DATA MINING APPLICATION AND COMPETITIVE ADVANTAGE OF COMMERCIAL BANKS IN KENYA

By

DANCUN NDUNG'U NG'ANG'A

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT OF MASTER OF BUSINESS ADMINISTRATION DEGREE, FACULTY OF BUSINESS AND MANAGEMENT SCIENCES, UNIVERSITY OF NAIROBI

November, 2022

DECLARATION

This research proposal is my original work and has not been presented for any award in any other university.

Signed: ________________________ Date: ______31st November 2022___________

DANCUN NDUNG'U NG'ANG'A

D61/18862/2019

This research proposal has been presented for examination with my approval as the University Supervisor

Sign _____________ Date 11th December 2022

Mr. Joel Kiplangat Lelei

Lecturer

Department of Management Science and Project Planning

Faculty of Business and Management Sciences

DEDICATION

I dedicate this project to my family. A special feeling of gratitude to my loving parents, David and Margaret Ng'ang'a, whose words of encouragement and push for tenacity ring in my ears. Thanks for keeping the interest rates low on everything I owe you. And to my sisters Laura, Florence, and Mary, who have never left my side, you are very special. Knowingly and unknowingly, you led me to an understanding of some of the more subtle challenges to our ability to thrive. Finally, to all my friends who kept pushing and inspiring me to move on, thank you very much.

ACKNOWLEDGEMENT

“Neither man nor machine can replace its creator.”

First, I thank God for His provision and for enabling me to finish my coursework and project albeit many challenges and setbacks. My most profound appreciation goes to Mr. Joel Lelei for his time, effort, and understanding during the execution of this study. His vast wisdom and wealth of experience were indispensable throughout the research project. I could not have finished this journey without my defense committee, who generously provided knowledge and expertise.

This endeavor would also not have been possible without the support of my family, who encouraged and supported me all along. I will forever remain indebted to you. Lastly, I would be remis in not mentioning my close friends who cheered me to the finish line and fellow MIS specialization students with whom we worked together. You have been incredible!

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

IT – Information Technology

IS – Information Systems

DBMS – Database Management Systems

CRM – Customer Relationship Management

KDD – Knowledge Discovery in Databases

OLAP – Online Analytical Processing

CBK – Central Bank of Kenya

GDP – Gross Domestic Product

KCB – Kenya Commercial Bank

PwC – PricewaterhouseCoopers

RBV – Resource-Based View

TAM – Technology Acceptance Model

IoT – Internet of Things

ANOVA – Analysis of Variance

ABSTRACT

Nowadays, banks strive to gain a competitive edge in the marketplace thanks to globalization and cut-throat competition. Except for business processes, creating and applying a knowledge base has become a strategic tool to compete. The volatile progression of transient and stored data in financial institutions has stirred the need for new systems and computerized techniques to scrutinize vast amounts of data to find implicit and potentially insightful information. Commercial banks have discerned the value of data mining in building a knowledge base and utilizing it in strategic planning to survive and thrive in today's highly competitive market. This study aimed to examine the correlation between data mining application and competitive advantage in the banking sector in Kenya. The objectives of the study were to determine the relationship between data mining application and competitive advantage; to establish the extent of data mining espousal; to examine the drivers of data mining adoption, and to investigate the challenges facing the implementation of data mining. The study applied the descriptive cross-sectional research design, conducted a census survey on the 38 commercial banks in the country, and used structured questionnaires in data collection. The data was analyzed using regression analysis and descriptive statistics. The results reveal that commercial banks in Kenya use data mining for deviation detection, predictive modelling, database segmentation, and link analysis in different functional areas. The application helps banks unravel hidden knowledge from vast volumes of data to realize such competitive advantages as cost leadership, customer focus, channel optimization, and informed decision-making. Thus, the application of data mining leads to considerable realization of competitive advantage. The study also found successful management of information systems, ICT advances, operational necessity, and cost reduction to be critical drivers of data mining adoption. Moreover, the findings show that data security and privacy concerns are key challenges facing data mining application in the Kenyan banking sector. Further research should be conducted on other financial institutions to look into the bearing of data mining usage on performance and competitive advantage.

CHAPTER ONE: INTRODUCTION

4.1.Background of the Study

The automation of today's society has significantly enriched the capabilities of generating and gathering information from various sources. Specifically, the development of information technology (IT) has generated large databases and vast data in multiple areas. As a result, a considerable volume of data has engulfed every single facet of our lives. This volatile progression of transient or stored data has stirred the need for new systems and computerized techniques that can smartly help convert extensive data into valuable knowledge and information (Han, Pei, & Kamber, 2011). Unfortunately, the massive data available in numerous storage mechanisms is overwhelming, notwithstanding the establishment of structured databases and database management systems (DBMS). Today, vast data can be managed, varying from transactions and texts to images and intelligence. However, data retrieval is not sufficient for making sound financial decisions. This condition accentuates the need to make well-informed and worldly-wise decisions, which encompasses summarizing data, extracting relevant information from databases, and discovering hidden patterns and information (Oladimeji & Oladimeji, 2020). Consequently, this need has resulted in the establishment of an effective and thriving frontier known as data mining and its numerous applications.

The financial industry is a data-driven business considering that massive data is generated with every transaction and on a daily basis. Accordingly, data mining has become an indispensable asset for banks as a resource for monitoring fraud, analyzing credit worthiness, managing risk, and reducing customer churn. For example, Hasheminejad and Khorrami (2018) observe that banks have realized that customer relationship management (CRM) is a practical approach that can significantly assist in establishing a solid relationship with customers as well as enhancing their revenues and profits. CRM puts emphasis on retaining customers while moving from client acquisition and guaranteeing the adequate allocation of money, time, and administrative resources to these critical tasks (Chitra & Subashini, 2013). Many banks grapple with the challenge of retaining the most lucrative clients and how to achieve that cost-effectively. At the same time, firms need to implement an appropriate and flexible resolution swiftly. As the rivalry in the banking sector gets stronger, customer reliability is becoming one of the most targeted marketing goals of CRM (Zakirov & Momtselidze, 2015). Thus, commercial banks apply data mining to improve customer relationships, retain clients, and enhance their marketing strategies.

The banking sector has significantly benefitted from advances in digital technology. The idea of storing data at branches has paved way to integrated databases, subsequently leading to numerous channels that access bank accounts. Besides, Pulakkazhy and Balan (2013) note that banking systems are customer-oriented and technically viable with online transactions, automatic teller machines, automatic wire transfers, and cash deposit machines. The increased number of channels has also amplified transaction volume and the relatable information stored. Banks currently have enormous data sources in their computing storage systems, with the data having increased in terms of amount and dimensionality. With advances in data mining methods and competence, the process of identifying concealed knowledge from a significant amount of data has become the most valuable asset in commercial banks. Financial institutions have recognized the enormous potential of data mining use in their administrative and decision-making practices in such areas as investment banking, credit risk management, marketing, and fraud detection.

1.1.1. Data Mining

There are various definitions of data mining as an interdisciplinary subject. Mining is a term used to characterize the process of finding precious nuggets from a considerable volume of raw materials. Data mining is the discovery of unknown, tacit knowledge and valuable information from a substantial amount of raw data. According to Han et al. (2011), data mining is the computerization or identification of patterns and relationships characterizing knowledge discreetly recorded or stored in massive databases, the web, data warehouses, data streams, or other large data depositories. Pulakkazhy and Balan (2013) describe it as the process of deriving hidden and unknown insights from vast amounts of raw data. This premise implies that the knowledge must be new, inconspicuous, relevant, and applicable in the domain where the information is obtained. Some scholars relate data mining to another widely applied term, knowledge discovery from data (KDD), whereas others view it solely as a stage in the process of discovering new knowledge. Even so, many users commonly see it as an intelligent tool that helps accumulate and process vast amounts of data and derive sense of it. Data mining bridges a variety of technical areas, for instance, human-computer interaction, statistics, databases, and machine learning (Pechenizkiy et al. 2005). The aspect of data mining practices employed to identify and corroborate useful and relevant information is the crux of the knowledge discovery process.

Thanks to technological advancements, data can now be stored in diverse databases and information depositories. A data warehouse is a data repository architecture with multiple heterogeneous data sources structured in an integrated schema at a specific site to ensure effective decision-making by the management (Dhaouadi et al., 2022). The technology encompasses the cleaning and integrating of data as well as online analytical processing (OLAP). Whereas OLAP tools are essential in supporting a multidimensional inquiry and decision-making, other data-analysis instruments are needed for detailed scrutiny. For instance, data mining techniques that facilitate data clustering, anomaly or deviation detection, and the classification of variations in data over time have become indispensable. Therefore, the expanding need for information and insights call for the systematic development and usage of data mining methods to convert data resources into 'golden nuggets' of knowledge.

1.1.2. Competitive Advantage

Achieving a competitive advantage is the precursor to the significant performance of a firm. Its pursuit is very much at the core of the strategic management literature. Wang et al. (2011) note that competitive advantages exist when an organization can deliver comparable benefits as counterparts at a reduced cost (cost advantage) or surpass their quality (differentiation). According to Kotler (2000), competitive advantage is a firm's capacity to outperform its competitors in the same industry. These definitions denote that competitive advantage is a company's trait that aids it in outperforming the competition. It allows an organization to attain more significant margins than the rivals and create value for the firm and its stakeholders. A competitive advantage, therefore, must be challenging, if not difficult, to imitate. Firms have a competitive advantage when they successfully attract customers, generate more profits, or return more worth to their stakeholders than competing firms (Wang et al., 2011). Essentially, organizations attain a competitive advantage by adding worth to their products and services or effectively reducing their costs over their rivals.

Michael Porter identified three generic business-level approaches that outline the primary ways of organizing to strive in a product and service market. These strategies for ensuring a competitive advantage include cost leadership, differentiation, and focus (differentiation-focus and cost-focus)

(Porter, 1998). The cost leadership approach entails ensuring that the company becomes the lowest-cost producer in the industry. It is attained through extensive production, whereby organizations take advantage of economies of scale. Differentiation ensures the delivery of unique and high-quality products and services. A study of 65 manufacturing organizations recorded on the Indonesia Stock Exchange reveals that firms that employed the differentiation strategy had a superior financial performance than their rivals that did not employ it (Widuri & Sutanto, 2019). In a focus approach, a company concentrates on a narrow or niche section instead of generally targeting the entire market (Porter, 1998). For example, some commercial banks use the focus approach to gain a viable competitive edge by targeting small businesses or high net-worth individuals (Gitonga, 2015).

1.1.3. Commercial Banks in Kenya

Machiraju (2008) defines a bank as a certified financial organization that receives checking and savings deposits and issues loans. It deals with money and its equivalents and offers other money-related services. The dictionary meaning of a bank is a commercial establishment that deals with receiving, exchanging, lending, and safeguarding money and, in some cases, dispensing notes and managing other financial business. Saini and Sindhu (2014) aver that the banking system plays an indispensable part in modern economic development. As a vital constituent of the financial system, Haralayya and Aithal (2021) observe that banks facilitate commerce by collecting deposits and savings and lending them out to people and businesses. They also foster internal and international trade by availing guarantees and references at the behest of customers, enabling merchants to sell goods on credit. Commercial banks are also the backbone of gross domestic product (GDP) growth. Loans offered to individuals and investors by banks helps to improve the well-being of people and decrease economic stagnation significantly. Moreover, banks help in capital realization, generate employment opportunities, and promote the industrial boom (Haralayya & Aithal, 2021).

The Kenyan banking sector is vast and constitutes thirty-eight banks, including local and foreign banks with divisions, agencies, and other outlets across the country (CBK, 2022). It is primarily dominated by seven tier 1 commercial banks: Kenya Commercial Bank (KCB), Absa, Standard Chartered, Equity Bank, Diamond Trust Bank, Cooperative Bank, and NCBA Bank. More than 10 Kenyan banks, including KCB and Equity Bank, have subsidiaries in East and Central Africa and

South Sudan. The growing access to financial services is heightened by such technological ingenuities as agent banking, which allows commercial banks to make use and take advantage of third-party services to perform specific functions on their behalf. Due to the increased use of mobile money and the integration with official banking systems, many Kenyans increasingly access electronic monetary services (CBK, 2022). Consumers have also amplified the utilization of mobile banking platforms to access an extensive range of financial services.

The Central Bank of Kenya (CBK) is the main controller and regulator of Kenya's financial institutions, including commercial banks. It classifies banks by ownership. While local firms and individuals own some banks, others are held by foreign persons or entities. A different overall grouping is by nature, that is, commercial banks and microfinance institutions (CBK, 2022). The CBK further categorizes banks based on the value of their total assets. For instance, tier 1 banks are the largest with substantial assets and are unlikely to slump or fail business-wise. Tier 2 banks are midsize, whereas tier 3 comprises small-scale banks. Moreover, mortgage finance institutions and commercial banks are accredited and regulated in line with and true to the provisions of the Banking Act, the Central Bank of Kenya Act, and other prudential and practical standards enforced by the CBK (PwC, 2022). As influential entities in the banking segment, commercial banks must heed supervisory requirements controlling their provident standing and operational conduct to protect the overall reliability and stability of the monetary system.

1.1.4. Challenges of Data Mining Application

The application of data mining to complex real-world tasks is far from straightforward and involves many pitfalls. Zain and Rahman (2017) note that the five general issues of data mining application in organizations revolve around skills, technology, complex and incomplete data, privacy, and data security. The fact that data mining entails the search for unknown, hidden knowledge makes the outcome of its application difficult to predict and an uncertain endeavor. According to Jaseena and David (2014), big data mining involves multiple phases, each with its challenges, such as timeliness, heterogeneity, scale, privacy, and complexity issues. Data mining often leads to serious data security, privacy, and governance issues. It involves tons of sensitive information about customers. This premise implies that there is always the risk of hackers with malicious intentions accessing such information because there is no comprehensive tool that swipes out unwelcomed onlookers. Data mining also deals with huge volumes of data, which are noisy, incomplete, and heterogeneous. Due to instrument and human errors, large data quantities can be unreliable and inaccurate. With organizations storing data in diverse platforms in distributed computing environments, it is always challenging to consolidate all the data into a unified data repository (Jaseena & David, 2014). Notwithstanding these determined challenges, many other issues concerning data mining application exist. Organizations grapple with different issues during the actual data mining process, but successful application lies in overcoming them.

