for collection development which found out that majority of the librarians used data mining tools during selection and acquisition of information resources.

4.6 Post-Graduate students' contribution in the collection development process by use of data mining tools.

The section sought to examine how postgraduate students used data mining tools to contribute towards developing a patron- driven acquisition. This was done by interrogating several aspects: postgraduate students' submission of requests, action taken on postgraduate students' requests and

postgraduate students' involvement in selection of information resources in their academic libraries.

4.6.1 Submission of requests, complaints or suggestions by postgraduate students

The study explored whether postgraduate students in the selected academic libraries submitted requests or recommendations of information resources to their libraries for incorporation in the libraries' collection and the nature of requests submitted.

The Postgraduate students were asked whether they submitted requests, complaints or suggestions to their libraries.

The results collected were as illustrated below in Figure 4.2 and Table 4.4 respectively.

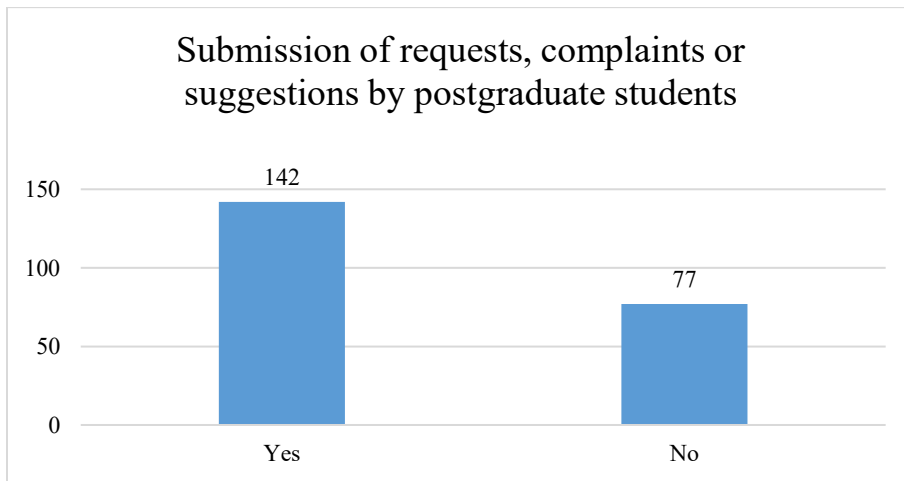


Figure 4.2: Submission of requests, complaints or suggestions by Post Graduate students

Source: Field data (2022)

As illustrated on Figure 4.2 above, 65% (142 out of 219 respondents) reported that they made suggestions, complaints and requests to their libraries while 35% (77 out of 219 respondents) answered that they did not make any submissions.

Table 4.4: Requests submitted by Post Graduate students

Submitted requests	Frequency	Percentage
Acquire books by specifically listed publishers	130	59
The library should several copies of a title	203	92
Unavailability of some research books	187	85
Satisfied with the library services as they are	57	26
Liaise with lecturers before purchasing books	145	66

Source: Field data (2022)

Table 4.4 shows a breakdown of the categories of requests submitted by Post Graduate students to their libraries. The most recurring suggestions were lack of adequate copies of physical information resources; at 92% and unavailability of the required resources at 85%. It is alarming to note that only 26% of the respondents indicated that they were satisfied by their library's collection. A study by (Shah, Khan, Ullah& Yar, 2021) reported that students were not satisfied with the resources and services offered in libraries. This was evidently displayed by this study, noting that majority of the suggestions made by the respondents were complaints about the inadequacy of the collection.

4.6.3 Postgraduate students' involvement in the collection development process

The study sought to establish from the Post graduate students if their library actively involved them in the information resources selection and acquisition process.

The findings were as illustrated below in Figure 4.3.

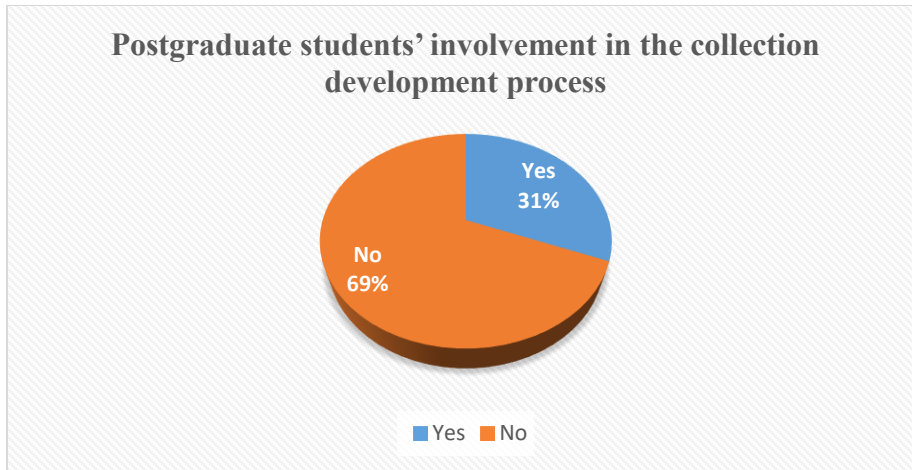


Figure 4.3: Postgraduate students' involvement in collection development

Source: Field data (2022)

From the data collected, 151 out of 219 (69%) of postgraduate students indicated that they were not actively involved in the selection process. These findings contradicted a study by (Kim, 2017) which argued that a library's collection is important to its users as much as the user is important in the selection process. According to the study, a library can claim to be developing a Patron-driven acquisition only if it involved its information users in the collection development process.