1.2. Statement of the Problem

The banking domain across the globe is increasingly realizing the need to employ data mining techniques to help them compete in today's highly turbulent and competitive market. Bhambri (2012) explored the industry status and found out that such banks as Chase Manhattan Bank in New York and Fllet Bank in Boston are increasingly using data mining to counter decreases in customer base, analyze customer profiles, and identify candidates for mutual fund offerings. Liu (2003) studied the China Merchant Bank to determine the implication of data mining application in banks and its interrelated importance. The author discovered that data mining empowers commercial banks to provide Internet-based monetary resolutions to their consumers through cost-effective services. The technology also enables financial institutions to provide services to remote markets and small banks with specialized niches to reach their consumers worldwide (Liu, 2003). Çaliş et al. (2015) studied the usage of data mining methods in the largest banks in Turkey and found that for commercial banks to attain a competitive advantage, they must understand their customers correctly and separate risky customers from others. These findings draw a connection between data mining application and the competitiveness of banks, highlighting the need to investigate the same in the Kenyan context.

Every firm aims to establish and preserve a competitive edge. This premise is especially relevant in the current competitive business environment in which companies must survive and gain a competitive edge against their rivals (Bal et al., 2011). They must also earn above-average returns to be prosperous. Competitive advantage is the centerpiece of a company's operation and performance in viable marketplaces. It stems from an organization's ability to establish worth for its consumers and exceed its operating costs (Porter, 1998). Thus, any profit-making firm must seek competitive advantage to survive, grow, and be profitable. Banks are not an exception to this fundamental premise. They have increasingly embraced technological and innovative solutions to open up efficiency delivery channels. Like many other financial institutions, banks exploit automated advances for competitive advantage, and data mining is one such solution. Microfinance institutions and banks are increasingly applying data mining techniques in such areas as fraud detection, due diligence analysis, revenue prediction, money laundering detection, quality control, customer retention, customer churn analysis, target marketing, and credit risk analysis to gain an edge amid cutthroat competition (Miller & Nyauncho, 2015).

Financial institutions across the globe have replaced traditional face-to-face customer contacts with electronic contact points to reduce the time and cost of their operations and improve their financial performance (Moin & Ahmed, 2012). The mechanization of monetary activities and the utilization of computerized software have totally revolutionized the basic notion of business and its operations. Nowadays, banks collect and store vast volumes of data that they gather every day, from transaction details, customer information, and transactional and risk profile records. However, existing statistical data analysis techniques are inadequate in managing and deriving insights from the vast amounts of data (Moin & Ahmed, 2012). This explosive growth has increased the application of novel technologies, such as data mining tools, to find hidden insights in the collected data. The discovery of knowledge from data helps banks develop strategies in such areas as customer satisfaction, risk management, and customer retention (Pascu, 2018). Banks are now realizing various benefits of data mining applications, which help them discover a competitive edge over their rivals. In this regard, there is a need to determine the extent and impact of data mining application on commercial banks' competitiveness and performance.

In Kenya, commercial banks have heavily invested in technology. Equity Bank, for example, disburses nearly 93% of its loans over mobile phones (Mugane, 2018). Notwithstanding the widespread adoption of technological solutions and data mining techniques, many commercial banks in the country are still determining the worth in vast datasets collected from mobile banking, e-banking transactions, and business operations. Some are still testing models that can, for example, predict customers' willingness to pay back their loans (Mastercard Foundation, 2018). This premise implies that the broad-based effect of data mining application on the competitive advantage in the Kenyan banking sector is relatively unknown. Besides, there is a literature gap in linking the usefulness of data mining to banks' competitiveness in the country. Existing literature, such as Mugane (2018) and Kagechu (2018), majorly focuses on the impact of big data on commercial banks in Kenya. Accordingly, this research gap underlines the need to determine the significance of data mining usage on banks' competitiveness. The overriding research question of the study is: what is the relationship between data mining application and the competitive advantage of commercial banks in Kenya?

1.3. Research Objectives

The overall objective of the research is to identify the relationship between data mining application and the competitive advantage of commercial banks in Kenya.

i) To establish the extent of data mining application in commercial banks in Kenya.
ii) To establish the drivers that motivated the adoption of data mining techniques in commercial banks in Kenya.
iii) To establish the challenges of adopting data mining in commercial banks in Kenya.
iv) To determine the relationship between data mining application and the competitive advantage of commercial banks in Kenya.

1.4. Value of the Study

This study is valuable to multiple stakeholders in the banking sector. First, there is a need to examine the extent of data mining application and its bearing on the competitive advantage in the banking sector. The results will shed light on how data mining influences the competitive advantage of banks. By focusing on banks, the study offers a point of reference for other financial institutions to benchmark against. Second, the results will help policymakers to gain value-added information concerning the role and significance of data mining technology in the competitiveness of banks. Third, the research contributes to the body of knowledge in data mining. Not so much research has been done about the connection between data mining application and the competitive advantage in the Kenya banking sector. Therefore, academicians and scholars can use the survey findings to support literary citations and identify new areas for future research. The study also enriches the academic knowledge repository on the current data mining trends and practices, drivers, challenges, and the extent of data mining use to the organizational performance of financial institutions. Lastly, investors can use the findings to evaluate the future of various commercial banks implementing strategic information technology solutions to add value and understand the worth of their investments.

CHAPTER TWO: LITERATURE REVIEW

2.1.Introduction

This chapter explores universally used theories to inform the application of information systems and ways of gaining competitive advantages in the market. It delves into the significance of data mining methods in the banking industry to gain an unbeatable competitive edge. In addition, drivers, challenges, and popular data mining technologies are explored accordingly. The books, journals, thesis, and dissertations referred to in this research paper focus on the objectives previously highlighted in the study and their relation to the operations of commercial banks in Kenya.

2.2. Theoretical Framework

2.2.1. Resource-Based View (RBV)

RBV contends that the possession of strategic resources is a source of competitive advantage that gives organizations an edge over their competitors. The thinking around the approach emerged in the 1980s and 1990s by Hamel and Prahalad (1996) and Barney (1986), and others. According to Barney (1991), strategic resources are imperfectly imitable, rare, valuable, and non-substitutable assets that are key sources of sustainable competitive advantage. The theorists note that such resources must ensure that a company performs its operations in a manner that results in low expenses, high sales, superior margins, or ways that enhance the firm's financial worth. Essentially, the model is an approach to attaining sustained competitive lead. It concentrates on the ideology of hard-to-mimic features of an organization as sources of competitive advantage. Companies use their resources to formulate strategies that improve their overall efficiency and performance, which can be pretty wide-ranging. Resources that are impossible to transfer or purchase and take a prolonged learning curve or significant change in enterprise-wide culture and setting will automatically be distinctive to the company and, thus, more problematic to emulate by competitors (Barney, 1991). Therefore, RBV assumes an inside-out perception or company-specific view concerning how corporations thrive or fail in the marketplace. Businesses often exploit unique resources to establish and uphold competitive advantages and higher performance.

The theory proposes that a resource is treasured if it aids businesses in capitalizing on their strengths and managing threats in the marketplace. RBV also claims that resources are useful when they help companies to identify or implement approaches that improve their proficiency and effectiveness (Barney 1991). Therefore, this theory enhances the understanding of why competencies are among the firms' indispensable assets in improving business performance. It provides valuable discernment on how monetary organizations can gain a competitive advantage through the application of distinctive digital resources. Financial companies, such as commercial banks and insurance firms, can apply data mining techniques and tools to get distinctive acumens that competitors cannot imitate. They can also attract and retain the best talent in the market who can combine their technical skills and business acumen to make empowered decisions and focus on things that matter most to the company.

2.2.2. Technology Acceptance Model (TAM)

TAM, which is largely credited to Fred Davis (1989), is an effective research model that deals with the prediction of the acceptability of an information system. It is widely utilized and substantiated by further research that explores the specific technology acceptance behavior in disparate information system concepts. TAM's main aim is to understand the processes sustaining the acceptance of technology to forecast the behavior of and give a theoretic description for the successful application of technology. In TAM, two facets, "perceived ease of use" and "perceived usefulness," are pertinent to computer usage behaviors. Davis explains that perceived usefulness is the potential user's subjective view that utilizing a particular system will increase job or life performance. On the contrary, perceived ease of use is the degree to which a likely user anticipates the proposed system to be easy, simple, or not demanding effort. According to the theory, apparent simplicity of use and usefulness are the main contributing factors to actual system use. Nonetheless, various external dynamics influence these two factors, including political, cultural, and social aspects (Davis, 1989).

TAM explains the general determinants of data mining application in the banking sector since it elucidates users' behavior in a broad spectrum of user populations and end-user computing technologies. If users in commercial banks in Kenya perceive data mining as valuable or vital in making informed decisions, then the adoption rate will be greater and vice versa. If the users perceive that data mining is easy to use, there is a higher likelihood of a high adoption rate. Various existing studies have investigated the use of business intelligence systems in firms using the TAM approach. Hart et al. (2007) examined the importance of cognitive and other dynamics on the apparent effectiveness of online analytical processing (OLAP). Huang, Liu, and Chang (2012) investigate how users observe and embrace data mining tools to increase their practical understanding of the business intelligence community. The study reveal that perceived ease of use and usefulness are integral elements that impact intentions to make use of data mining tools.

2.2.3. Schumpeter's Theory of Innovation

This framework asserts that business novelty is the main reason for economic dynamics – boosted investments and business variations. According to Schumpeter, the cyclical process is nearly complete in the outcome of innovation in industrial and commercial organizations (Schumpeter, 1934). By innovation, Schumpeter means the variations in the production of new products, industrial organization, and entering new markets, among others. According to the theory, innovation entails adopting new methods or creations in the administrative, cultural, and social medium. Firms can express their innovation as a function of entrepreneurship. Schumpeter splits the model into four components: innovation, invention, diffusion, and imitation. He claims that invention is the creation of ideas that can be adopted in the economy, while innovation is the conversion of new ideas into marketable products and is contingent on economic and technological conditions. Diffusion involves converting new ideas into products and spread out into potential markets. Imitation refers to when competitors copy an innovation, thereby making it cease to be new or novice (Śledzik, 2013). The theory emphasizes that while the creation stage has less impact; the diffusion and imitation process greatly influences the economic state.

In this study, data mining can be seen as the novelty while commercial banks that have embraced the technology are the entrepreneurs. According to Schumpeter's theory, entrepreneurs' activities, which draw upon the creations of inventors and scientists, establish entirely new employment, investment, and growth prospects. Accordingly, what matters in the Kenyan banking sector is not the innovation of data mining but rather its diffusion and imitation. As entrepreneurs, commercial banks are using data mining as a means of invention to attain competitive lead by targeting and acquiring more customers, detecting fraudulent activities, making customer-oriented strategies, identifying emerging trends, and providing segment-based products.

2.3. Data Mining Application in the Banking Sector

Commercial banks that use data mining tools are reaping huge profits and gaining considerable competitive advantage. Pulakkazhy and Balan (2013) claims that banking information systems encompasses massive amount of both historical and operational data. Accordingly, banks are increasingly recognizing the significance of customer information, which includes a large volume of personal information, risk profiles, transactional data, collateral details, demographic information, credit card usage patterns, compliance, anti-money laundering information, and so forth (Ranjbarfard & Ahmadi, 2020). Thanks to appropriate data mining tools, commercial banks provide personalized and tailored products and services to consumers after analyzing and summarizing the collected data from different perspectives and unraveling hidden patterns. Besides, banks make thousands of decisions daily concerning issues of fraud, credits, investments, and money laundering, among others. These decisions fall in different areas, including marketing, fraud detection, CRM, risk management, and mobile banking.

Banks use data mining tools for risk management and fraud detection. All important decisions a financial institution makes involves some element of risk. According to Pulakkazhy and Balan (2013), quantifying this risk streamlines and improves the risk management process and limits the threat of monetary loss to the bank. Understanding the ability of customers to refund their debt can significantly improve a credit manager's decisions. Jayasree and Balan (2013) note that credit authorization consultants in financial organizations apply data mining systems to identify the risk aspects in loaning and credit decisions by examining data concerning the nationality, reimbursement capacity, and among other details. The retail marketing unit applies data mining procedures to determine reliability and the behavior of credit card applicants when selling the credit cards. Moreover, banks and their customers can easily fall prey to fraudulent activities. Financial institutions apply various data mining tools to identify fraud. One notable technique is record screening modeling method that enables the identification of odd and unexpected cases and data patterns (Zakirov & Momtselidze, 2015). While using the tool does not inevitably show a doubtful activity, it provides information that leads to additional inquiry. Banks also use supervised and unsupervised models to detect fraud.

Data mining is applicable in all stages of the customer affiliation cycle: client acquisition, client value addition, and client retention. In the era of increased competitiveness, the customer is considered king. Accordingly, Jayasree and Balan (2013) observe that financial institutions, especially banks, recruit relationship managers who pay close attention to their customers. They use data mining to create customer profiling and group consumers in segments to deal with them accordingly. The collected information is used for numerous functions including making new marketing campaigns, creating market segments, undertaking risk analysis, and reviewing customer policies based on consumers' needs (Bhambri, 2012). Chitra and Subashini (2013) claim that banks use data mining to improve CRM by creating long-term affiliations with customers and increasing incomes and proceeds. This application is indispensable in the banking industry as it extends beyond attracting and retaining clienteles.

Another area of data mining application is commercial and customer product marketing. The marketing and sales departments in financial institutions make use of data mining algorithms to scrutinize current clients, identify their tastes and preferences, and determine how they can market their products and services (Jayasree & Balan, 2013). They also use the technology to evaluate past patterns, identify existing customer demands, and forecast client behavior on different products to achieve more business prospects and establish and enhance their market position. In a quest to position strategically in the competitive market, monetary institutions focus on recommending exceptional products with effective service via use of data mining techniques.