From the data collected, it is evident that postgraduate students were of the opinion that their input was not put into consideration. This beats the whole purpose of a patron driven acquisition where the user should be the focus.

4.6.4 Post Graduate Students' opinion on desire to be involved by their academic libraries in the collection development process

Data analysed and presented in figure 4.3 above revealed that postgraduate students were not actively involved in the selection process during collection development.

The study went a step ahead to establish whether the postgraduate students desired to be more involved in their libraries' collection development process or they were comfortable with the process as it was.

The data was analysed and presented in figure 4.4 below.

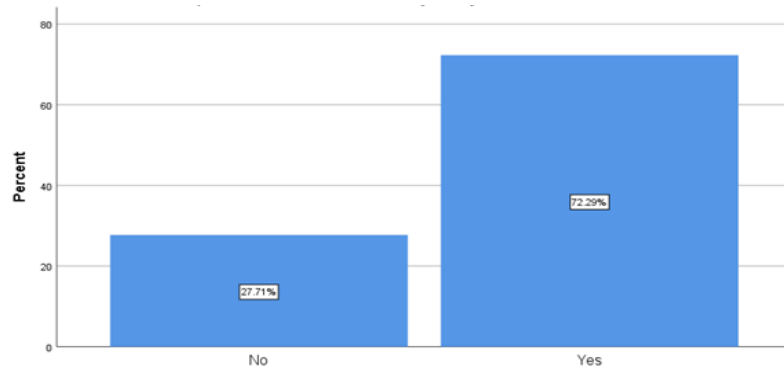


Figure 4.4: Post-Graduate students' opinions on involvement in collection development

Source: Field data (2022)

Majority of the students (72.29%) wanted to be actively involved in the collection development processes while 27.71% indicated that they were satisfied with the process as it was and therefore did not need to be involved in the collection development process.

This analysis showed that the students were willing to be involved in the user content collection process but the library was not doing enough to actively involve the library users. This was from the students' perspective.

4.7 Current acquisition processes for development of a Patron Driven Acquisition

(Schroeder, Wright & Murdoch, 2012) reported that Patron Driven Acquisitions shift the purchasing decision from the librarian to library users. Traditionally, libraries acquired information resources just in case needs arose. This meant that a considerably large portion of the collection

was under-utilized, which translated to poor usage of the already limited finances. In their article, (Crawford, Condrey, Avery & Enoch, 2020) argued that acquisition librarians had to switch their mode of acquisition from 'just in case' to 'just in time'.

In order understand the selection and acquisition process in place for developing a patron driven acquisition.

A few variables were explored:

- i. Factors considered during selection of information resources
- ii. Did library users submit of suggestions for information resources to be acquired their libraries?
- iii. Did the library acquire the information resources requested?

4.7.1 Factors considered during selection of information resources

The ultimate goal of this study was establishing the use of data mining tools in developing a patron driven collection. The study collected data from University librarians and Acquisition librarians to identify the various factors considered by librarians during collection development.

The findings were as summarized in Figure 4.5 below.

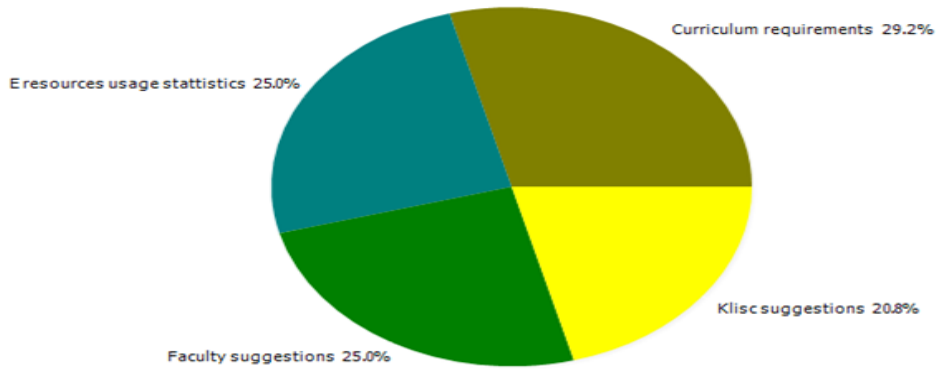


Figure 4.5: Factors considered during information resources selection process

Source: Field data (2022)

Figure 4.5 shows the responses obtained from acquisition librarians and university librarians on the considerations made during the selection and acquisition of information resources. Curriculum requirements was the major factor at 29.2%; the librarians defined this as the reading list recommended by the ministry of education for every course offered in the institution. These were the basic information resources and the bare minimum that were required in every faculty. Another factor the librarians considered was the statistics for usage of electronic resources which was at 25% of the total responses and tied with faculty suggestions. The librarians indicated that they analysed how the libraries' electronic resources were used and they were able to establish which subject areas were mostly used. They used this data during the selection process and they gave priority to the highly consulted information resources. The teaching staff was also consulted and they submitted recommendations for information resources that would be useful to their students' further reading. Finally, the librarians reported that they considered the Kenya Libraries and Information Services Consortium (KLISC) suggestions which were at 20%. The consortium comes

up with information resources for all the subject areas offered by their member libraries and these libraries are required to have them as part of their collection.

These findings are in congruence with (Ashikuzzaman, 2018), who reported that, as a guide to future acquisitions, librarians analyse the check-in/ check-out data of physical information resources, database usage and changes in curriculum. This provides the opportunity to assess the significance and worth of their library's collection. A study by (Kamau & Elegwa, 2021) further pointed out the need to have the faculty involved in the selection and acquisition process. However, faculty involvement often resulted in delay of the acquisition process. This was echoed by this study's respondents who felt like the faculty did not give weight to submitting their recommendations. The librarians pointed out that lack of faculty support often posed a challenge during acquisition especially when making decisions on which publishers to acquire resources from.

4.7.2 Submission of suggestions for information resources to be acquired

The study sought to determine whether post graduate students submitted suggestions of information resources to be incorporated in their academic libraries' collection. This data was collected librarians; both acquisition and university libraries. The results obtained were as shown in Figure 4.6 below.

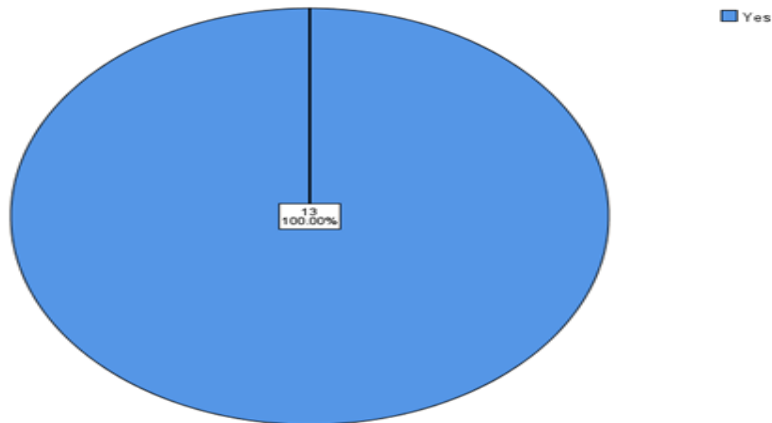


Figure 4.6: Librarians' response on Post Graduate students' submission of suggestions to their university libraries

Source: Field data (2022)

As shown above, all librarians admitted that their library users submitted suggestions of information resources they would like their libraries to acquire. 62% of the Post Graduate respondents also reported that they submitted requests for acquisition to their libraries. (Galbraith, 2020) stressed the 'power of the masses'. According to the study, librarians are able to achieve user satisfaction as well as maximize proper usage of the already limited funds allocated to the libraries when their users' input is put into consideration. The study noted that library users felt involved and appreciated whenever their libraries sought their opinions and contribution during the selection and acquisition process.

4.7.2 Did the library acquire the information resources requested

Having reported that library users submit suggestions of information resources to acquire, the study additionally wanted to establish whether the librarians acted on these requests. The data collected was summarized as shown on figure 4.7 below.

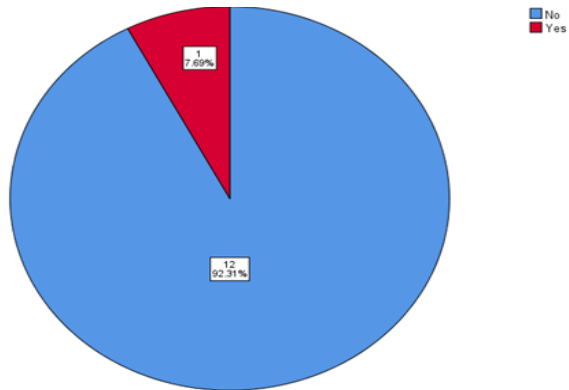


Figure 4.7: University libraries reaction towards information resources requested

Source: Field data (2022)

On whether the academic libraries took action on their users' suggestions, majority of the librarians responded no at 92.31% while 7.69% responded yes. This showed that the most of the user requests were not acted upon according to the librarians. The librarians indicated that they did not purchase information resources suggested by their users. Instead, they used inter-library loaning and borrowed the information resources from other libraries for a limited period of time.

4.7.3 User-centeredness of the library collection

The study sought the librarians' opinion on whether their academic library's collection was developed with their users' information needs as the centre of concern. The results are as shown below in figure 4.8

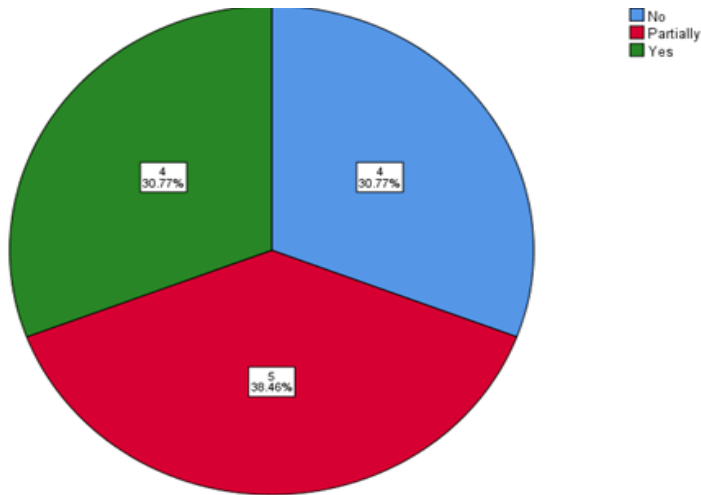


Figure 4.8: User-centeredness of the current library collection

Source: Field data (2022)

(Schroeder et al., 2010) defines a user centred collection as a library collection developed with the users as the centred of focus, where their input is appreciated and put into consideration and information needs are identified and acknowledged. From the data analysed, 30.77% of the responses collected from librarians indicated that their library's collection was user centred and the same percentage indicated that it was not user centred. 38.46% were stated that the library's collection was partially user centred. With this data, it was evident that the librarians were not confident enough on the level of user centeredness of their libraries' collection. This meant that their library users' information needs were not met and that the libraries did not maximize on the limited financial resources they have at their disposal. According to (Dahl, 2012), with unmet information needs questions the value of a library. In addition, the value of a library is no longer on how much information resources it has in its collection but on the importance of the collection to the library's users.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter gives a summary of what was found in the previous chapter. In addition, it draws conclusions from the study's findings and makes recommendations for future research.