2.4. Drivers of Data Mining Application

With the computing power doubling every eighteen months and the price of computing and data store reducing, data mining adoption has become desirable for firms as a tool for enabling competitive advantage. Globalization has become increasingly widespread, and global competition among enterprises has risen significantly now than in the past (Huang, Liu, & Chang, 2012). Many organizations and managers worldwide use IT and IS to remedy business challenges. The adoption of IS tools enables them to be more accurate and effective in acquiring information or making decisions. Zide and Jokonya (2022) note that businesses are increasingly embracing budding technologies to ensure competences in their business operations, retain customers, reduce costs, and uphold a competitive advantage in their respective markets. According to Almoqren and Altayar (2016), various organizational, technological, and environmental factors have an effect on the espousal and usage of data mining tools in the banking sphere.

2.4.1. Organizational Factors

Organizational factors are enterprise-wide dynamics that are indispensable in promoting the success of an organization. They refer to descriptive measures within the organizational context that encompasses various aspects like human resources, higher management, and IT programs

(Almoqren & Altayar, 2016). In the wake of disparate data sources such as social networking platforms, mobile devices, and the Internet, data mining has become a revolution in a society obsessed with data with high capacity, velocity, veracity, and variety (Chan, 2013). According to Nunan and Domenico (2013), the administration level of information is among the perceptions confronting that companies must grapple with. Companies collect massive data and information requiring a high level of data exploration in terms of storage, examination, and retrieval.

The successful controlling of information in firms impacts the adoption and application of data mining technology. Khan et al. (2011) investigated the aspects facilitating the espousal of IT in the Saudi banking segment and established that a number of commercial bank executives fail to examine their policies, plans, and measures and embrace the latest technology. The study recommends that banks must evaluate their strategies at least every year, considering the variations and evolution of technology and prevailing trends. This premise implies that proper information system management can significantly influence the banking industry's acceptance and usage of data mining. Another significant organizational factor driving data mining application is top management support. Top management plays a critical role in supporting information systems management since they give financial organizations greater scope to formulate strategies and allows a more agile and flexible form in a business organization (Almoqren & Altayar, 2016). Coumaros et al. (2014) also exemplify how upper management practices in strategically deploying data mining inhibit the productive adoption of data mining.

Staffing and training influence the assimilation of data mining technology in organizations. Al Ruwaili et al. (2013) identified that personnel management in Saudi Arabia banks has transitioned from a conventional structure to a focus on talent-driven administration. This talent administration has increased the personnel training to enhance their skills and competencies and enable them to obtain new expertise and understanding. Besides, Kugel's (2013) investigation reveals that staffing and training are among the hurdles impeding the use of big data in monetary organizations. Organizational structure is also a driver of data mining technology in organizations. A centralized organization structure adversely impacts the usage of big data mining tools. According to a study by Askool and Nakata (2012), aspects affecting the amalgamation of Web 2.0 technologies with conventional CRM systems in Saudi banks are contingent on organization, environment, and technology framework. The researchers discovered that the formalization and centralization of organizational structure also play a vital impact in the decisions to assimilate social media, which can be a deliberation concerning big data mining.

2.4.2. Technological Factors

Technological aspect plays an indispensable impact in guaranteeing fruitful information systems management. This context entails dynamics, such as IT infrastructure, technology integration, technical resources, and technology readiness. First, wireless technology influences the espousal of data mining systems and practices. Ahmad et al. (2014) note that the period of wireless technology associates with numerous gadgets that are self-controlled data generating units. The authors observe that these devices are critical sources of big data, which means wireless technology is a considerable facet in the acceptance and execution of data mining technology. The transmission of wireless technology data occurs through indicators passed by electromagnetic waves over a transmission path, implying that it is a vital aspect of data mining technology (Ahmad et al., 2014). Second, the ability to decipher and shed light on unstructured data influences data mining implementation. Since unstructured textual data is created via web 2.0 and wireless technologies, there is a need to have powerful tools to find hidden knowledge and patterns. Coumaros et al. (2014) reviewed the obstacles affecting the fruitful embracing of big data from the practical perspective and identified the trouble of exploicating and decoding unstructured data as a notable problem. It emerges that big data platforms need potent data mining tools that contribute to the interpretation of unstructured data.

Third, IT infrastructure is a notable driver of data mining adoption in organizations. Companies grapple with various technological issues when they seek to extract predictive and descriptive knowledge from huge and voluminous data. These challenges encompass fast processing, reliable and efficient storage, and integration, all of which need an original infrastructure for scattered data mining and integration (Brezany et al., 2003). Fourth, the level of data control is an essential driver of the acceptance of data mining tools. With systems integration and wireless technology, organizations must coordinate and control data both from outside and inside sources. Askool and Nakata (2012) established that internal regulator is among the aspects affecting social media approval and a significant source of big data. Furthermore, there are various apprehensions associating internal controls with the time workers utilize on social media, consequently affecting overall workplace productivity (Askool & Nakata, 2012). Lastly, system integration also impacts the approval and application of data mining technologies. System technology is crucial to leadership in a company, especially when applying big data knowhows. Coumaros et al. (2014) discovered that the failure to pool data for the advantage of the whole organization is another notable challenge facing the adoption of data mining.

2.4.3. Environmental Factors

Environmental aspects are concerned with the setting where a firm conducts its operations, including information intensity, market competition, and government involvement (Almoqren & Altayar, 2016). Competition pressures and market strategies promote the approval of data mining technology. Taghva et al. (2011) aver that banks utilize data mining tools to achieve a competitive advantage for the reason that the technology has been embraced extensively in modern financial institutions to counter the challenge of increased rivalry. According to a survey on the influence of Twitter in the banking industry in Saudi Arabia, the results show that Twitter is an effective platform that promotes commercial banks' products and services and ensures practical dealings with clients. Further, information intensity also influences the adoption of the technology. The deployment of original marketing tactics to handle and cope with amorphous data, including Twitter-based records, coupled with the upsurge in the amount of other data sources, has resulted in a rise in the size of data – information intensity (Al Kalaf, 2011). Therefore, the increase in information intensity is a notable driver of the approval and application of data mining technology.

2.5. Data Mining and Competitive Advantage

Competitive advantage is the centerpiece of companies' operations in the current challenging commercial landscape. Choudhary et al. (2010) observe that there is fierce competition to find, nurture, and maintain trustworthy clients, enhance processes, adapt swiftly to fluctuating business environments, and find actional knowledge rapidly to grow business and increase profitability. A company that desires to establish a competitive edge must build and leverage its proficiencies (Bal et al., 2011). One of the critical ways of achieving a competitive advantage is creating original knowledge and transferring it to several organizational areas and levels. Financial organizations are increasingly embracing and installing data mining know-hows to manipulate collected data to augment their decision-making capabilities (Bal et al., 2011). They have realized the usefulness of utilizing data mining tools as a means of ensuring a competitive advantage. Brusilovsky and Brusilovskiy (2008) posit that data mining is now an important aspect of the business decision support system as it helps firms achieve a competitive advantage and reveal clients' data that cannot be gotten in other ways. This evidence shows that enterprise information obtained from data mining is a crucial success component for businesses intending to gain a competitive lead.

Data mining helps companies to establish and maintain a competitive advantage. Profit-seeking entities must identify a competitive advantage to survive, grow, and be profitable. According to the RBV theory, enterprise-wide capabilities and resources that are concurrently imperfectly imitable, valuable, rare, and non-substitutable form the crux of a company's viable competitive advantage (Bal et al., 2011). Various studies show that data mining can identify information and patterns that most traditional business analysis and statistical methods fail to deliver. Its application maximizes the worth of data warehouses by changing the expensive amount of data into treasured assets for upcoming strategic and tactical business expansion. A study by Afolabi and Adegoke (2014) found that data mining methods, including k-means clustering, enable firms to create competitive intelligence and make competitive advantage-based inferences. Similarly, Srivastava et al. (2019) claim that CRM programs facilitate customer information acquisition, which helps spread customized service depending on individuals' needs, thus ensuring client satisfaction and retention. In Saudi SMEs' research on the adoption of data mining for appropriate decision-making, Mian and Ghabban (2022) found that data mining, as well as training and advancement, plays an indispensable part in achieving competitive advantage via personnel role and knowledge management. Thus, data mining is an essential tool for creating and maintaining an organization's competitive advantage.

The application of data mining has numerous benefits for individuals, businesses, and society. From a business standpoint, Bal et al. (2011) note that it helps firms tailor their provision to particular customer wants, improve customer satisfaction, decide on product and service features critical to consumers, attract and retain customers, and customize marketing plans. It also helps distinguish marginal customers, identify new marketing opportunities, develop insights into changing customer requirements, and enhance productivity. In the banking sector, Khder et al. (2021) observe that banks that implement data mining strategies significantly profit and hold an advantage over those that do not. For instance, data mining helps banks detect and prevent fraud and meet customer needs. It enables them to gather data about customer needs and patterns, making it easier to flag instances where a pattern breaks to detect suspicious activities and protect customers' money. Bal et al. (2011) as well claim that the individual benefits of big data mining techniques involve reaching conclusions that are far beyond ordinary human scrutiny, reacting to customers promptly, understanding consumer needs, serving modified products, giving better products and services to customers, and improving customer relationship. At the societal level, data mining application facilitates the collection of intelligent information and the detection of criminal activities (Renushe et al., 2012).

2.6. Challenges of Adopting Data Mining

Data mining is critical in the banking sector but comes with its fair share of challenges. Since the technology is continuously evolving for large-scale handling of data, notable issues come along with scalability and automation. Almoqren and Altayar (2016) note that numerous vital trends contribute to the upsurge of generated data and associated data mining challenges. First, the development of traditional dealings related majorly to an extension of frequency and granularity that create numerous data forms. This trend has contributed to the disruption in the work setting, increased competitiveness, and upsurge in customers' demands. Accordingly, firms must counter these issues by making data mining tools more precise and comprehensive. Second, the rise in multimedia content in the current economy is another trend contributing to increased data generation in the banking sector. Third, the growth of IoT, which involves a system of interconnected computing gadgets, digital machines, objects, and individuals associated with interchange data with other devices and systems over the Internet, introduces new complexities in the handling of collected data (Almoqren & Altayar, 2016; Verma & Nashine, 2012). These trends have contributed to numerous data mining challenges in the financial industry. Fourth, Jaseena and David (2014) posit that incomplete data creates uncertainties during data analysis, which necessitates its management, and doing so correctly is also a challenge. Besides, the authors note that managing large and rapidly increasing data volumes is also a challenging issue. Conventional software tools cannot sufficiently handle the increasing data volumes. Data analysis, organization, modelling, and retrieval are also notable problems on account of the complexity and scalability of the data that need to be explored (Jaseena & David, 2014).

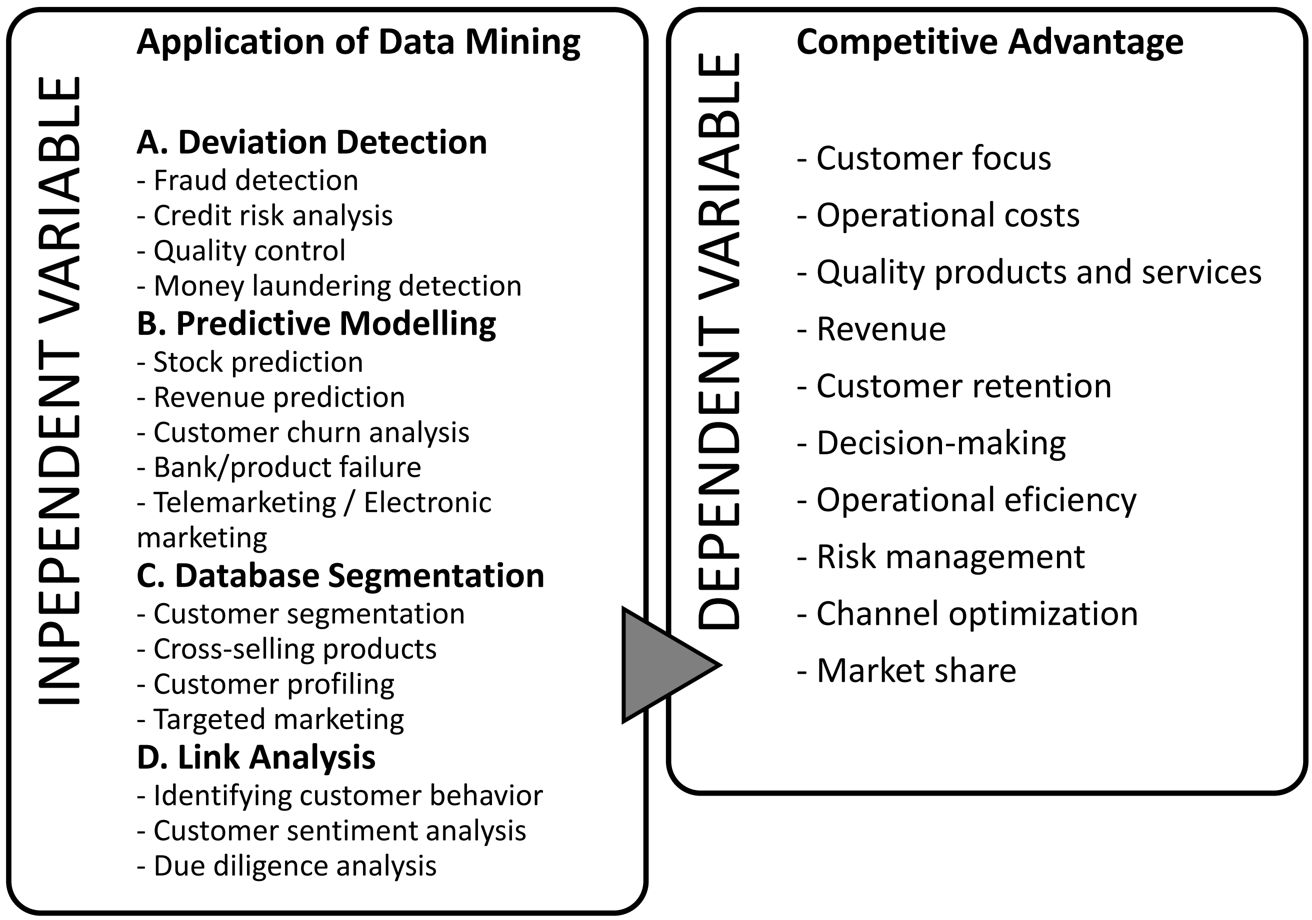
In a review of the data mining issues and challenges, Ragavi et al. (2018) note that data mining systems rely on databases to supply the raw input, which brings about complications since databases can be incomplete, dynamic, large, and noisy. The authors also note that other challenges arise from the irrelevance and inadequacy of the stored information. Missing and noisy data are pervasive in data mining applications. Ahmed et al. (2018) claim that flouting cases with missing values every so often causes the loss of information, which is contrary to building a good quality data mining model. Equally, Ragavi et al. (2018) observe that higher dimensionality in databases generates obstacles by raising the size of the search space to efficiently construct a model to perform data mining tasks.

Various organizational issues can hinder the adoption and application of data mining in organizations. Weiss (2009) avers that numerous administrative and non-technical concerns, such as political issues, can arise in data mining projects. Characteristically, the closer the data mining specialists are to the architects of the data mining endeavor, the larger the support they can anticipate, and the reverse is true. This premise suggests that the dearth of or inadequate top management support can undermine the successful adoption and execution of data mining initiatives. An observation from a group of data mining practitioners clearly sums up this proposition, “never underestimate the power of politics and turf battles” (Weis, 2009). Another key challenge is access to data mining specialists or domain experts. Data mining knowhow basically cannot compensate for lack of domain expertise. Having adequate access to domain connoisseurs is a lot more challenging than getting access to the appropriate data, even in instances where the experts reside in the organization, due to such reasons as limited time, job security, power play, lack of budget, and pride (Weiss, 2009).

2.7. Conceptual Framework

Figure 2.7.1 below depicts the theoretical outline employed in the research. It illustrates the connection between the application of data mining in Kenyan commercial banks and the resultant gain from competitive advantage.

Figure 2.7. 1: Conceptual Framework



CHAPTER THREE: METHODOLOGY

3.1. Introduction

This chapter covers the research methodologies applied in the survey. It explains the research design, the population and sampling design, the data collection techniques and research procedures, and the data analysis methods applied.

3.2. Research Design

The researcher used a descriptive cross-sectional research design to explore the use of data mining in the Kenyan banking sector. This type of research design seeks to obtain information to systematically describe a situation, population, or phenomenon. It entails examining and explaining subjects' behavior without influencing them in any way. De Lima (2011) avers that descriptive research design presents the distribution of current variables, irrespective of causation. Thus, it is appropriate for determining the correlation between data mining applications and banks' competitive advantage. The research design revealed the connection between the predictor and explanatory variables.

3.3 Population and Sample Design

A study population is the entire set of elements with common characteristics from which to choose the sample. The study's target population comprised of commercial banks in Kenya. At the point in time of the survey, there were 38 foreign and domestic banks in the country (CBK, 2022). The sampling design is applied to select items from a given population. Accordingly, the sample size encompassed all 38 commercial banks in Kenya, taking into account that the target population is considerably low. A census is best suited for this study because the population is noticeably small.

3.4. Data Collection

The researcher used structured questionnaires was utilized to gather data. The questionnaires were dispensed by the "drop and pick up later" approach, and online questionnaires were sent to the target participants. The questionnaires were directed to IT managers and managers dealing with the bank's strategies since they oversee and control data mining execution. The researcher targeted one respondent from each bank, bringing the total number of respondents to 38. The questionnaire had several parts. Section A covered the demographics of individual respondents and their organizations and included age, gender, duration of service in their position, number of employees, and value of the firm in terms of assets. Section B contained the areas of data mining applications in the targeted banks. Section C consisted of the drivers or contributing factors that prompted the adoption of data mining in commercial banks in Kenya. Section D included the competitive advantages resulting from data mining application in the organizations. Lastly, Section E covered the challenges of using data mining in the organization.

3.5. Data Analysis

After receiving the questionnaires, the data was checked and verified to validate its correctness and consistency. The data demographics were examined through percentages and frequencies. The areas of data mining applications, driving factors, and the challenges of data mining application were analyzed through means and standard deviation. Lastly, the relationship between data mining and competitive advantage was investigated using the following regression formula:

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3 + \beta 4X4 + \mathcal{E}$$

Where:

Y = Competitive advantage

B_0 = Constant

X_1 = Deviation detection

X_2 = Predictive modelling

X_3 = Database segmentation

X_4 = Link analysis

Ɛ = Error

CHAPTER FOUR: DATA ANALYSIS, FINDINGS, AND DISCUSSION

4.1. Introduction

This chapter describes and presents the analysis and results of the collected data and proceeds to make interpretations. The administered questionnaires contained five sections; demographic information, areas of data mining application, factors driving the espousal of data mining, challenges faced in the application of data mining, and the resultant competitive advantages from data mining usage. The chapter will interpret and discuss the results in line with pertinent theories and literature as established by other scholars.

4.2.Response Rate

The study was a census of all 38 banks operating in the country at the time of the investigation. The target respondents were IT managers and managers dealing with the bank's strategies and the execution of information systems. A total of 33 commercial banks responded, making the response rate 86.84%. This response rate constitutes the number of respondents in the census that successfully completed the survey. The response rate for this study is ideal, considering that a reasonable survey response rate is above average. Kothari (2006) observes that a 50% response rate is sufficient, 60% is good, and 70% and above is excellent. This premise means that the resultant response rate of 86.84 is perfect, thus, a reliable representation of the target population. As afore-stated, the response rate offers ample and satisfactory data to progress with the analysis.

Table 4.3. 1: Response Rate

Response	Frequency	Cumulative Percent
Responded	33	86.84
Did Not Respond	5	13.16
Total	38	100.00

Source: Author, 2022

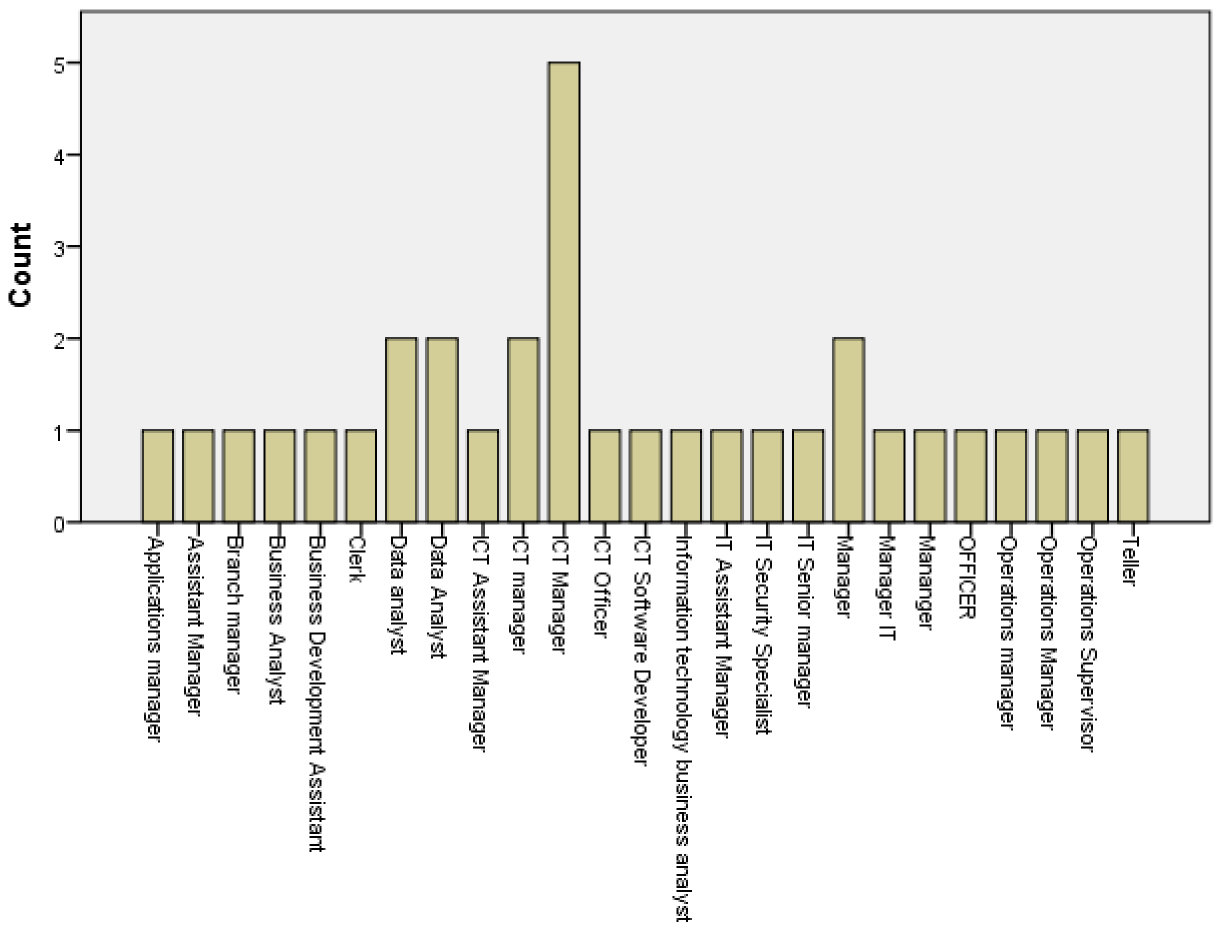
4.3.Demographics

The demographic section of this study provides background information about the respondents and the commercial banks. The sought information about the respondents includes gender, age, period employed in the position, position in the firm, and academic qualification. The organization demographic encompasses the number of people working in the surveyed firm, the length of the organization has been operating, and the organization's value in terms of total assets in Kenya shillings.

4.3.1. Position in the Organization

The majority of participants in the study were ICT managers, as shown in Figure 4.3.1 below. This observation implies that the study managed to get most of the target respondents to partake in the research.

Figure 4.3. 1: Position of the Respondents



Source: Author, 2022

4.3.2. Gender of the Respondents

The research aimed to establish the respondents' distribution in terms of gender. Table 4.3.2 illustrates the outcomes in this respect.

Table 4.3. 2: Gender of the Respondents

Gender		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	24	72.7	72.7	72.7
	Female	9	27.3	27.3	100.0
	Total	**33**	**100.0**	**100.0**	

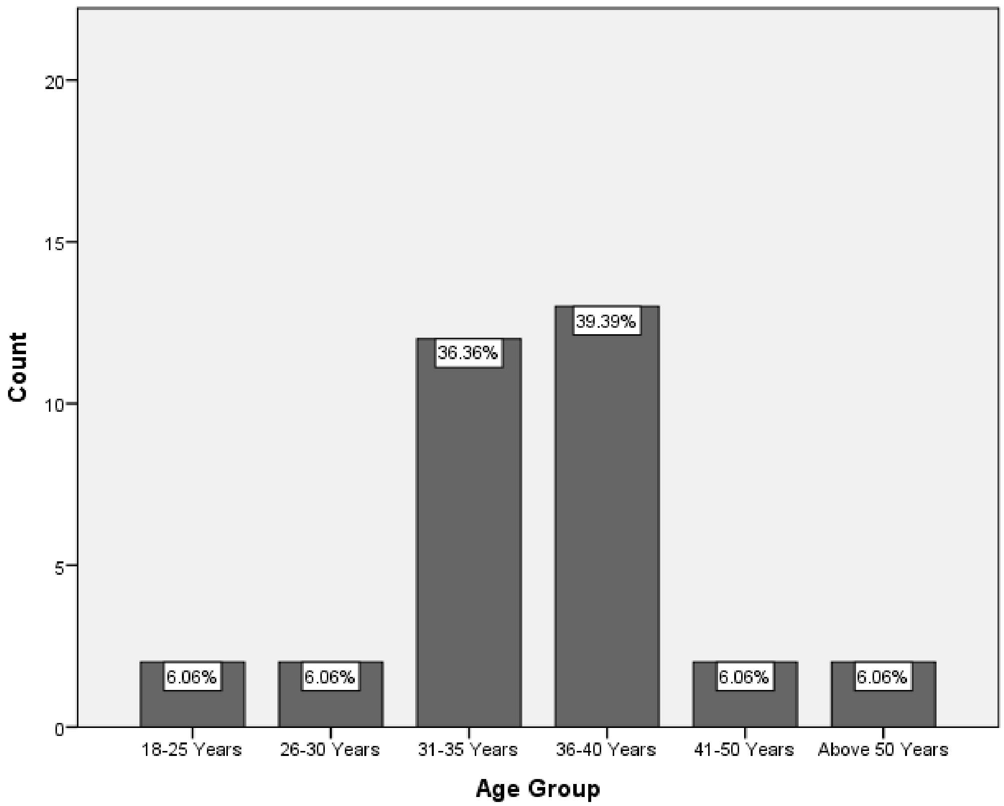
Source: Author, 2022

Of the respondents, 72.7% were male, whereas 27.3% were female. This finding shows that the organizations have both male and female personnel at the management level and are in charge of implementing information systems. It is fundamentally critical to notice that male personnel are more than female staff members, which implies that the former are majorly involved in data mining applications in their respective firms. The rationale for considering gender in the study was manifold. Gender is imperative in communication, administrative functions, and stakeholder predilections and engagements for the uptake or application of interventions. According to Tannenbaum et al. (2016), gender roles influence the way in which execution strategy functions, under what conditions, why, and for whom.

4.3.3. Age of the Respondents

The survey also studied the age distribution of respondents. Figure 4.3.3 shows how respondents are distributed in terms of age.

Figure 4.3.3. 1: Age of the Respondents



Source: Author, 2022

Many respondents were aged between 31 and 35 years with 36.4% and 36 and 40 years with 39.4%. The percentage for those aged between 18 and 25 years is 6.1%, 26 and 30 years is 6.1%, 41 and 50 years is 6.1%, and above 50 years is 6.1%. These findings show that the respondents aged between 30 and 40 are aware of technological advancements, generally, and data mining techniques, specifically. Moreover, they can constructively use technological advances to contribute to the competitive advantages of their organizations. These findings on age distribution are useful because an individual's knowledge and experience about a topic or subject are contingent on their age.