5.2 Summary

The purpose of the study was use of data mining tools in determining patrons' information needs for collection development in selected academic libraries in Kenya. In order to reach this goal, the research was based on four goals.

- i. To establish various data mining tools available for collection development in selected academic libraries in Kenya.

From the data collected, the study established several data mining tools were available in libraries to facilitate collection development. The data mining tools available in the libraries were: library website, email, library live chat, library social media platforms; specifically, Facebook, phone calls and library help desk.

Based on the data collected, the researcher organized the data mining tools in terms of popularity as follows: the most used data mining tool was library live chat, followed by Facebook. library email, library website, phone calls and the least used tool was the library help desk.

- ii. To assess librarians' awareness of the existence of data mining tools for collection development in selected academic libraries in Kenya.

Having established that the libraries had several data mining tools, the study sought to assess whether the librarians were aware that these data mining tools could be used to identify the library

users' information needs. The data collected from librarians revealed that the librarians were aware that these data mining tools could be used to identify the categories of users their libraries serve and the information needs of these users. However, the librarians only named a few of the data mining tools and not the entire list as they had named them earlier. They reported to being aware of the following data mining tools for identification of their users' information needs: library website, email, library live chat, library social media platforms; specifically, Facebook, phone calls and library help desk.

- iii. To determine contribution of postgraduate students in the collection development process by use of data mining tools in selected academic libraries in Kenya.

The aim of this objective was to establish whether the library users in this study used data mining tools to contribute towards collection development in their library. The postgraduate students, who represented the library users in this study, reported that they used data mining tools to submit their requests and suggestions to their academic libraries. The postgraduate students submitted their requests through the library email, library live chat and phone calls. This data was in line with the responses given by the librarians in the second objective.

- iv. To assess the current acquisition processes for collection development in use by selected academic libraries in Kenya.

This objective sought to establish the factors put into consideration during the selection and acquisition process of information resources. Data was collected from acquisition librarians and university libraries. The respondents listed the following factors:

- a) Curriculum requirements - these are the minimum reading materials required by the ministry of education for each subject area covered by an academic institution

- b) Electronic resources usage statistics - librarians reported that they kept a record of how the libraries' electronic information resources were utilized. They collected details of the subject areas that were mostly used. This acted as a guide when renewing subscriptions to e-resources or making new purchases. E-resources with the highest usage were given preference since they were considered vital to the library users' information needs.
- c) Faculty suggestions -the librarians stated that teaching staff submitted requests for information resources to be used for further reading in their subject areas.
- d) Kenya Library and Information Services Consortium (KLISC) recommendations -according to the librarians, KLISC comes up with a list of information resources which all member libraries should to have in their collection for all the subject areas offered in their parent institutions.

From the data collected on this objective, it was clear that the library users were not involved in the collection development process.

5.3 Conclusion

This study yielded several conclusions. First, collaboration between acquisition librarians and library users is paramount if a library is to develop a user-centred collection. Involving information users in the selection process is essential for developing a patron-driven collection because librarians can identify the specific information needs of each category of their users. Secondly, acquisition librarians should incorporate the library users' suggestions of new acquisitions during the selection process. From the study, both the librarians and library users admitted that the users submitted requests but they were not considered at all during the selection and acquisition process. This leaves a lot to be desired since the librarians on their own cannot successfully determine their

users' information needs due to their diversity and ever changing information needs. The researcher also concludes that despite there being data mining tools in the libraries, which the librarians admitted to being aware of, they were not adequately used during collection development. The primary goal of data mining tools is to identify library users' categories, their subject areas and preferences of information resources. This therefore meant that failure of using this data mining tool for their intended purposes translated to developing a less useful collection than would have been developed had the librarians utilized the data mining tools to the maximum.

5.4 Recommendations

According to the results of this study, the researcher recommended the following:

1. The collection development process should be undertaken with the users' information needs in mind. Achieve this; the users ought to be given a chance to submit suggestions of information resources to be acquired.
2. Librarians should make an effort to incorporate their users' suggestions during the selection process. This will ensure that the libraries acquire information resources that are in response to their users' just in time information needs rather than hosting a collection for just in case an information need occurs. This will minimize wastage of the already limited financial resources.
3. Librarians should hold sessions with their users where they hold discussions on which ways the users would like to be involved in their academic libraries. This will enable librarians identify areas they could be weak and come up with ways of improving them in order to offer enhance during service delivery to their users.
4. From the data collected, librarians reported that they mostly used the library website to communicate to their users about new acquisitions. The researcher recommends that librarians

expand their communication channels in order to reach out to all categories of their users, especially those who rarely visit the library website.

5.5 Suggestion for further research

This study was conducted based on the use of data mining tools to identify users' information needs in order to develop a patron driven collection. It was carried out in selected academic libraries and limited to a specific user group (postgraduate students) in a specific subject area.

The researcher suggests that the research be carried out on a wider target group cutting across all subject areas.

REFERENCES

- Ajibade, P., & Muchaonyerwa, N. (2022). The importance of data mining, user information behaviour and interaction audit for information literacy. *Library Hi Tech News*, 40(4), 12–14. <https://doi.org/10.1108/LHTN-09-2022-0109>
- Alamri, W. A. (2019). Effectiveness of Qualitative Research Methods: Interviews and Diaries. *International Journal of English and Cultural Studies*, 2(1), 65. <https://doi.org/10.11114/ijecs.v2i1.4302>
- Ashikuzzaman, M. (2018, June 20). Acquisition Policy of Library Materials. *Library & Information Science Education Network*. <https://www.lisedunetwork.com/acquisition-policy/>
- Barnsbee, L., Nghiem, S., Circulatory, M., & Sup-, R. (2018). *Cost-effectiveness New and Emerging Testing Technology for Efficacy and Safety Evaluation of Personal Care Delivery Systems*.
- Carlsson, H., & Torngren, T. (2020). Understanding Library Users via Surveys and Other Methods—Best Practices for Evidence-Based Library Development. *Journal of Library Administration*, 60(8), 925–944. <https://doi.org/10.1080/01930826.2020.1820276>
- Chang, C. C., & Chen, R. S. (2006). Using data mining technology to solve classification problems: A case study of campus digital library. *Electronic Library*, 24(3), 307–321. <https://doi.org/10.1108/02640470610671178>
- Chitra, K., & Subashini, B. (2013). Data Mining Techniques and its Applications in Banking Sector. *International Journal of Emerging Technology and ...*, 3(8), 219–226.
- Cortês, C., Gabriel, G., & Martins, A. (2011, May 1). *Communication in the Library: The Role of Librarians*. <https://doi.org/10.13140/RG.2.1.3731.2407>

- Crawford, L. S., Condrey, C., Avery, E. F., & Enoch, T. (2020). Implementing a just-in-time collection development model in an academic library. *The Journal of Academic Librarianship*, 46(2), 102101. <https://doi.org/10.1016/j.acalib.2019.102101>
- Dahl, C. (2012). Primed for Patron-Driven Acquisition: A Look at the Big Picture. *Journal of Electronic Resources Librarianship*, 24(2), 119–126. <https://doi.org/10.1080/1941126X.2012.684557>
- Das, S. (2020). *Library is a Heart of an Institution*.
- Diwani, S. A., & Sam, A. (2017). *Data Mining Awareness and Readiness in Healthcare Sector: A case of Data Mining Awareness and Readiness in Healthcare Sector: A case of Tanzania. January 2014*, 1–8.
- Diwani, S., & Sam, A. (2014). Data Mining Awareness and Readiness in Healthcare Sector: A case of Tanzania. *Advances in Computer Science*, 3.
- Fallon, H. (2016). Librarian as Communicator: Case Studies and International Perspectives. *New Review of Academic Librarianship*, 22(2–3), 107–111. <https://doi.org/10.1080/13614533.2016.1216216>
- Faulk, N., & Crist, E. (2020). *A Mixed-Methods Study of Library Communication with Online Students and Faculty Members | Faulk | College & Research Libraries*. <https://doi.org/10.5860/crl.81.3.361>
- Galbraith, Q. (2020). *Letting the readers have a say. October 2018*. <https://doi.org/10.5860/crln.79.9.502>
- Ishtiaq Ahmed & Sundas Ishtiaq. (2021). Reliability and Validity: Importance in medical research. *Journal of the Pakistan Medical Association*, 71(10), 2401–2406. <https://doi.org/10.47391/JPMA.06-861>

- Jabbar, A. (2017). Research methodology. In *SpringerBriefs in Applied Sciences and Technology* (Issue 9783319654560). https://doi.org/10.1007/978-3-319-65457-7_3
- Kamau, G. W., & Elegwa, A. L. (2021). Factors influencing collection development process at the University of Nairobi Library. *Library Management*, 43(3/4), 207–217. <https://doi.org/10.1108/LM-09-2020-0127>
- Kassim, N. A. (2009). Evaluating users' satisfaction on academic library performance. *Malaysian Journal of Library and Information Science*, 14(2), 101–115.
- Khemakhem, S., & Boujelbene, Y. (2018a). Predicting credit risk on the basis of financial and non-financial variables and data mining. <https://doi.org/10.1108/RAF-07-2017-0143>
- Khemakhem, S., & Boujelbene, Y. (2018b). Predicting credit risk on the basis of financial and non-financial variables and data mining. *Review of Accounting and Finance*, 17(3), 316–340. <https://doi.org/10.1108/RAF-07-2017-0143>
- Kim, J. A. (2017). User Perception and Use of the Academic Library: A Correlation Analysis. *Journal of Academic Librarianship*, 43(3), 209–215. <https://doi.org/10.1016/j.acalib.2017.03.002>
- Kovacevic, A., Devedzic, V., & Pocajt, V. (2010). Using data mining to improve digital library services. *The Electronic Library*, 28(6), 829–843. <https://doi.org/10.1108/02640471011093525>
- Kulp, C., & Rupp-Serrano, K. (2007). Organizational approaches to electronic resource acquisition: Decision-making models in libraries. *Collection Management*, 30(4), 3–29. https://doi.org/10.1300/J105v30n04_02