Table 4.3. 3: Age of the Respondents

Age		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-25 Years	2	6.1	6.1	6.1
	26-30 Years	2	6.1	6.1	12.1
	31-35 Years	12	36.4	36.4	48.5
	36-40 Years	13	39.4	39.4	87.9

	41-50 Years	2	6.1	6.1	93.9
	Above 50 Years	2	6.1	6.1	100.0
	Total	**33**	**100.0**	**100.0**	

Source: Author, 2022

4.3.4. Employment Period

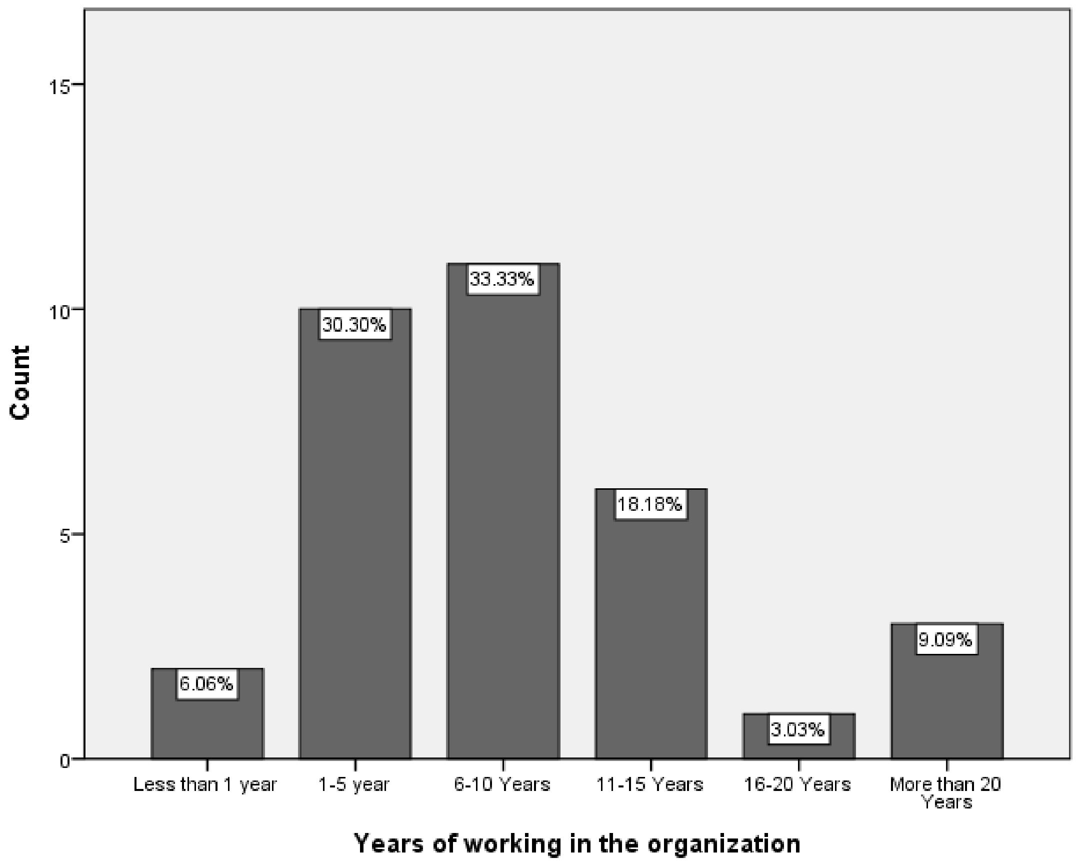
The information about the respondents' length or years of service is integral in assessing the dependability and reliability of the gathered data from a given population under study. The results in Figure 4.3.4.1 reveal respondents' working experience in commercial banks in Kenya. As regards the study, 33.3% and 30.3% of the participants had worked between 6 and 10 years and 1 and 5 years, correspondingly. 18.2% of the participants worked for a period of between 11 and 15 years. Moreover, 9.1% of the ICT managers worked for their organizations for more than 20 years, while 6.1% only worked for less than one year. Only 3% worked for a span of between 16 and 20 years.

Table 4.3. 4: Employed Period

Employed		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	Less than 1 year	2	6.1	6.1	6.1
	1-5 year	10	30.3	30.3	36.4
	6-10 Years	11	33.3	33.3	69.7
	11-15 Years	6	18.2	18.2	87.9
	16-20 Years	1	3.0	3.0	90.9
	More than 20 Years	3	9.1	9.1	100.0
	Total	**33**	**100.0**	**100.0**	

Source: Author, 2022

Figure 4.3.4. 1: Employed Period



Source: Author, 2022

4.3.5. Level of Education

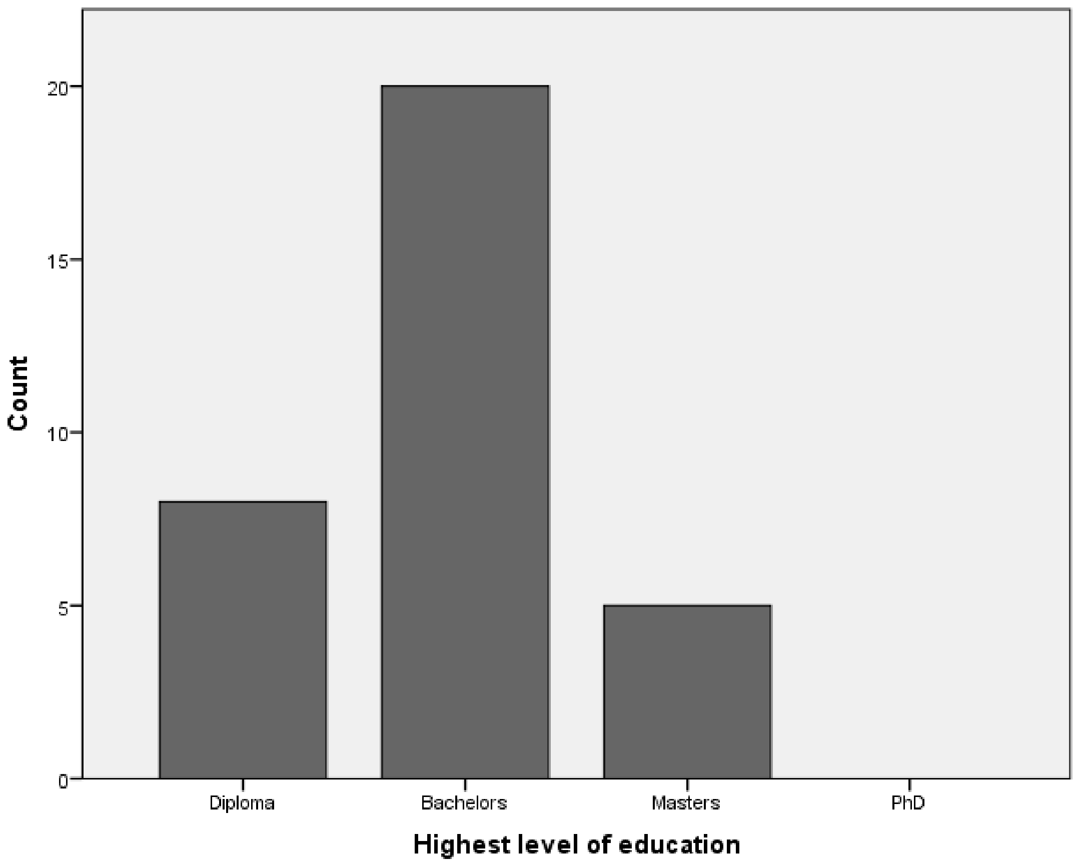
The survey inquired about the utmost achieved level of academic qualification from the respondents. 60.6% of the ICT managers and those in charge of bank's strategies and information systems implementation have bachelor's degrees. While 24.2% have diplomas, 15.2% have master's degrees, and none had a doctorate degree or other educational qualifications.

Table 4.3. 5: Level of Education

Education Level		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Diploma	8	24.2	24.2	24.2
	Bachelors	20	60.6	60.6	84.8
	Masters	5	15.2	15.2	100.0
	Total	**33**	**100.0**	**100.0**	

Source: Author, 2022

Figure 4.3.5. 1: Level of Education



Source: Author, 2022

4.3.6. Number of Employees

The study also investigated the number of staff working in commercial banks in Kenya. Accordingly, 45.5% of the respondents indicated that their organization used less than 1000 workforces. 15.25, 9.1%, 12.1%, 6.1%, and 6.1% of the respondents disclosed that their organizations had between 1001 and 2000, 2001 and 3000, 3001 and 4000, 5001 and 6000, and over 7000 employees, correspondingly. Additionally, 3% of the respondents work in banks with between 5001 and 6000 workforces and 6001 and 7000 employees. These results indicate that the organizations are sufficiently resourced to deal with the rising demand for financial services. The high number of banks with less than 1000 workers also denotes that organizations are increasingly adopting technological advances to cut costs by reducing the operational expenses incurred from salaries. Table 4.3.6 reveals the number of employees working in the surveyed banks.

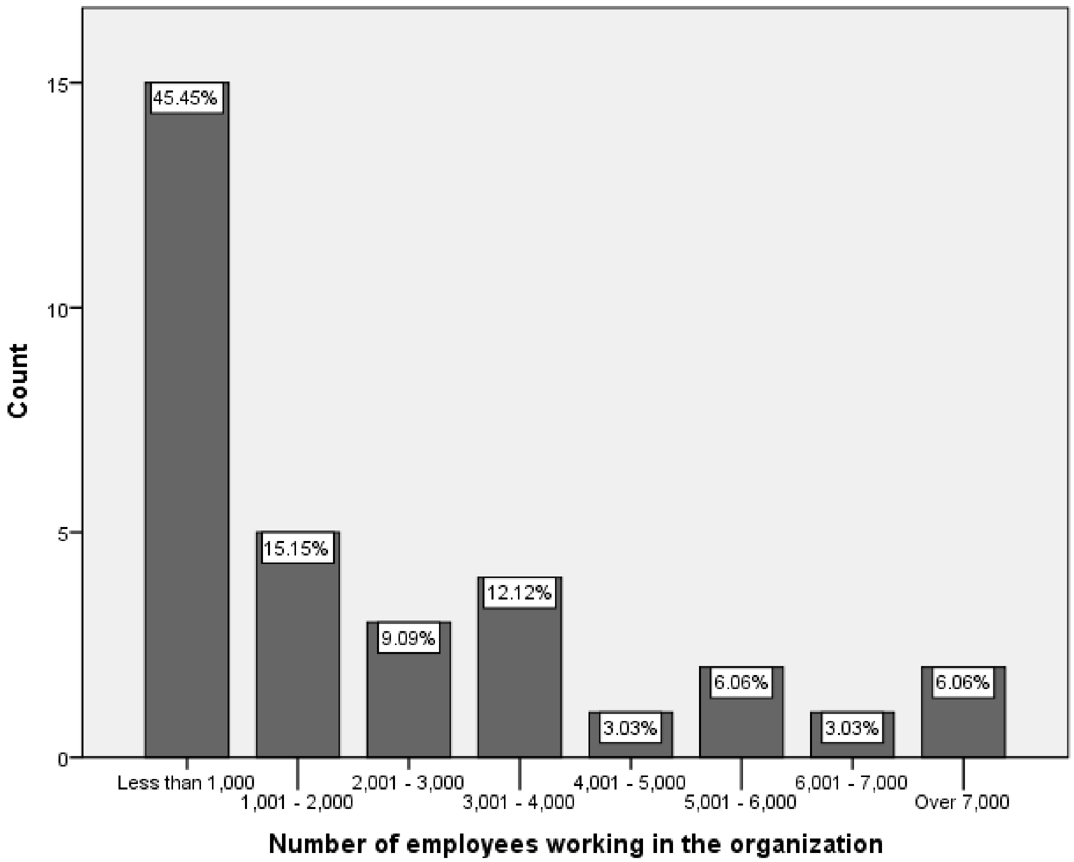
Table 4.3. 6: Number of Employees

No. of Employees	**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**

Valid	Less than 1,000	15	45.5	45.5	45.5
	1,001 - 2,000	5	15.2	15.2	60.6
	2,001 - 3,000	3	9.1	9.1	69.7
	3,001 - 4,000	4	12.1	12.1	81.8
	4,001 - 5,000	1	3.0	3.0	84.8
	5,001 - 6,000	2	6.1	6.1	90.9
	6,001 - 7,000	1	3.0	3.0	93.9
	Over 7,000	2	6.1	6.1	100.0
	Total	**33**	**100.0**	**100.0**	

Source: Author, 2022

Figure 4.3.6. 1: Number of Employees



Source: Author, 2022

4.3.7. Organizations' Length of Operation

The study further explored the period the commercial banks have been operational in Kenya. 72.7% of the respondents say that their organizations have operated for more than two years.

Whereas 24.2% of the respondents note that their organizations have been in business for between 11 and 20 years, only 3% say that theirs have operated for between 5 and 10 years. None of the commercial banks that participated in the survey operated for less than 5 years. These results demonstrate that many banks in Kenya have operated for more than two decades.

Table 4.3. 7: Organizations' Length of Operation

Length of Operation		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5-10 Years	1	3.0	3.0	3.0
	11-20 Years	8	24.2	24.2	27.3
	More than 20 Years	24	72.7	72.7	100.0
	Total	**33**	**100.0**	**100.0**	

Source: Author, 2022

4.3.8. Value of the Organization in Terms of Total Assets

Commercial banks in Kenya contribute to the GDP through their operations. They play a critical role in creating credit, which increases consumer spending, employment, and production, thereby boosting the economy. Banks also aid in accumulating capital, availing funds, mobilizing savings, implementing new technology, providing essential services, and ensuring the optimum use of resources. On this point, the study sought to ascertain the value of the firms with respect to their total assets. The results show that 36.4% of the respondents say their banks have less than 30 billion worth of assets, while 24.2% note that their organizations have assets valued between 30 and 100 billion Kenya shillings. 24.2% and 15.2% of the respondents noted that their organizations had total assets worth more than 300 billion and between 100 and 300 billion Kenya shillings. These results imply that the organizations have the capacity to embrace new technologies, generally, and data mining techniques, specifically to enhance their competitiveness in the industry.