- Morton, S. M. B., Bandara, D. K., Robinson, E. M., & Carr, P. E. A. (2012). In the 21st Century, what is an acceptable response rate? *Australian and New Zealand Journal of Public Health*, 36(2), 106–108. <https://doi.org/10.1111/j.1753-6405.2012.00854.x>
- Mugenda, O. M., & Mugenda, A. G. (2003). *Research methods: Quantitative & qualitative approaches (Vol. 2, No. 2)*. Nairobi: Acts press. (n.d.).
- Nicholson, S., Hwang, S. Y., Keezer, P., O'Neill, E. T., Silwattananusarn, T., Njagi, P. R., Okwu, E., Echem, M. E., Mickelsen, A., Goyal, M., Vohra, R., Forcht, K. A., Cochran, K., Siguenza-guzman, L., Saquicela, V., Cattrysse, D., Chitra, K., Subashini, B., Renaud, J., ... Kim, J. A. (2015). General Conference. *DESIDOC Journal of Library and Information Technology*, 2(2), 13–24. <https://doi.org/10.1002/meet.1450400184>
- Nzivo, C. N. (2012). User perception on library services and information resources in Kenyan Public Libraries. *Library Review*, 61(2), 110–127. <https://doi.org/10.1108/00242531211220744>
- Okogwu, F. I. (2018). *Challenges of Collection Development of Electronic Resources in University Libraries in South East Nigeria*.
- Okogwu, F., & Ozioko, E. (2018). *Challenges of Collection Development of Electronic Resources in University Libraries in South East Nigeria*.
- Oyeleye, A. O. (2015). *Acquisition patterns in academic libraries: A case study of Michael Otedola College of Primary Education , Noforija Epe Lagos State*. 7(2), 40–46. <https://doi.org/10.5897/IJLIS2014.0438>
- Rath, M. S., & Rath, P. (2014). Collection Development in Libraries: Challenges and Solutions for Library Professionals. *Pearl : A Journal of Library and Information Science*, 8(1), 53. <https://doi.org/10.5958/j.0975-6922.8.1.009>

- Renaud, J., Britton, S., Wang, D., & Ogihara, M. (2015a). Mining library and university data to understand library use patterns. *The Electronic Library*, 33(3), 355–372. <https://doi.org/10.1108/EL-07-2013-0136>
- Renaud, J., Britton, S., Wang, D., & Ogihara, M. (2015b). Mining library and university data to understand library use patterns. *Electronic Library*, 33(3), 355–372. <https://doi.org/10.1108/EL-07-2013-0136>
- Schroeder, R., Wright, T., & Murdoch, R. (2010). *Patron Driven Acquisitions: The Future of Collection Development ?*
- Schroeder, R., Wright, T., & Murdoch, R. (2012). Patron Driven Acquisitions: The Future of Collection Development? *Anything Goes*, 168–172. <https://doi.org/10.5703/1288284314834>
- Scout, H. (2022). *Customer Feedback: Why It's Important + 7 Ways to Collect It - Help Scout*. <https://www.helpscout.com/blog/customer-feedback/>
- Sengupta, S. (2017, January). *Applications of Data Mining in Library & Information Centres: An Overview* (Journal Article (Paginated) 01). *International Journal of Current Research*. <https://doi.org/10.1/Data%20Mining%20IJCR%20Jan2017.pdf>
- Shah, U. A., Khan, M. N., Ullah, T., & Yar, M. S. (n.d.). *User Satisfaction on Library Resources and Services; a Case Study of Government Degree College Sabir Abad (Karak), Khyber Pakhtunkhwa*.
- Sharma, G. (2017). Pros and cons of different sampling techniques Gaganpreet. *International Journal of Applied Research*, 3(7), 749–752.
- Siddiqui, M. A. (2003). Management for change in acquisitions in academic libraries. *Electronic Library*, 21(4), 352–357. <https://doi.org/10.1108/02640470310491577>

- Silwattananusarn, T., & Tuamsuk, K. (2012). Data Mining and Its Applications for Knowledge Management: A Literature Review from 2007 to 2012. *International Journal of Data Mining & Knowledge Management Process*, 2. <https://doi.org/10.5121/ijdkp.2012.2502>
- Simkus, J. (2023, July 31). *Pilot Study in Research: Definition & Examples*. <https://www.simplypsychology.org/pilot-studies.html>
- Smit, B., & Onwuegbuzie, A. J. (2018). Observations in Qualitative Inquiry: When What You See Is Not What You See. *International Journal of Qualitative Methods*, 17(1), 1609406918816766. <https://doi.org/10.1177/1609406918816766>
- Sohrabi, B. (2019). *A predictive analytics of physicians prescription and pharmacies sales correlation using data mining*. 13(3), 346–363. <https://doi.org/10.1108/IJPHM-11-2017-0066>
- Sohrabi, B., Raeesi Vanani, I., Nikaein, N., & Kakavand, S. (2019). A predictive analytics of physicians prescription and pharmacies sales correlation using data mining. *International Journal of Pharmaceutical and Healthcare Marketing*, 13(3), 346–363. <https://doi.org/10.1108/IJPHM-11-2017-0066>
- Strong, J., & Galbraith, Q. (2018). *Letting the readers have a say: Crowd theory in collection development* | Strong | College & Research Libraries News. <https://doi.org/10.5860/crln.79.9.502>
- Taherdoost, H. (2017). Determining sample size; How to calculate survey sample size. *International Journal of Economics and Management Systems*, 2(2), 237–239.
- Taherdoost, H. (2021). Data Collection Methods and Tools for Research; A Step-by-Step Guide to Choose Data Collection Technique for Academic and Business Research Projects. *International Journal of Academic Research in Management (IJARM)*, 10(1), 10–38.

- Warokka, A. (2020). Data Mining, Financial Performance and Financial Decisions on Business: Evidence from Digital Marketing. *International Journal of Psychosocial Rehabilitation*, 24(1), 1160–1168. <https://doi.org/10.37200/IJPR/V24I1/PR200217>
- You, S., Joo, S., & Katsurai, M. (2022). Data mining topics in the discipline of library and information science: Analysis of influential terms and Dirichlet multinomial regression topic model. *Aslib Journal of Information Management*, 76(1), 65–85. <https://doi.org/10.1108/AJIM-05-2022-0260>
- Zong, C., Xia, R., & Zhang, J. (2021). *Text Data Mining*. Springer. <https://doi.org/10.1007/978-981-16-0100-2>
- Zook, M., Graham, M., Ojanpera, S., Straumann, R., & De Sabbata, S. (2017). Engagement in the knowledge economy: Regional patterns of content creation with a focus on Sub-Saharan Africa. *Information Technologies and International Development*, 13(2017), 33–51.

APPENDICES

Appendix I: Informed Consent

My name is Margaret Wamuyu Kinyua, a Masters student from Kenyatta University. I am conducting a study titled "Use of Data Mining Tools in Determining Patrons' Information Needs for Collection Development in Selected Academic Libraries In Kenya." The information will be used to establish how selected academic libraries employ various data mining techniques in determining the information needs of their patrons in order to facilitate development of a patron-driven collection. The findings of this study will be used by librarians as an educational tool on the various data mining techniques available and how to utilize them to sort through their huge chunks of data to collect information relevant to the acquisition process.

This study will also provide a baseline to librarians who would want to incorporate data mining in their acquisition process. In order to employ data mining techniques effectively, specialized training is required.

Librarians can use this study when making requests for funding by illustrating to their parent organizations or financiers the importance of data mining in developing a user-driven collection.

In addition, this study will demonstrate the need for academic libraries to provide user-friendly platforms where their patrons can post feedback, make enquiries and suggestions etc. Since data collected through data mining is that which has been previously provided by library patrons.

Procedures to be followed

Participation in this study will require that I ask you some questions which will be recorded in a questionnaire.

Voluntarism

You have the right to refuse participation in this study. You will get the same services and care whether you agree to join the study or not and your decision will not change the care you will receive. Please remember the participation in this study is voluntarily. You may ask questions related to the study at any time.

You may refuse to respond to any questions and you may stop an interview at any time. You may also stop being in the study at any time without any consequences to the services you receive here or any other organization now or in the future.

Discomforts and Risks

Some of the questions you will be asked are on intimate subject and may be embarrassing or make you uncomfortable. If this happens, you may refuse to answer these questions if you so choose. You may also stop the interview at any time.

Benefits

If you participate in this study, you will help us to learn how to improve provision of information services to users, with the users’ needs in mind. This will go a long way in improving users’ satisfaction and at the same time maximize on the financial resources allocated to the library.

Reward

There are no rewards or any payment to you if you participate.

Confidentiality

Your personal details will not be recorded on the questionnaire or the interview schedule. The data collected will be accessed and used only by the researcher.

Contact Information

If you have questions about the study call Margaret Kinyua (researcher) 0705124556 or Supervisor, Dr. Rose Njoroge 0722978982.

However, if you have questions about your rights as a study participant: You may contact Kenyatta University Ethical Review Committee Secretariat on chairman.kuerc@ku.ac.ke,

Participant’s statement

The above information regarding my participation in the study is clear to me. The study has been explained to me and I have been given a chance to ask questions and my questions have been answered to my satisfaction. My participation in this study is entirely voluntary. I understand that my records will be kept private and that I can leave the study at any time.

Name of Participant: _____

Signature or Thumbprint

Date

Name of Representative/Witness (where necessary)

Relationship to Subject

Investigator’s statement

I, Margaret Kinyua, have explained to the volunteer in a language s/he understands, the procedures to be followed in the study and the risks and benefits involved

Margaret Kinyua

02/04/2025.

Appendix II: Introductory Letter

Margaret Wamuyu Kinyua,
Kenyatta University, School of Education,
Department of Library and Information Science,
P.O BOX 43844-00100,
Nairobi, Kenya.

To

Name of institution

Postal address

Dear Sir/Madam

RE: REQUEST TO COLLECT DATA IN THE LIBRARY

I am a postgraduate student at Kenyatta University, pursuing Master's degree in Library and Information Science. I am conducting research on use of data mining tools in determining patrons' information needs for collection development in selected academic libraries in Kenya. This research will collect data from the library management, acquisition librarians and Post Graduate students.

The purpose of this letter is to seek your consent to take part in the research. I will use interviews to collect data from the library management and questionnaires will be issued to the librarians and Post Graduate students. In adherence to the MoH Covid-19 guidelines, the researcher intends to administer the questionnaires in soft copy through the participants' corporate emails. On completion of the study, the findings and recommendations will be shared with the library. Information provided by the participants will be only used to accomplish the goals of the study and confidentiality will be maintained.

Thank you.

Yours Faithfully,

Margaret Kinyua.

Appendix III: Interview Schedule for University Librarians

My name is Margaret Wamuyu Kinyua, a post-graduate student at Kenyatta University pursuing a Master's degree in Library and Information Science. I am carrying out a study on the **use of data mining tools in determining patrons' information needs for collection development in selected academic libraries in Kenya.**

I kindly request you to spare a few minutes for this interview.

Your input will be highly appreciated and will be treated with confidentiality and meant for the purpose of the research only.