Table 4.3. 8: Value in Total Assets

Total Assets		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Less than 30 billion	12	36.4	36.4	36.4
	30-100 Billion	8	24.2	24.2	60.6

	100 - 300 billion	5	15.2	15.2	75.8
	More than 300 billion	8	24.2	24.2	100.0
	Total	**33**	**100.0**	**100.0**	

Source: Author, 2022

4.4.Data Mining Application in Commercial Banks

The foremost objective was to examine the extent to which data mining is employed in Kenyan banks. Respondents revealed the degree to which their organizations applied data mining using the following scale: No Extent (1), Little Extent (2), Moderate Extent (3), Large Extent (4), and Very Large Extent (5). Means of between 3.30 and 4.67 were registered. The use of data mining in commercial banks falls within four dimensions: database segmentation, deviation detection, predictive modelling, and link analysis. The application of deviation detection has an average mean of 4.46, predictive modelling has an average mean of 3.82, database segmentation has an average mean of 4.35, and link analysis has an average mean of 4.28. Table 4.4 illustrates the use of data mining in the surveyed organizations.

Table 4.4. 1: Data Mining Application in Commercial Banks

Applications		**N**		**Mean**		**Std. Deviation**	**Variance**
		Valid	**Missing**				
Deviation Detection	Credit scoring & approval	33	0	4.42	4.46	0.663	0.439
	Fraud detection	33	0	4.48		0.667	0.445
	Quality control	33	0	4.39		0.496	0.246
	Money laundering detection	33	0	4.55		0.617	0.381
Predictive Modelling	Stocks prediction	33	0	3.64	3.82	0.822	0.676
	Revenue prediction	33	0	4.06		0.747	0.559
	Loan default prediction	33	0	4.39		0.609	0.371
	Bank/product failure estimation	33	0	3.30		1.212	1.468
	Telemarketing & electronic marketing	33	0	3.64		0.994	0.989
	Customer churn analysis	33	0	4.24		0.830	0.689
	Telemarketing outcomes prediction	33	0	3.48		0.939	0.883
Database Segmentation	Customer relationship management	33	0	4.48	4.35	0.508	0.258
	Customer segmentation	33	0	4.55		0.564	0.318
	Cross-selling products	33	0	4.24		0.708	0.502
	Customer profiling	33	0	4.24		0.708	0.502

	Targeted marketing	33	0	4.24		0.751	0.564
Link Analysis	Identifying customer behavior	33	0	4.48	4.28	0.755	0.570
	Customer sentiment analysis	33	0	3.70		1.015	1.030
	Due diligence analysis	33	0	4.67		0.595	0.354

Source: Author, 2022

These results reveal that many of the respondents note that their organizations greatly use data mining in their organizations. The high means of deviation detection and link analysis shows that commercial banks mainly use data mining technology in their respective areas. Specifically, it emerges that commercial banks in Kenya largely use data mining for due diligence analysis, customer segmentation, customer relationship management, money laundering detection, and fraud detection. The respondents also noted that their organizations moderately use data mining for a bank or product failure estimation and telemarketing outcomes prediction modestly, with means of 3.30 and 3.48, in that order.

4.5.Drivers of Data Mining Adoption

The second objective of this study was to establish the degree to which various driving forces inspired the adoption of data mining in the organizations. The researcher used the following scale: No Extent (1), Little Extent (2), Moderate Extent (3), Large Extent (4), and Very Large Extent (5). Means of between 4.36 and 3.21 were registered.

Table 4.5. 1: Drivers of Data Mining Adoption

Drivers	**N**		**Mean**	**Std. Deviation**	**Variance**
	Valid	**Missing**			
Successful management of information systems	33	0	4.24	0.663	0.439
ICT advances	33	0	4.33	0.736	0.542
Ability to interpret unstructured data	33	0	3.70	0.810	0.655
Level of data control	33	0	3.76	0.708	0.502
System integration	33	0	4.03	0.883	0.780
Increased competition	33	0	4.09	0.805	0.648
Operational necessity	33	0	4.30	0.847	0.718
Cost reduction	33	0	4.42	0.708	0.502
Market intelligence	33	0	4.36	0.653	0.426
Query complexity	33	0	3.21	0.960	0.922

Source: Author, 2022

Table 4.5 results imply that the respondents considered the factors as notable influencers in adopting data mining in their organizations. Of note is that cost reduction, the successful management of information systems, operational necessity, and market intelligence are key influential facets that encourage the use of data mining in banks with means of 4.42, 4.24, 4.30, and 4,36, respectively. Nonetheless, the majority of respondents consider query complexity, ability to interpret unstructured data, and level of data control as moderate driving factors toward adopting of data mining in their organizations.

4.6.Challenges Faced in the Use of Data Mining

To identify the challenges faced in using data mining in the banking sector in Kenya, the survey required participants to select the degree to which their organizations face various challenges in data mining use. The scale used was No Extent (1), Little Extent (2), Moderate Extent (3), Large Extent (4), and Very Large Extent (5). Table 4.6 indicate the extent to which commercial banks in Kenya grapple with various challenges in the use of data mining.

Table 4.6. 1: Challenges of Data Mining Use

Challenges	**N**		**Mean**	**Std. Deviation**	**Variance**
	Valid	**Missing**			
Data security	33	0	3.67	0.924	0.854
Employee resistance	33	0	2.09	1.042	1.085
Privacy concerns	33	0	3.58	1.001	1.002
Poor quality data	33	0	2.58	0.936	0.877
Lack of required infrastructure	33	0	2.48	1.064	1.133
Noisy & incomplete data	33	0	2.73	1.098	1.205
Regulatory risks	33	0	1.94	1.391	1.934
Shortage of skills & technical knowhow	33	0	2.55	1.034	1.068
Inadequate management support	33	0	2.21	1.193	1.422
Cost of implementation	33	0	2.00	1.299	1.688
Misalignment of IT & corporate strategy	33	0	2.61	1.088	1.184

Source: Author, 2022

With respect to the findings on Table 4.6, the respondents indicated that data security and privacy concerns are significant challenges of data mining usage in their organizations, with means of 3.67 and 3.58, respectively. They also revealed that regulatory risks, cost of implementation, and employee resistance are the most minor hindrances in using data mining in their organizations, with means of 1.94, 2.00, and 2.09, respectively.

4.7. Competitive Advantages

The study aimed to determine the degree to which the organizations achieved different competitive advantages due to the implementation of data mining. Accordingly, the respondents specified the degree to which their organizations realized different competitive advantages as a result of applying data mining. The following Likert scale was applied: No Extent (1), Little Extent (2), Moderate Extent (3), Large Extent (4), and Very Large Extent (5). Table 4.7 shows the extent to which commercial banks realized competitive advantages due to data mining technology.

Table 4.7. 1: Competitive Advantage

Competitive Advantages	N		Mean	Std. Deviation	Variance
	Valid	Missing			
Customer focus	33	0	4.64	0.549	0.301
Cost leadership	33	0	4.42	0.751	0.564
Quality products and services	33	0	4.55	0.711	0.506
Revenue	33	0	4.06	0.864	0.746
Customer retention	33	0	4.30	0.684	0.468
Informed decision making	33	0	4.67	0.540	0.292
Operational efficiency	33	0	4.73	0.517	0.267
Comprehensive risk management	33	0	4.55	0.666	0.443
Channel optimization	33	0	3.73	0.801	0.642
Market share	33	0	3.73	0.839	0.705

Source: Author, 2022

With respect to Table 4.7, most of the surveyed staff indicated that their organizations extensively realized such benefits as customer focus, cost leadership, quality products and services, increased revenues, customer retention, informed decision-making, heightened operational efficiency, and comprehensive risk management. In other words, the application of data mining substantially yielded numerous competitive advantages. The respondents also noted that their organizations experienced modestly improved channel optimization and market share because of data mining, with means of 3.73 and 3.73, respectively.

4.7.1. Relationship Between Data Mining and Competitive Advantage

The last and key objective was to investigate the connection between data mining application and the competitive advantage of banks in Kenya. Data mining application was quantified based on the degree to which commercial banks used data mining, while competitive advantage was quantified through the resultant competitive gains from data mining technology. Accordingly, the researcher applied a multiple regression analysis to establish the connection between the predictor variable (application of data mining) and the explanatory variable (competitive advantage). Table 4.7.1.1 illustrates the variables utilized in the regression model.

Table 4.7.1. 1: Variables Entered/Removed

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Link Analysis, Database Segmentation, Predictive Modelling, Deviation Detection[b]	.	Enter

a. Dependent Variable: Competitive Advantage

b. All requested variables entered.

Source: Author, 2022

Table 4.7.1.2 illustrates the fitness of the regression model. The coefficient of determination R^2, account for the relative amount of change in the dependent variable that a change in the independent variable can justify. Simply put, it represents the difference in the dependent variable (competitive advantage) that the statistical model predicts. Since R^2 is between 0 and 1, the model allows the prediction of the competitive advantage resulting from the change in the data mining application. The 0.958 coefficient of determination value is high, implying that the observations are close to the model's predictions, and the predictor variable describes 95.8% of the variance in the explanatory variable in the model. This outcome clearly indicates that the four independent variables (deviation detection, link analysis, predictive modelling, and database segmentation) contribute to approximately 95.8% of the competitive advantage. Hence, there is a strong correlation between data mining application and the competitive advantage of commercial banks.

Table 4.7.1. 2: Model Summary

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.979[a]	.958	.952	.0913

a. Predictors: (Constant), Link Analysis, Database Segmentation, Predictive Modelling, Deviation Detection

Source: Author, 2022

The ANOVA test results illustrated in Table 4.7.1.3 signals that the model was statistically significant, considering the small p-value below 0.05. These results imply that the application of data mining is a decent predictor of the competitive advantage of the commercial banks examined.

Table 4.7.1. 3: The Analysis of Variance (ANOVA)

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5.343	4	1.336	160.285	.000[b]
	Residual	.233	28	.008		
	Total	5.576	32			

a. Dependent Variable: Competitive Advantage

b. Predictors: (Constant), Link Analysis, Database Segmentation, Predictive Modelling, Deviation Detection

Source: Author, 2022

All the four independent variables were included in the model to examine how they act together to affect the competitive advantage of the studied organizations. This analysis was to allow the documentation of cumulative effects. With reference to the table generated by the SPSS program, the equation ($Y = \beta 0 + \beta 1 X 1 + \beta 2 X 2 + \beta 3 X 3 + \beta 4 X 4 + \mathcal{E}$) is translated to:

$Y = 0.304 + 0.229X_1 + 0.199X_2 + 0.229X_3 + 0.273X_4 + 0.163$.

Y is the competitive advantage, while X_1 is deviation detection, X_2 is predictive modelling, X_3 is database segmentation, and X_4 is link analysis.

Table 4.7.1. 4: Coefficient of Determination

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95.0% Confidence Interval for B Lower Bound	95.0% Confidence Interval for B Upper Bound
1	(Constant)	.304	.163		1.860	.073	-.031	.639
	Deviation Detection	.229	.109	.240	2.097	.045	.005	.452
	Predictive Modelling	.199	.084	.226	2.361	.025	.026	.371
	Database Segmentation	.229	.086	.251	2.667	.013	.053	.404
	Link Analysis	.273	.103	.306	2.645	.013	.062	.485

a. Dependent Variable: Competitive Advantage

Source: Author, 2022

The regression coefficient results in Table 4.7.1.4 denote that data mining application considerably influences the competitive advantage of the organizations studied. Considering the established regression model, which comprises all variables (deviation detection, predictive modelling, database segmentation, and link analysis) constant at zero, the competitive advantage of the commercial banks would be 0.304. The data-backed outcomes reveal that when all other predictor variables are zero, a unit rise in deviation detection brings about a 0.229 increase in the competitive advantage of commercial banks. A unit upsurge in predictive modelling causes a 0.199 rise in competitive advantage, while a unit increase in database segmentation generates a 0.229 increment in the competitive advantage of the organizations. Lastly, a unit increase in link analysis gives rise to a significant increment of 0.273 in the competitive advantage of the banks in Kenya.

With respect to the independent sample tests, the analysis reveal that there is no statistical significance between the two variables encompassing the predictor variable and the explanatory variable since the t-test values are more significant than 0.05, which is the significance level. The biggest disparity lies in the database segmentation variable (t-value = 2.667), while the most negligible variance was registered in the deviation detection variable (t-value = 2.097). These findings infer that database segmentation significantly contributes to the competitive advantage of banks, followed by link analysis, predictive modelling, and finally, deviation detection. Overall, the study results denote that the application of data mining has a positive connection with the competitive advantage of commercial banks in Kenya.

4.8. Discussion of Results

This study has established that commercial banks in Kenya aggressively deploy data mining tools and techniques in disparate functional areas to realize various competitive advantages. These results coincide with the study findings of Afolabi and Adegoke (2014) that reveal data mining methods and techniques, such as K-means clustering, help firms create competitive intelligence and make competitive advantage-based inferences. The results also mirror the findings of Srivastava et al. (2019) that showed CRM programs facilitate the acquisition of hidden customer information, which assists in enhancing customer service based on individual needs, thus ensuring the gratification and retention of customers. The findings confirm the revelations of Bal et al. (2011) that data mining application helps firms tailor their provisions to particular customer wants and needs, improve customer satisfaction, and decide notable products and services critical to consumers. Therefore, commercial banks that implement data mining strategies significantly profit and hold a competitive advantage over those that do not. The results infer that data mining application in deviation detection, predictive modelling, link analysis, and database segmentation are strong contributors to competitive advantages in the banking realm in the country. They reinforce the prevailing understanding that the application of data mining in banks leads to notable competitive benefits and improved performance.