1. How does the university library decide on what information resources to acquire?
2. Which are the acquisition practices currently in use by your university library.
3. What policies are in place that facilitate user involvement during the information resources selection process in your university library?
4. Do your users submit suggestions of information materials to be acquired in the library?

Yes	[]
No	[]

b) If yes, how do they submit their suggestions?

- i. Through the library help desk
- ii. Library website

- iii. Library live chat
- iv. Phone calls
- v. Through their departments
- vi. Library email
- vii. Library social media accounts

Others, kindly state them

5. Are there special measures taken by your university library in order to achieve a Patron Driven Acquisition?

Appendix IV: Questionnaire for Post Graduate Students

My name is Margaret Wamuyu Kinyua, a post-graduate student at Kenyatta University pursuing a Masters in Library and Information Science. I am carrying out a study on the **use of data mining tools in determining patrons' information needs for collection development in selected academic libraries in Kenya**. You have been selected together with others to participate and contribute to this study.

I kindly request you to assist me by filling in the questionnaire honestly and freely by ticking or briefly explaining where necessary. Your input will be highly appreciated and will be treated with confidentiality and meant for the purpose of the research only.

1. Kindly indicate your institution

2. In your opinion, do the current information resources in your university library satisfy your information needs?

3. If your answer to number 2 is no, what suggestions would you offer to your university library in order to improve satisfaction?

4. In the course of your study, how do you reach out to your university library to make requests, complaints, questions or suggestions to your university library regarding the available information resources?
 - i. Library help desk []
 - ii. Library e-mail []
 - iii. Phone calls []

- iv. Library chat []
- v. Library's social media channels (kindly specify the channel)
- Facebook []
- LinkedIn []
- b. In your opinion, are the platforms listed above user friendly?
- i. Yes []
- ii. No []
5. How does your university library to notify you of newly acquired information resources?
- i. Through the website []
- ii. By sending emails []
- iii. Through the library social media platforms []
- iv. By holding exhibitions []
- v. Through the department []
6. Has your university library involved you during selection information resources to acquire?
- i. Yes []
- ii. No []
- b. If yes, how did you submit your contributions?
- i. Library email []
- ii. Through phone calls []
- iii. Library website, []
- iv. Library live chat []

7. In your opinion, does university library incorporate your input when acquiring information resources?

Yes

No

8. If no, would you advise your library to incorporate your input in the selection and acquisition process?

Yes

No

THANK YOU.

Appendix V: Questionnaire for Acquisition Librarians

My name is Margaret Wamuyu Kinyua, a post-graduate student at Kenyatta University pursuing a Master's degree in Library and Information Science. I am carrying out a study on the **use of data mining tools in determining patrons' information needs for collection development in selected academic libraries in Kenya.**

I kindly request you to assist me by filling in the questionnaire honestly and freely by ticking or briefly explaining where necessary.

Your input will be highly appreciated and will be treated with confidentiality and meant for the purpose of the research only.

1. Kindly state the name of your institution.

2. How does your university library identify the users' information needs?

3. What factors does your library put into consideration during the information resources selection process?
 - i. Curriculum requirements []
 - ii. Users' suggestions []
 - iii. Faculty suggestions []
 - iv. Circulation statistics []
 - v. E-resources usage statistics []

Others, kindly list them

4. In your opinion, is the library collection user centred?

i. Yes []

ii. No []

b) If yes, what would you attribute it to?

c) If no, what would you advise your library implement in order to ensure the collection is built focusing on your users' information needs?

5. Which data mining tools do you use to inform your users when the library acquires new information resources?

i. Library website []

ii. Users' emails []

iii. Phone calls []

iv. Library's social media accounts []

v. Liaison with departments []

Others, kindly list them

6. Do your library users submit suggestions of information resources to be acquired in the library?

i. Yes []

ii. No []

7. Which data mining tools are available for your users to submit their suggestions of information resources they would like added to the library collection?

- i. Library help desk []
- ii. E-mail []
- iii. Phone []
- iv. Library live chat []
- v. Library's social media accounts []
- vi. Through their departments []

Other (kindly list them)

8. Does your library acquire information resources requested by your users?

- i. Yes []
- ii. No []

10. Kindly state your opinion on whether your university library should or should not put users' requests into consideration during the acquisition process?

THANK YOU.


Appendix VI: Observation Checklist


My name is Margaret Wamuyu Kinyua, a post-graduate student at Kenyatta University pursuing a Master's degree in Library and Information Science. I am carrying out a study on the use of data mining tools in determining patrons' information needs for collection development in selected academic libraries in Kenya.

The researcher visited the library in person to check the availability of various data mining tools in the selected academic libraries and ticked against the checklist below.

Data mining tool	Available	Not available
Library website		
Library Face Book page		
LinkedIn		
Library email address		
Phone numbers		
Suggestion box		
Library live chat		
Library help desk		


Appendix VII: Research Permit


REPUBLIC OF KENYA


NATIONAL COMMISSION FOR
SCIENCE, TECHNOLOGY & INNOVATION

RefNo: 204109 Date of Issue: 09/November/2021


RESEARCH LICENSE




This is to Certify that Miss. MARGARET WAMUYU KINYUA of Kenyatta University, has been licensed to conduct research in Nairobi on the topic: USE OF DATA MINING TOOLS IN DETERMINING PATRONS' INFORMATION NEEDS FOR COLLECTION DEVELOPMENT IN SELECTED ACADEMIC LIBRARIES IN NAIROBI COUNTY, KENYA for the period ending : 09/November/2022.

License No: NACOSTI/P/21/13908

204109
Applicant Identification Number


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.