This study has found that many commercial banks exploit Porter's (1998) generic strategies to not only survive but also thrive in the banking industry. According to Porter (1998), cost leadership allows firms to minimize the costs incurred in providing value (product or service) to customers. Porter also notes that the focus strategy enables firms to concentrate on niche markets and, by discerning customers' unique needs and market dynamics, develop distinctively low-cost or specific products and services for the market. Companies applying this strategy tend to develop robust brand loyalty amongst their clients and become more customer-oriented, making a particular market less appealing to the competition. Per Porter's (1998) generic strategies, commercial banks in the country apply data mining to become low-cost providers of financial services and minimize the outlays incurred in delivering value to clients. They also apply a differentiation strategy to make their financial products and services unique and outstanding relative to competitors. Besides, it emerges that a significant number of commercial banks in Kenya concentrate on particular niche markets by using database segmentation, generally, and customer profiling, CRM, customer segmentation, and targeted marketing, specifically, to understand market dynamics and satisfy specific customer needs and preferences.

While there is a clear-cut constructive and beneficial correlation between data mining application and the competitive advantage of banks, the results also recognized that various organizational, technological, and environmental factors lead to the espousal of data mining methods and tools. Banks in Kenya broadly embrace data mining techniques and tools because of the need to reduce costs and gain market intelligence. ICT advances, operational necessity, system integration, successful management of information systems, and increased competition are also key drivers shaping the implementation of data mining in the Kenyan banking industry. Considering the resultant competitive advantages of data mining application, it is a no-brainer that individual banks will continuously exploit other areas of data mining usage to realize more benefits and heighten their competitiveness in the highly turbulent banking sector. The drivers are also not pro tem or stopgap influencing factors; they will continuously persuade the organizations to apply data mining in different areas to realize competitive advantages and benefits that come with its application.

Lastly, as is the case with the application of other information systems, commercial banks grapple with various challenges in the use of data mining. Privacy concerns and data security top the list of notable challenges. The findings echo the outcomes of the existing works of Zain and Rahman (2017), which found that the general issues surrounding data mining revolve around technology, skills, complex and incomplete data, privacy, and data security. A study by Jaseena and David (2014) also discloses that data mining involves multiple phases, each with challenges concerning timeliness, scale, privacy, heterogeneity, and complexity issues. Accordingly, commercial banks in Kenya grapple with similar challenges in their quest to realize various competitive advantages from the usage of data mining. Notwithstanding these challenges, the successful application of data mining lies in overcoming them.

CHAPTER FIVE: SUMMARY, CONCLUSION, AND RECOMMENDATIONS

5.1. Introduction

This segment summarizes the research results, draws deductions, and proposes recommendations on the grounds of the survey findings. It anchors on the background information, the problem statement, the significance of the research, and the research objectives. The segment also covers the constraints of the survey and the suggestions for further research.

5.2. Summary of Findings

A census of all 38 commercial banks in Kenya was conducted, out of which 33 participated in the survey, representing an 86.84% response rate. Of the respondents, 72.7% were male, whereas 27.3% were female. This outcome implies that the organizations have both genders represented at the management level and participate in implementing information systems. While many of the participants fall in the 30 to 40 age bracket, most have worked in the organizations for between 1 and 10 years. The study also disclosed that many of the respondents, who were mainly ICT managers, had a bachelor's degree as their uppermost academic qualification. Concerning the organizations' information, many banks in Kenya have less than 1000 employees. Those with between 1001 and 2000 and 3001 and 4000 of the respondents are 15.2% and 12.1%, respectively. Most respondents also noted that their organizations started more than 20 years ago, considering they are 72.7% of the surveyed banks. Only a handful (24.2%) had been operational for between 11 and 20 years. Lastly, the study also unearthed the value of commercial banks in Kenya and found many of them have assets worth less than 30 billion Kenya shillings. While 24.2% have between 30 and 100 billion Kenya shillings, 24.2% have more than 300 billion Kenya shillings are fall in the tier-one category of banks in the country.

With reference to the first objective, which sought to investigate the extent of data mining applications, banks essentially use data mining for deviation detection, predictive modelling, link analysis, and database segmentation. Deviation detection is widely used with a mean of 4.46, followed by database segmentation at 4.35 and then link analysis at 4.28. Predictive modelling is equally used with a mean of 3.82. The specific areas in which data mining is extensively applied include money laundering detection, fraud detection, market intelligence quality control, and credit scoring and approval. The next objective aimed to establish the drivers that encouraged data mining adoption in Kenya's commercial banks. The survey ascertained the driving factors that strongly influenced the adoption of data mining to be ICT advances, cost reduction, successful implementation of information systems, and operational necessity. The third objective intended to pinpoint the challenges of adopting data mining in the banking sector. Two main challenges stand out: data security, with a mean of 3.67, and privacy concerns, with a mean of 3.58. However, regulatory risks, cost of implementation, employee resistance, and inadequate management support were minor challenges for commercial banks concerning data mining application. Lastly, the study purposed to ferret out the connection between data mining application and the competitive advantage of the banking sector in Kenya. The findings show that numerous competitive advantages result from data mining technology. Commercial banks benefit from customer focus (4.64), cost leadership (4.42), informed decision-making (4.67), comprehensive risk management (4.55), and quality products and services (4.55). Of import is that there is a positive correlation between the use of data mining and competitive advantages in the banking industry in Kenya.

5.3. Conclusion

This study found that data mining application in Kenyan banks has a positive effect and pay-off on their competitive advantage. In this regard, banks seeking to remain competitive should extensively adopt and implement data mining techniques to create an internal value that sets their business undertakings apart from the competition. By using data mining in deviation detection, predictive modelling, database segmentation, and link analysis to look for implicit, hitherto indefinite, and potentially valuable information from massive data, commercial banks learn more about their customers, increase operational efficiency, make informed decisions, retain customers, increase revenues, and develop quality products and services. The application helps banks to deliver financial services and products better than competitors. It allows firms to achieve superior margins and generate value for stakeholders. Thus, banks in Kenya comprehensively sort through vast volumes of data to identify patterns and relationships that help solve business problems. Furthermore, the study has shown that commercial banks in Kenya are well-furnished with respect to stewardship through the successful implementation of management information systems and the espousal of ICT advances. This propensity has been pivotal in their implementation of data mining tools and techniques. Nonetheless, other factors have also driven the adoption of the technology in the banking sector, such as increased competition, operational necessity, system integration, cost reduction, and high level of data control. These dynamics have led to advanced data mining use in the banking division, thanks to its resultant competitive advantages.

Moreover, the study reveals that banks in Kenya extensively use data mining in various functional areas. They include credit scoring and approval, fraud detection, due diligence analysis, customer profiling, targeted marketing, customer segmentation, customer churn analysis, and money laundering detection. Besides, the banks are taking a practical and business-oriented approach to data mining, which is reflected in how they deploy it to gain competitive advantages over their rivals. They are also considering using data mining to solve issues imminent in the financial industry, such as money laundering and fraud, which present a complex and dynamic challenge across the globe. Commercial banks recognize that these problems have potentially devastating economic, security, and social consequences. Therefore, data mining has a broad domain in almost every functional area in commercial banks, which makes it a crucial frontier in database and information systems and a promising source of competitive advantages.

5.4. Recommendations

Technological advances, such as data mining techniques, have contributed to faster, easier, and more insightful knowledge discovery. The more massive and more multifaceted the data sets are, the greater the prospects of discovering pertinent insights and precious nuggets of previously unknown information. By applying data mining, firms can exploit invaluable information in decision-making, accomplish predefined goals, and attain a competitive advantage in the marketplace. Therefore, commercial banks and other financial organizations, including microfinance and insurance companies, should implement data mining in disparate functional areas to create an internal value that gives them a competitive advantage. Using data mining to understand and extrapolate data and information can lessen the chances of fraud, reduce operational costs, create cost leadership, optimize channels, retain customers, increase revenues, heighten operational efficiency, and help develop unique and quality products and services. Financial institutions should prioritize investments in technology and embrace a test-to-learn mentality to explore and realize untapped opportunities and competitive advantages. They must consider the resultant benefits from data mining as a foothold towards solving problems, mitigating risks, and seizing new opportunities.

5.5.Limitations of the Study

The researcher of the present study grappled with more than a few limitations during the research process. The major limitation concerns the nature of the banking sector. Commercial banks are very sensitive and secretive with their information. The respondents in the commercial banks that did not participate asked to be excluded from the process owing to their organization's policy regarding confidentiality of information. Thus, there were notable difficulties soliciting input from the respondents. With respect to the above, it is also possible that some respondents could be biased or dishonest, considering the non-disclosure policies in their organizations. Furthermore, the time allotted for the research was a constraining factor. Many respondents seemed busy and faled to show the urgency to fill out the questionnaires. This situation prompted the researcher to regularly remind and visit the head offices and branches several times to check whether the questionnaire had been completed. Lastly, the research scope is also constrained since it only investigated the application of data mining and the competitive advantage in the Kenyan banking sector. As such, it cannot be generalizable for the broader financial services industry.

5.6.Suggestions for Further Research

Further research can be conducted to determine the correlation between the use of data mining and the competitive advantage of firms in the broader financial services industry, including microfinance, SACCOs, and insurance firms in Kenya. It is imperative to investigate the conception of data mining and its application to improve the performance of some of the fundamental business processes in the banking industry. It is also worthwhile to investgate the connection between data mining application and the performance of financial institutions. Moreover, further research can also focus on the data mining methods used in different banking domains and how the methods and practices make the decision-making process productive.

REFERENCES

Afolabi, I., & Adegoke, F. (2014). Analysis of customer satisfaction for competitive advantage using clustering and association rules. *International Journal of Computer Science and Engineering*, *3*, 141-150.

Ahmad, M., Rashid, T., & Mishra, D. K. (2014, March). Impact of wireless technology on future of big-data industry. In *2014 Conference on IT in Business, Industry and Government (CSIBIG)* (pp. 1-4). IEEE.

Ahmed, V., Aziz, Z., Tezel, A., & Riaz, Z. (2018). Challenges and drivers for data mining in the AEC sector. *Engineering, Construction and Architectural Management*, 1-31. https://usir.salford.ac.uk/id/eprint/47294/1/PDF_Proof%20(002).pdf.

Al Kalaf, A. (2011). *Announcement of the Basel III Accord and SAMA Plans for its Implementation of Basel II and III in 2011*, Saudi Arabian Monetary Agency (SAMA).

Al Ruwaili, N. F., Bright, D., & Alhameed, A. (2013). To what extent talent management in Saudi Arabian banks. *International Journal of Advances in Management Science*, *2*(1), 9-15.

Almoqren, N., & Altayar, M. (2016). Identify challenge factors in the adoption of big data mining technology In Saudi Banks. *International Journal of Management and Applied Science*, *2*(7), 33-39.

Askool, S., & Nakata, K. (2012, June). Investigation into the adoption intention of social CRM in Saudi banks. In *International Conference on Information Society (i-Society 2012)* (pp. 402-407). IEEE.

Bal, M., Bal, Y., & Demirhan, A. (2011). Creating competitive advantage by using data mining technique as an innovative method for decision making process in business. *International Journal of Online Marketing (IJOM)*, *1*(3), 38-45.

Barney, J. B. (1986), "Strategic factor markets: Expectations, luck, and business strategy", *Management Science*, *32*(10), 1231-1242.

Barney, J. B. (1991), "Firm resources and sustained competitive advantage", *Journal of Management*, *17*(1), 99-121.

Bhambri, V. (2012). Role of data mining in banking sector. *International Indexed & Referred Research Journal*, *3*, 70-71.

Brezany, P., Hofer, J., Tjoa, A. M., & Woehrer, A. (2003, October). Gridminer: An infrastructure for data mining on computational grids. In *Proceedings of Australian Partnership for Advanced Computing Conference (APAC)*.

Brusilovsky, D., & Brusilovskiy, E. (2008). Data mining: The means to competitive advantage. *Business Intelligence solutions*.

Çaliş, A., Boyaci, A., & Baynal, K. (2015, March). Data mining application in banking sector with clustering and classification methods. In *2015 International Conference on Industrial Engineering and Operations Management (IEOM)* (pp. 1-8). IEEE.

CBK. (2022). *Bank supervision | CBK*. Centralbank.go.ke. Retrieved 20 June 2022, from https://www.centralbank.go.ke/bank-supervision/.

Chan, J. O. (2013). An architecture for big data analytics. *Communications of the IIMA*, *13*(2), 1.

Chitra, K. & Subashini, B. (2013). Data mining techniques and its applications in banking sector. *International Journal of Emerging Technology and Advanced Engineering* 3 (8), 219-226.

Choudhary, A. N., Narayanan, R., & Zhang, K. (2010, March). High-end analytics and data mining for sustainable competitive advantage. In *International Conference on Information Systems, Technology and Management* (pp. 8-15). Springer, Berlin, Heidelberg.

Coumaros, J., Buvat, J., Auliard, O., Roys, S., Kvj, S., Chretien, L., & Clerk, V. (2014). Big data alchemy: How can banks maximize the value of their customer data? *Capgemini Consulting*.

de Lima, D. V. M. (2011). Research design: A contribution to the author. *Online Brazilian Journal of Nursing*, *10*(2).

Dhaouadi, A., Bousselmi, K., Gammoudi, M. M., Monnet, S., & Hammoudi, S. (2022). Data warehousing process modeling from classical approaches to new trends: Main features and comparisons. *Data*, *7*(8), 1-38.

Gitonga, K. M. (2015). *Strategies for gaining competitive advantage of commercial banks listed at the NSE in Kenya* (Doctoral dissertation).

Hamel, G. and Prahalad, C. (1996), Competing for the future. Harvard Business School Press,Paperback edition, Boston (Massachusetts).

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques*. Elsevier.

Haralayya, D., & Aithal, P. S. (2021). Implications of banking sector on economic development in India. *George Washington International Law Review*, *7*(1), 631-642.

Hart M., Esat F., Rocha M., & Khatieb, Z. (2007). Introducing students to business intelligence: Acceptance and perceptions of OLAP software. *Informing Science, 4,* 105-123.

Hasheminejad, S. M. H., & Khorrami, M. (2018). Data mining techniques for analyzing bank customers: A survey. *Intelligent Decision Technologies*, *12*(3), 303-321.

Huang T.C.K., Liu C. C., & Chang D. C. (2012). An empirical investigation of factors influencing the adoption of data mining tools. *International Journal of Information Management*, *32*, 257-270.

Jaseena, K. U., & David, J. M. (2014). Issues, challenges, and solutions: Big data mining. *CS & IT-CSCP*, *4*(13), 131-140.

Jayasree, V., & Balan, R. V. S. (2013). A review on data mining in banking sector. *American Journal of Applied Sciences*, *10*(10), 1160 1165. doi:10.3844/ajassp.2013.1160.1165.

Kagechu, I. (2018). *Impact of big data on organizational performance: A case of listed commercial banks in Kenya* (Doctoral dissertation, United States International University-Africa).

Khan, M. A., Zedan, H., & Al Ghadeer, A. A. (2011). Influential factors in adoption of information technology in financial business sector in Saudi Arabia. *Canadian Journal of Pure and Applied Sciences*, *5*, 1413-1417.

Khder, M. A., Abu-Alsondos, I. A., & Bahar, A. Y. (2021). The impact of implementing data mining in business intelligence. *International Journal of Entrepreneurship*, *25*, 1-7.

Kothari, C. R. (2006). *Research methodology methods and techniques*. New Age International (P) Ltd., Publishers.

Kotler, P. (2000). *Marketing management*, 10th Edition. Prentice Hall, New York.

Kugel, R. (2013). Finance can get a big advantage from big data. *Ventana Research2013*.

Liu, Y. (2003). Case study of data mining application in banking industry. ICEB Proceedings (Singapore), 77. https://aisel.aisnet.org/iceb2003/77/.

Machiraju, H. R. (2008). *Modern commercial banking*. New Age International Ltd.

Marikyan, D. & Papagiannidis, S. (2022). Technology acceptance model: A review. In S. Papagiannidis (Ed), *Theory Hub Book*. http://open.ncl.ac.uk

Mastercard Foundation (2018, April). *Can big data shape financial services in East Africa?* https://www.financedigitalafrica.org/2018/03/19/can-big-data-shape-financial-services-in-east-africa/.

Mian, T. S., & Ghabban, F. (2022). Competitive advantage: A study of Saudi SMEs to adopt data mining for effective decision making. *Journal of Data Analysis and Information Processing*, *10*(3), 155-169.

Miller, M., & Nyauncho, E. (2015). *Effective data mining and analysis for SME banking*. Financial Sector Deeping. http://fsd-circle.s3.amazonaws.com/wp-content/uploads/2015/05/30094431/15-07-20_GrowthCap_technical_note_Effective_data_mining_and_analysis_for_SME_banking.pdf.

Moin, K. I., & Ahmed, D. Q. B. (2012). Use of data mining in banking. *International Journal of Engineering Research and Applications*, *2*(2), 738-742.

Mugane, G. M. (2018). *Big data analytics and competitive advantage of commercial banks and fintech companies in Kenya* (Doctoral dissertation, University of Nairobi).

Nunan, D., & Di Domenico, M. (2013). Market research and the ethics of big data. *International Journal of Market Research*, *55*(4), 505-520.

Oladimeji, O. O., & Oladimeji, O. O. (2020). Exploring data mining research in West Africa: A bibliometric analysis. *SLIS Connecting*, *9*(2), 9. https://doi.org/10.18785/slis.0902.05.

Pascu, A. I. (2018). Data mining. Concepts and applications in banking sector. *Annals of Constantin Brancusi University of Targu-Jiu. Economy Series*, (1). https://ideas.repec.org/a/cbu/jrnlec/y2018v1p159-166.html.

Pechenizkiy, M., Puuronen, S., & Tsymbal, A. (2005, August). Competitive advantage from Data Mining: Some lessons learnt in the Information Systems field. In *16th International Workshop on Database and Expert Systems Applications (DEXA'05)* (pp. 733-737). IEEE.

Porter, M. (1998). *Competitive advantage*. The Free Press, New York.

Pulakkazhy, S., & Balan, R. S. (2013). Data mining in banking and its applications – A review. *Journal of Computer Science*, *9*(10), 1252-1259.

PwC. (2022). *Banking*. Retrieved 20 June 2022, from https://www.pwc.com/ke/en/industries/banking.html.

Ragavi, R., Srinithi, B., & Sofia, V. A. (2018). Data mining issues and challenges: A review. *International Journal of Advanced Research in Computer and Communication Engineering*, (7), 4-7. 118-121. https://doi.org/10.17148/IJARCCE.2018.71125.

Ranjbarfard, M., & Ahmadi, S. (2020). A Study of data requirements for data mining applications in banking. *Journal of Digital Information Management, 18*(3), 109-117.

Renushe, H. N., Rasal, P. R., & Desai, A. S. (2012). Data mining practices for effective investigation of crime. *International Journal of Computer Applications in Technology, 3*, 865-870.

Saini, P., & Sindhu, J. (2014). Role of commercial bank in the economic development of India. *International Journal of Engineering and Management Research*, *4*(1), 27-31.

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and business cycle* (Vol. 46). Cambridge: Harvard University Press.

Śledzik, K. (2013). Schumpeter's view on innovation and entrepreneurship. *Management Trends in Theory and Practice, (ed.) Stefan Hittmar, Faculty of Management Science and Informatics, University of Zilina & Institute of Management by University of Zilina*. https://doi.org/10.2139/ssrn.2257783.

Srivastava, S. K., Chandra, B., & Srivastava, P. (2019). The impact of knowledge management and data mining on CRM in the service industry. In *Nanoelectronics, Circuits and Communication Systems* (pp. 37-52). Springer, Singapore.

Taghva, M., Bamakan, S., & Toufani, S. (2011). A data mining method for service marketing: A case study of banking industry. *Management Science Letters*, *1*(3), 253-262.

Tannenbaum, C., Greaves, L., & Graham, I. D. (2016). Why sex and gender matter in implementation research. *BMC Medical Research Methodology, 16*(1), 1-9.

Verma, D., & Nashine, R. (2012). Data mining: Next generation challenges and future directions. *International Journal of Modeling and Optimization*, *2*(5), 603.

Wang, W. C., Lin, C. H., & Chu, Y. C. (2011). Types of competitive advantage and analysis. *International Journal of Business and Management*, *6*(5), 100-104. https://doi.org/:10.5539/ijbm.v6n5p100.

Weiss, G. M. (2009). Data mining in the real world: Experiences, challenges, and recommendations. In *DMIN* (pp. 124-130). https://storm.cis.fordham.edu/~gweiss/papers/dmin09-real-world.pdf.

Widuri, R., & Sutanto, J. E. (2019, January). Differentiation strategy and market competition as determinants of earnings management. In *International Conference on Tourism,* *Economics, Accounting, Management, and Social Science (TEAMS 2018)* (pp. 171-176). Atlantis Press. https://doi.org/10.2991/teams-18.2019.30.

Zain, M. S. I. M., & Rahman, S. A. (2017). Challenges of applying data mining in knowledge management towards organization. *International Journal of Academic Research in Business and Social Sciences*, *7*(12), 405-412.

Zakirov, D. & Momtselidze, N. (2015). Application of data mining in the banking sector. *Journal of Technical Science and Technologies*, *4*(1), 13-16. https://doi.org/10.31578/.v4i1.68.

Zide, O., & Jokonya, O. (2022). Factors affecting the adoption of data management as a service (DMaaS) in small and medium enterprises (SMEs). *Procedia Computer Science*, *196*, 340-347.

APPENDICES

Appendix I: Letter of Introduction

UNIVERSITY OF NAIROBI
FACULTY OF BUSINESS AND MANAGEMENT SCIENCES
OFFICE OF THE DEAN

Telegrams: "Varsity",
Telephone: 020 491 0000
VOIP: 9007/9008
Mobile: 254-724-200311

P.O. Box 30197-00100, G.P.O.
Nairobi, Kenya
Email: fob-graduatestudents@uonbi.ac.ke
Website: business.uonbi.ac.ke

Our Ref: **D61/18862/2019** November 09, 2022

National Commission for Science, Technology and Innovation
NACOSTI Headquarters
Upper Kabete, Off Waiyaki Way
P. O. Box 30623- 00100
NAIROBI

RE: INTRODUCTION LETTER: DUNCAN NDUNG'U NG'ANG'A

The above named is a registered Masters of Business Administration candidate at the University of Nairobi, Faculty of Business and Management Sciences. He is conducting research on ***"Data Mining Application and Competitive Advantage of Commercial Banks in Kenya."***

The purpose of this letter is to kindly request you to assist and facilitate the student with necessary data which forms an integral part of the Project.

The information and data required is needed for academic purposes only and will be treated in **Strict-Confidence**.

Your co-operation will be highly appreciated.

Dean's Office
University of Nairobi
Faculty of Business
and Management Science
P O Box 30197-00100, Nairobi

PROF. JAMES NJIHIA
DEAN, FACULTY OF BUSINESS AND MANAGEMENT SCIENCES

JN/jkm

Appendix II: Research Questionnaire

QUESTIONNAIRE FOR BANK ICT MANAGERS

PROJECT TITLE: DATA MINING APPLICATION AND COMPETITIVE ADVANTAGE OF COMMERCIAL BANKS IN KENYA

Instructions

This questionnaire is for research on Data Mining Application and Competitive Advantage of Commercial Banks in Kenya. It is part a post-graduate research project presented in partial fulfilment of Master of Business Administration degree, Faculty of Business and Management Sciences, University of Nairobi.

Kindly respond to the following questions and where applicable, mark the relevant box with a tick (✓).

SECTION A: DEMOGRAPHIC INFORMATION

1. What is your position within the organization? ____________________________________

2. What is your gender?

   Male **[ ]** Female **[ ]**

3. In which age group do you fall?

   [ ] 18 – 25 years **[ ]** 26 – 30 years **[ ]** 31 – 35 years

   [ ] 36 – 40 years **[ ]** 41 – 50 years **[ ]** Above 50 years

4. How long have you worked in the organization?

   [ ] Less than 1 year **[ ]** 1 – 5 years **[ ]** 6 – 10 years

   [ ] 11 – 15 years **[ ]** 16 – 20 years **[ ]** More than 20 years

5. What is your highest level of education?

   [ ] Diploma **[ ]** Bachelor's Degree

   [ ] Master's Degree **[ ]** Doctorate Degree

   [ ] Other (specify)……………………………………………..

6. How many employees work in the organization?

[ ] Less than 1,000 **[ ]** 1,001 – 2,000 **[ ]** 2,001 – 3,000

[ ] 3,001 – 4,000 **[ ]** 4,001 – 5,000 **[ ]** 5,001 – 6,000

[ ] 6,001 – 7,000 **[ ]** Over 7,000

7. How long has the organization been in business?

   [ ] Less than 5 years **[ ]** 5 – 10 years

   [ ] 11 – 20 years **[ ]** More than 20 years

8. What is the value of the organization in terms of total assets in Kenya shillings?

   [ ] Less than 30 billion **[ ]** 30 – 100 billion

   [ ] 100 – 300 billion **[ ]** More than 300 billion

SECTION B: DATA MINING APPLICATION IN YOUR ORGANIZATION

9. Kindly indicate with a (✓) the extent to which data mining is used in the organization for each of the following purposes. Use the scale of 1 to 5, where;

   (1) No Extent, (2) Little Extent, (3) Moderate Extent,

   (4) Large Extent, and (5) Very Large Extent.

Applications	**No Extent (1)**	**Little Extent (2)**	**Moderate Extent (3)**	**Large Extent (4)**	**Very Large Extent (5)**
Credit scoring & approval					
Fraud detection					
Quality control					
Money laundering detection					
Stocks prediction					
Revenue prediction					
Loan default prediction					
Bank/product failure estimation					
Telemarketing & electronic marketing					
Customer churn analysis					
Telemarketing outcomes prediction					
Customer relationship management					
Customer segmentation					
Cross-selling products					
Customer profiling					
Targeted marketing					
Identifying customer bank behavior					
Customer sentiment analysis					
Due diligence analysis					
Other (specify and tick accordingly)					

SECTION C: DRIVING FACTORS THAT INFLUENCE THE ADOPTION OF DATA MINING

10. Please rate the extent to which the following factors were the drivers in the adoption of data mining in the organization. Use the scale of 1 to 5, where;

(1) No Extent, (2) Little Extent, (3) Moderate Extent,
(4) Large Extent, and (5) Very Large Extent.

Driving Factors	**No Extent (1)**	**Little Extent (2)**	**Moderate Extent (3)**	**Large Extent (4)**	**Very Large Extent (5)**
Successful management of information systems					
ICT advances					
Ability to interpret unstructured data					
Level of data control					
System integration					
Increased competition					/
Operational necessity					
Cost certainty and cost reduction					
Market intelligence					
Query complexity					
Other (specify and tick accordingly)					

SECTION D: CHALLENGES FACED IN THE APPLICATION OF DATA MINING

11. Please indicate the extent to which the organization has faced each of the following challenges in the application of data mining. Use the scale of 1 to 5, where;

(1) No Extent, (2) Little Extent, (3) Moderate Extent, (4) Large Extent, and (5) Very Large Extent.

Challenges	**No Extent (1)**	**Little Extent (2)**	**Moderate Extent (3)**	**Large Extent (4)**	**Very Large Extent (5)**
Data security					
Employee resistance					
Privacy concerns					
Poor quality data					
Lack of required infrastructure					
Noisy and incomplete data					
Regulatory risks					
Shortage of skills and technical knowhow					
Inadequate management support					
Cost of implementation					
Misalignment of IT and corporate strategy					
Other (specify and tick accordingly)					

SECTION E: COMPETITIVE ADVANTAGE RESULTING FROM THE APPLICATION OF DATA MINING

12. Kindly indicate the extent to which the organization has realized each of the following of advantages as a result of applying data mining. Use the scale of 1 to 5, where;
    (1) No Extent, (2) Little Extent, (3) Moderate Extent,
    (4) Large Extent, and (5) Very Large Extent.

Competitive Advantage	**No Extent (1)**	**Little Extent (2)**	**Moderate Extent (3)**	**Large Extent (4)**	**Very Large Extent (5)**
Customer focus					
Cost leadership					
Quality products and services					
Revenue					
Customer retention					
Informed decision making					
Operational efficiency					
Comprehensive risk management					
Channel optimization					
Market share					
Other (specify and tick accordingly)					

THANK YOU FOR YOUR PARTICIPATION